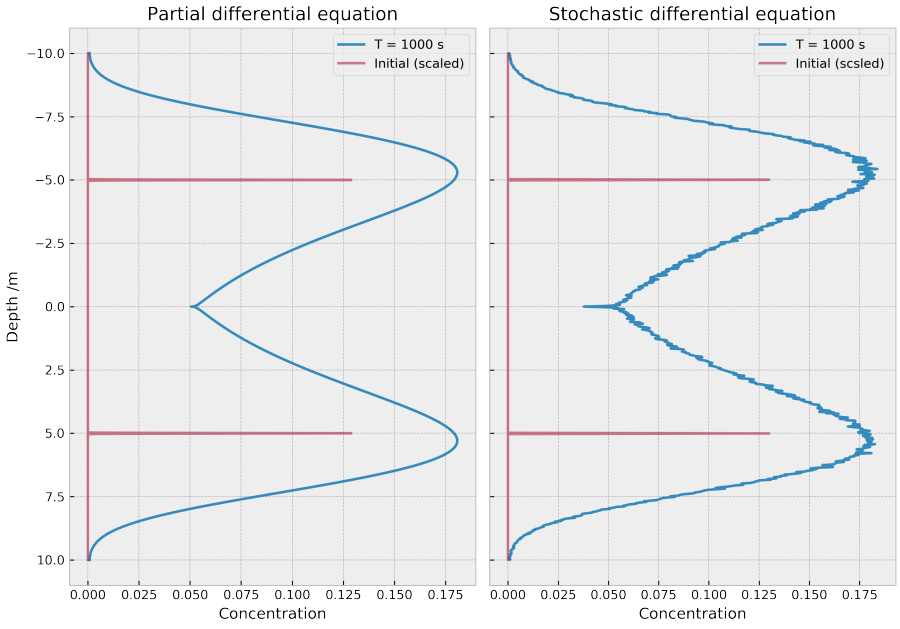
Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Physics

Mauhing Yip

# Boundary artifact analysis in diffusion modelling with stochastic differential equations

Master's thesis in Technical physics
Supervisor: Tor Nordam
June 2019

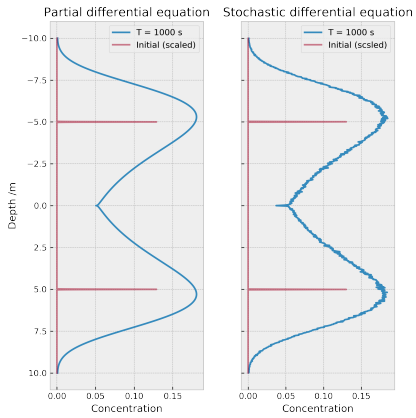



Diffusion process: comparison between partial differential equation and stochastic differential equation

NTNU
Norwegian University of
Science and Technology

Mauhing Yip

# Boundary artifact analysis in diffusion modelling with stochastic differential equations




Master's thesis in Technical physics
Supervisor: Tor Nordam
June 2019

Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Physics

NTNU
Norwegian University of
Science and Technology

**Abstract**

In many research areas, it is common to model advection-diffusion problems with Lagrangian particle methods. This is the same as solving a stochastic differential equation, with drift and diffusion coefficients derived from the advection-diffusion equation. But there is also a necessary condition for the particle method to be equivalent to the Eulerian advection-diffusion equation, is that it satisfies the well-mixed condition (Thomson, 1987), which says that if particles are well mixed, they have to stay well mixed later on. This is just a statement with respect to second law of thermodynamics, which is entropy. A commonly used implementation of reflecting boundary conditions for particle methods is analysed. We find that in some cases, this reflecting scheme will give rise to oscillations in concentration close to the boundary, which we call the boundary artifact.

We analyse the reflection scheme in the Lagrangian model, and compare it to Neumann boundary conditions in the Eulerian model. We find that if the diffusivity has a non-zero derivative at the boundary, this violates one of the conditions for equivalence with the advection-diffusion equation, which is that the drift coefficient in the SDE must be Lipschitz continuous. This seems to be the origin of the boundary artifact. We analyse the artifact further, and describe two different types of boundary artifact.

We suggest different approaches to dealing with the problem, and find that the problem can in practice be handled by adjusting the diffusivity close to the boundary. Support and motivation for such a change is found in the concept of the "unresolved basal layer" (Wilson & Flesch, 1993), which is a pragmatic idea stating that closer than some distance from the boundary, we simply cannot know the details of the turbulent motion.

# Preface

This thesis is written during my final semester of the MSc programme in Physics and Mathematics at NTNU, and is a continuation of work I did during my Specialization project in the autumn of 2018. During this last semester, I have also spent one month as a guest to Dr. Colin Cotter at Imperial College in London. This visit was funded by HPC-Europa3 and SINTEF Ocean.

I feel that I have an obligation to inform the reader that I did not have any experience with stochastic differential equation (SDE) and Python code before starting my Specialization project. However, my experience with ordinary differential equations and C++ has helped me a lot with this topic. Therefore, if the reader finds any errors, please contact me. I will be very grateful to receive any corrections and suggestions. As English is not my native language, the reader may find some grammatical errors, and again, I am very happy to receive corrections and suggestions. My email address is mauhy@stud.ntnu.no or mauhingyip@gmail.com

# Acknowledgments

Here, I would like to especially thank my supervisor, Dr. Tor Nordam, for being incredibly patient to help me in academically, linguistically and mentally. Beside all these, he has also spent a lot of time to teach me about the Linux operating system and Python. When there was a funding crisis to the one month research visit to England, he actively helped me to find a solution. He also went all the way from Norway to England to visit me and Dr. Collin Cotter and engaged to help me with full passion. In addition, he taught me human social interaction, also known as social norms. Perhaps, this is the most important knowledge that I was missing. I am grateful that I have a very good mentor.

During the London visit, Dr. Collin Cotter patiently analyzed the boundary artifact with me. He inspired me to perform statistical sampling of particles crossing the boundary during simulations, which became a very important of this thesis (see Section 3.5 for details).

This work has been performed with support from the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme. In particular, I gratefully

acknowledge the support of Dr. Colin Cotter of the Department of Mathematics at Imperial College, and the computer resources and technical support provided by the EPCC, at the University of Edinburgh. I especially want to thank the staff at EPCC who gave me a lot suggestions and taught me how to use a supercomputer. In additional, I want to thank SINTEF Ocean their help in funding my research stay in the UK.

Lastly, I want to thank my step farther (Saichun Wong), mother (Wanha Lau), step bothers (Kingyu Wong and Kingho Wong), aunt (Kuanseen Yip), cousin (Meelyn Leong) and all my family and friends for being encourage and support my education all ways long. Moreover, I would also want to thank Norwegian education system to give huge economical support to all students and make cross cities education possible. I also want to thank two of my friends and fellow students Venstein Dahl and Steinar Gundersen for showing their care to my thesis and formatting for this thesis. To my girlfriend, Carmen Chan, thank you for the constant supports and the gentle love. You gave me a lot courage.

# List of Symbols

### Number Sets (Space domain Time domain)

| | |
|---|---|
| $\mathbb{D}$ | Space domain |
| $\mathbb{T}$ | Time domain |
| $\mathbb{E}$ | $\mathbb{E} \subset \mathbb{D}$. This set is designed in such way that to tell whether a particle is near boundary or not. $\mathbb{E}$ stands for Edge. |
| $\mathbb{B}$ | $\mathbb{B} := \mathbb{D}/\mathbb{E}$ This set is designed in such way that to tell whether a particle is near boundary or not. $\mathbb{B}$ stands for Body. |

### Parameter

| | |
|---|---|
| $N_p$ | Number of particles launch in simulation |
| $\Delta t$ | Time step in simulation |
| $T_{end}$ | Simulation time, this is not wall-time |
| $\Delta W$ | Wiener process, $\Delta W \sim \mathcal{N}(0, \Delta t)$ |
| $H$ | The length of space in simulation |
| $H_h$ | The interested length from boundary. It is useful to zoom the boundary artifact. |
| $N_b$ | Number of bins in between $[0, H_h]$ to make a histogram |
| $\Delta z$ | Space step, use only in partial differential equation and numerical derivative in Section 6.1. |
| $\omega$ | Advection |

**Other Symbols**

FTCS Forward time central space. It is widely used numerical partial differential scheme.

PDE Partial differential equation

SDE Stochastic differential equation

PDF Probability density function

PRS Perfect reflection scheme. This is a common reflection boundary condition scheme for SDE to imitate the NBC in PDE (see Eq. (3.2)

NBC Neumann boundary condition. A common boundary condition scheme to converse the quantity in the domain of interest (see Eq. (2.32)).

WMC Well-mixed condition (see Section 2.10)

UBL unresolved basal layer (see Wilson and Flesch (1993))

# Contents

# Chapter 1

# Introduction

## 1.1 Motivation

Particle tracking in fluids has been of interest in many different academic areas. For example, the diffusion of plankton in the ocean, which is the base of the largest food-web in the world, is crucial for most life forms in the marine ecosystem. Moreover, plankton is also important for global oxygen production, and the marine economy. Another example is that of sediment transport (Singer, Atkinson, Manley, & McLaren, 2008). Some industrial activity would impact the river/marine ecosystem. Some of the pollutants have been discharged to the river and mixed within the sediments. Those sediments may typically contain heavy metal, and organic compounds and may cause some impairments including poor water quality, and loss of habitat. For more applications please refer to Gräwe, Deleersnijder, Shah, and Heemink (2012); van Sebille et al. (2018).

Marine transport modelling is mainly divided into two different numerical approaches, that is the Eulerian model and Lagrangian model. The Eulerian model focuses on how the local vector field, which is the flow direction, changes with time, whereas the Lagrangian model focuses on how individual particles move with the fluid. Mathematically speaking, an Eulerian model is $\vec{v}(x, y, z, t)$ and, Lagrangian model is $\vec{v}(x(t), y(t), z(t))$, where $\vec{v}$ is the flow velocity, $x, y, z$ are position coordinates, and $t$ is time of the particle. In this thesis, we will mainly focus on stochastic differential equations (SDEs), in the Lagrangian model. In van Sebille et al. (2018), they show that Lagrangian ocean analysis is a powerful way to analyse the output of ocean circulation

models, and it is widely used in ocean modelling. For more detail please refer to van Sebille et al. (2018); Yunus and Cimbala (2006, chapter 4).

In an Eulerian model, it is very common to use a finite difference method or a finite volume method to calculate concentration of particles on a discrete grid. But these methods do not exhibit the individual path of each single particle. If some of these particles have different properties, it will be difficult to implement in an Eulerian model since it does not focus on individual particles but the field. In contrast, a Lagrangian model based on stochastic differential equations can be easily implemented even if some of these particles have different properties.

A necessary condition for Eulerian consistency is that, a particle tracking scheme has to satisfy a condition called the well-mixed condition. This was introduced by Thomson (1987). In short, well-mixed condition is just a statement with respect to entropy. If all the particles are already well mixed, they should stay well mixed in pure diffusion environment, which means no advection. It turns out, Eulerian models can satisfy the well mixed condition without problem, but in a Lagrangian model, a common scheme used to implement reflecting boundary conditions can lead to an artifact, where one or more rapid oscillations occur in the concentration field close to the boundary. This is hereafter referred to as the boundary artifact. Because of this boundary artifact, Lagrangian models may fail to satisfy the well-mixed condition. The purpose of this thesis is to find out the cause of the boundary artifact, analyze it and give some suggestions to eliminate or reduce boundary artifact in Lagrangian models.

## 1.2 Outline of this thesis

The thesis is divided in two parts, part one is to understand the boundary artifact and part two is to give some suggested solutions.

In part one, in order to analyze the boundary artifact, we need some theory. In Chapter 2, we introduce background knowledge of stochastic differential equations (SDE) and derive two crucial coefficients of SDE from the advection-diffusion equation based on physics. We will also introduce the well-mixed condition and explain why a good particle tracking scheme has to satisfy this. In Chapter 3, we will demonstrate the boundary artifact of SDE schemes and explain why these schemes do not satisfy the well-mixed condition. We will also give a comprehensive analysis of why this boundary

artifact occurs by using the knowledge from Chapter 2. We will also show that in order for the common reflecting boundary scheme in Lagrangian to not cause boundary artifact, there is a special requirement of the diffusivity. If one makes the right change to the diffusivity, the boundary artifact problem is gone. We will also discuss the legitimization of changing the diffusivity near the boundary.

In part two, in Chapter 4, we will be looking for analytical solutions to satisfy the well-mixed condition. In Chapter 5, we will be looking for more practical solutions in real-life applications with minimum change of diffusivity. Furthermore, we will make some suggestions for how to satisfy the well-mixed condition. In Chapter 6, we will try to approach the problem without changing the diffusivity, but we still try to use the knowledge that has been obtained from Chapter 3. In Chapter 7, we will briefly summarize the most important results of this thesis, and also discuss in what circumstances the boundary artifact is insignificant.

For all simulation of this thesis, the configuration is specified. I hope it should be very clear for the reader. But in case I have miss something. The reader can always goes by to my GitHub site. My GitHub site is in the Appendix Chapter D. In my code, all the configuration of simulation should be very clear for the reader because I always specific my configuration just before the simulation.

# Part I

# Understanding the problem

# Chapter 2

# Theory

In this chapter, we will guide the reader through from the partial differential equation of diffusion processes to a stochastic differential equation of diffusion processes.

Throughout this thesis, we consider a one-dimensional system, which represents vertical water column ($z$ axis) in the ocean. Depth is positive downwards, and the water surface is located at $z = 0$.

## 2.1 Diffusion processes

As Newton's second law is a fundamental building block for the classical mechanic, one may start there to establish a differential equation. However, using Newton's second law to estimate diffusion of particle clouds in fluids is not practical due to the interacting potential between each particle, which will cause a gigantic calculation. Therefore, we may try to use another method to approach the problem. Assuming we can obtain vertical diffusion profile by measuring time average turbulence in the ocean, then, by fluid dynamics, we can solve the advection-diffusion Eq. (2.1) numerically in the Eulerian description.

$$\partial_t C = -\partial_z(\omega C - K\partial_z C) \tag{2.1}$$

where $C, \omega, K, z, t$ are concentration, advection, diffusivity, vertical position and time correspondingly. Concentration is an intensive and macroscopic physical quantity, and it is completely deterministic. However, diffusivity is a statistical concept. If we consider the trajectory of a single particle, there is some randomness in it. But each individual particle contributes to

the concentration. Here I would like to quote from a stochastic differential equations book:

> "If we allow for some randomness in some of the coefficients of a differential equation we often obtain a more realistic mathematical model of the situation." (Øksendal, 2003)

Therefore, We would like to establish a model that focuses on individual particles with some random displacement. Gräwe (2011) presents a basic idea that one can obtain a stochastic differential equation (SDE) from Partial differential equation (PDE) Eq. (2.1) by using Fokker-Planck equation. What is a Fokker-Planck equation? You may ask. In the next section, we will have a brief introduction of the Fokker-Planck equation.

## 2.2 Fokker-Planck equation

When we approach a physical problem, often there are two treatments: macroscopic treatment and microscopic treatment. The macroscopic treatment is used when the fluctuation to our system is negligible, for example, the gas pressure. The microscopic treatment is used when we want to look at the subsystem of the total system, for example, the velocity of each individual gas particle. It is especially interesting when the total system is built from many identical subsystems. For example, for a non-ideal gas in a fixed volume, each gas molecule is our subsystem. However, as we have mentioned above, such physical systems usually contain a tremendous number of molecules and calculating the forces between each another needs extreme computing capacity and often is not practical. Besides, the exact initial condition in microscopic treatment is hard to obtain or create.

It seems both of the treatments do not fit our problem, because we do not want a macroscopic solution and calculating forces between each molecule is completely hopeless task. What we need is something in between. It is called stochastic treatment. In this treatment, our macroscopic variable, for example position $z(t)$, is no longer deterministic, instead, it depends on a probability density function $\rho$, where $\int_a^b \rho(z)dz$ means the probability to find the variable $z$ in the interval between $a$ and $b$. Moreover, the expectation value is defined in such a way: $\langle z \rangle := \int_{-\infty}^{\infty} z\rho(z)dz$. The PDE describing the development of the concentration can be derived from microscopic treatment with a large ensemble and rigorous mathematics derivation. Interestingly, it

can also be derived from macroscopic treatment with heuristic arguments, for example, equipartition theorem from statistical mechanics. Such an equation describes the time evolution of probability density function and is called the Fokker-Planck equation (Risken, 1996).

In our case, we will try to find the probability density function from our macroscopic treatment, which is our PDE Eq. (2.1). Intuitively, the concentration field should scale with the probability density function because if we mark a particular particle and let it diffuse, there is a higher probability of finding this particular particle in a high concentration region than in a low concentration region. Therefore, if we normalize the concentration field $C(z,t)$ in such way:

$$\rho(z,t) := \frac{C(z,t)}{\int_{-\infty}^{\infty} C(z,t)dz}, \tag{2.2}$$

then $\rho(z,t)$ is actually our probability density function. If we divide the equation Eq. (2.1) with the constant $\int_{-\infty}^{\infty} C(z,t)dz$, it is actually a Fokker-Planck equation because it describes how the probability density function evolves in time. When the Fokker-Planck equation is known, one can try to solve it by a stochastic differential equation, see Arnold (1974); Risken (1996).

Before we go any further, we want to show an example of how an SDE and a PDE can describe the same situation. The Fig. 2.1 shows the results of two different models with constant diffusivity. The red line is the initial condition, and the green line is the concentration field that is two hours later calculated by the PDE given by Eq. (2.1). The blue area is an initial condition of 2000000 particles, where the particle density is calculated by counting the particles in 100 bins, and dividing by the total number of particles. The purple area is using the same setting as blue area, but two hours later and calculated by the SDE which will be introduced next section.

## 2.3 Stochastic differential equation (SDE)

According to Gräwe et al. (2012), a stochastic differential equation which is equivalent to our Fokker-Planck Eq. (2.1) is

$$dZ(t) = (w + \partial_z K_z(z))dt + \sqrt{2K_z(z)}dW(t), \tag{2.3}$$

where $Z(t)$ is the position of a particle, $w$ is advection, $t$ is time started

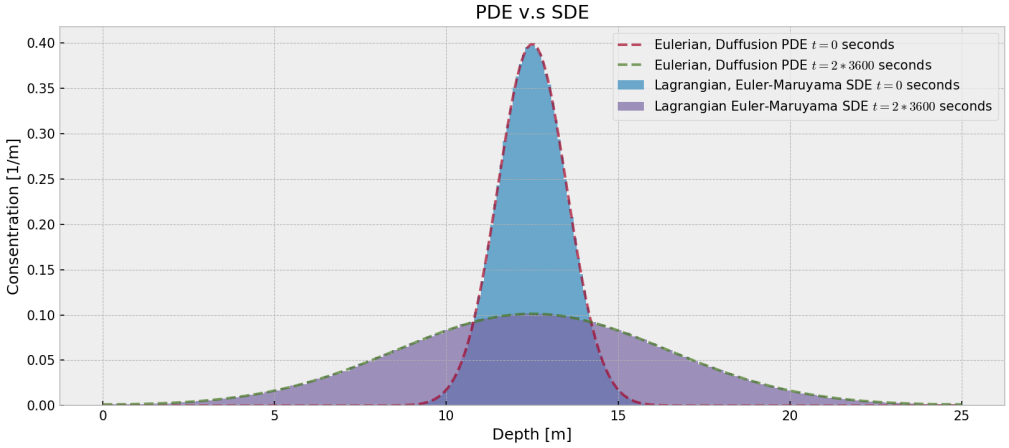



Figure 2.1: The red line is the initial condition and the green line is concentration field that is two hours later calculated by PDE Eq. (2.1). The blue area is an initial condition of 2000000 particles counted on 100 bins, and the purple area is 2000000 particles two hours later, calculated by the Euler-Maruyama numerical SDE scheme (Eq. (2.24)), which will be introduced very soon.

from 0, $K_z(z)$ diffusivity coefficient, and $dW(t)$ is a Wiener noise increment with the following properties:

- $W(t)$ is a Wiener process.
- $\langle W(t) \rangle = 0$ for all $t$.
- $\text{Std}(W(t) - W(s)) = \sqrt{|t - s|}$

The $\langle \cdot \rangle$ is the expectation value operator. Std is the standard deviation. If we define

$$dt := t - s, \quad dW := W(t) - W(s),$$

the expectation value of $dW^2$ is

$$\langle dW^2 \rangle = \text{Var}(dW) + \langle dW \rangle.$$

The second term on the right-hand side is zero by the Wiener processes property, and the first term on the right is the variance, which is the square of the standard deviation by definition. We obtain that

$$\langle dW^2 \rangle = dt. \tag{2.4}$$

This tells us that the expectation value of the "size" of $dW$ is $\sqrt{dt}$. Mathematically speaking, it is $\sqrt{\langle dW^2 \rangle} = \sqrt{dt}$. Consequently, $\sqrt{2K_z}dW$ has order $\sqrt{2K_z dt}$. The first term on the right-hand side of Eq. (2.3) is deterministic, and it depends on the spatial derivative of diffusivity coefficient $\partial_z K_z(z)$. The second term on the right-hand side is stochastic, and it has order $\sqrt{2K_z dt}$ as mentioned above. Both terms depend on the diffusivity profile $K(z)$. If $K_z(z) = 0$ for all $z$ , then it just leads us to an ordinary differential equation (ODE). However, the ocean is a turbulent environment, therefore, it can not be zero everywhere. We simplify Eq. (2.3) as

$$dZ(t) = a(z)dt + b(z)dW(t), \tag{2.5}$$

where $a = w + \partial_z K_z(z)$, stands for the deterministic part and is called the drift coefficient, and $b = \sqrt{2K_z(z)}$ represents the stochastic part and is called the diffusion coefficient. Eq. (2.5) is on differential form, we can also obtain an integral form, which is addressed below.

$$Z_t = Z_0 + \int_0^t a(Z_{\tilde{t}})d\tilde{t} + \int_0^t b(Z_{\tilde{t}})dW_{\tilde{t}}, \tag{2.6}$$

where the $\tilde{t}$ is a dummy variable. The natural question is what the last term on the right-hand side means and how we interpret it. It turns out there is not a uniform definition for this term. Here we will introduce only two different definitions in Section 2.5 which we will use later.

## 2.4 Derive drift and diffusion coefficient

After reading many articles of particle tracking, none of them try to derive the drift coefficient $a(z,t)$ and diffuse coefficient $b(z,t)$ explicitly. Most of them just refer to a comprehensive SDE book such as Arnold (1974), but I could not find any derivation in this book that can lead us directly to the derivation of coefficients $a$ and $b$ from Eq. (2.1). Therefore, let us derive them once for all. It turns out it will become useful in Section 3.4.

We start with the book Kloeden and Platen (2013, p. 36). It states that a Markov process with transition density function $p(x,s;y,t)$ is defined in such a way

$$P(x,s;[y_0,y_1],t) = \int_{y_0}^{y_1} p(x,s;y,t)dy \qquad \text{(Markov process)}, \tag{2.7}$$

where $P(x, s; [y_0, y_1], t)$ is the probability to find the particle between position $y_0$ and $y_1$ at time $t$ when the position is given $x$ at time $s$. The transition density function gives us the probability to find the particle between $[y_0, y_1]$ at time $t$ when we have given the particle position $x$ at time $s$, where $t > s$. A Markov process with transition density $p(x, s; y, t)$ is called a diffusion process if the following three limits exist for all $\epsilon > 0$, $s > 0$ and $x \in \mathbb{R}$:

$$\lim_{t \downarrow s} \frac{1}{t-s} \int_{|y-x|>\epsilon} p(x, s; y, t) dy = 0 \tag{2.8}$$

$$\lim_{t \downarrow s} \frac{1}{t-s} \int_{|y-x|<\epsilon} (y-x) p(x, s; y, t) dy =: a(x, s) \tag{2.9}$$

$$\lim_{t \downarrow s} \frac{1}{t-s} \int_{|y-x|<\epsilon} (y-x)^2 p(x, s; y, t) dy =: b^2(x, s), \tag{2.10}$$

where the symbols $t \downarrow s$ means $t \to s; t > s$. $x$ is fixed in the integration, the integration domain of Eq. (2.8) ($|y - x| > \epsilon$) means for all y has distance from $x$ bigger than $\epsilon$. Similar for the integration domain of Eq. (2.9) and Eq. (2.10), but $y$ has distance from $x$ smaller than $\epsilon$. The $a(s, x)$ is called the drift coefficient, and $b(s, x)$ is called the diffusion coefficient. Following the definition of expectation value $\langle \cdot \rangle$, we can rewrite Eq. (2.9) and Eq. (2.10) as follows:

$$a(s, x) = \lim_{t \downarrow s} \frac{1}{t-s} \langle Z(t) - Z(s) | Z(s) = x \rangle \tag{2.11}$$

and

$$b^2(s, x) = \lim_{t \downarrow s} \frac{1}{t-s} \langle (Z(t) - Z(s))^2 | Z(s) = x \rangle. \tag{2.12}$$

From Eq. (2.11), it is not hard to see that $a(x, s)$ is the instantaneous rate of mean change of the process. Similarly, from Eq. (2.12), $b^2(x, s)$ is the instantaneous rate of squared fluctuation of the process at position $x$ and time $s$. To find the expressions of $a(x, s)$ and $b(x, s)$, we need the transition probability function $p(x, s; y, t)$.

As mentioned in chapter 2.2, we can reinterpret the Eq. (2.1) as a Fokker-Planck equation by dividing a constant. This means that we substitute $C$ in Eq. (2.1) to $\rho$ in Eq. (2.2); The constant is total number of particles of our domain, $\int_{-\infty}^{\infty} C(z, t) dz$, which it will not change with time. Dividing Eq. (2.1) by this constant and combining with Eq. (2.2) gives us the following:

$$\frac{\partial}{\partial t} \rho = -\frac{\partial}{\partial z} (\omega \rho - K \frac{\partial}{\partial z} \rho) \tag{2.13}$$

where $\rho$ is the probability density function to find the particles in our domain. Let time $s$ be our initial time and the function $\delta(x)$ be our initial condition. The PDF at the time $s$ is $\rho(x,s) = \delta(x)$. Solving the partial differential Eq. (2.13) with initial condition gives us $\rho(x,s;y,t)$ which is the transition density function. Therefore, the $\rho(x,s;y,t)$ is equivalent to $p(x,s;y,t)$. However, solving Eq. (2.13) analytically is often very difficult, or even impossible, but we need it to compute the coefficients $a$ and $b$. It seems we have trouble here. However, the book Kloeden and Platen (2013) states that :

> "When the drift $a$ and diffusion coefficient $b$ of a diffusion process are moderately smooth functions, then its transition density $p(x,s;y,t)$ also satisfies partial differential equations, these are the Kolmogorov forward equation and Kolmogorov backward equation." (Kloeden & Platen, 2013, p. 37)

In our case, we only need the Kolmogorov forward equation. That is

$$\frac{\partial}{\partial t}p(x,s;y,t) + \frac{\partial}{\partial y}[a(y,t)p(x,s;y,t)] - \frac{\partial^2}{\partial y^2}[b^2(y,t)p(x,s;y,t)] = 0, \quad \text{fixed } (x,s). \tag{2.14}$$

The variable $y$ in our case is $z$. We change $y$ to $z$. Now, let us expand Eq. (2.14) and Eq. (2.13). And then, we compare these two equations term by term such as $p$, $\frac{\partial}{\partial z}p$ and $\frac{\partial^2}{\partial z^2}p$. These two equations become the following:

$$\frac{\partial}{\partial t}p + (\frac{\partial a}{\partial z} - \frac{1}{2}\frac{\partial^2 b^2}{\partial z^2})p + (a - \frac{\partial b^2}{\partial z})\frac{\partial}{\partial z}p - \frac{b^2}{2}\frac{\partial^2}{\partial z^2}p = 0 \tag{2.15}$$

and

$$\frac{\partial p}{\partial t} + \frac{\partial \omega}{\partial z}p + (\omega - \frac{\partial K}{\partial z})\frac{\partial}{\partial z}p - K\frac{\partial^2}{\partial z^2}p = 0. \tag{2.16}$$

Of course, Eq. (2.15) and Eq. (2.16) have to be equal at every position $z$ and time $t$. Therefore, each term of $\frac{\partial^i}{\partial z^i}p$ for $i \in \{0,1,2\}$ has to be equal in both equations as well. Consequently, we obtain

$$\frac{\partial a}{\partial z} - \frac{1}{2}\frac{\partial^2 b^2}{\partial z^2} = \frac{\partial \omega}{\partial z} \tag{2.17a}$$

$$a - \frac{\partial b^2}{\partial z} = \omega - \frac{\partial K}{\partial z} \tag{2.17b}$$

$$\frac{b^2}{2} = K \tag{2.17c}$$

The last equation gives us $b^2 = 2K$. And then, we substitute $b^2$ of $K$ in second equation. It gives us $a = \omega + \frac{\partial K}{\partial z}$. Let's check that the expressions of $a$ and $b^2$ are indeed satisfying Eq. (2.17a).

$$\frac{\partial \omega}{\partial z} + \frac{\partial^2}{\partial z^2} K - \frac{1}{2} \frac{\partial^2}{\partial z^2} (2K) = \frac{\partial \omega}{\partial z}$$

Congratulations! We finally have obtained the drift $a$ and diffusion coefficient $b$, but we are not finished yet. Now, we have to link these two coefficients to the SDE Eq. (2.3). To do this, we need the three assumptions from Kloeden and Platen (2013, p. 128-129) about the coefficients $a$ and $b$. They are *1. Lipschitz condition*, *2. Linear growth bound* and *3. Initial value*. Here we will only introduce the Lipschitz condition, which will be crucial later.

*Lipschitz condition:*

There exists a $L > 0$ such that:

$$|a(x,t) - a(y,t)| < L|x - y| \tag{2.18}$$

and

$$|b(x,t) - b(y,t)| < L|x - y| \tag{2.19}$$

for $x, y \in \mathbb{R}$ and $t \in [t_0, T_{end}]$.

We will not express *linear growth* and *initial value* here since they are not crucial for this thesis, but the reader can refer to Kloeden and Platen (2013, p. 128-129) for more details. In addition, we also need one more theorem,

> "Assume that $a$ and $b$ are continuous and that the three assumptions hold, then the solution $Z(t)$ of [Eq. (2.5)] for any initial value is a diffusion process on $t \in [t_0, T]$ with drift $a$ and diffusion $b$ coefficient. " (Kloeden & Platen, 2013, p. 142)

Here, $t_0$ is the initial time and $T$ is maximum time such that those three assumptions hold. In our case, $T$ is infinity, because our $a$ and $b$ do not depend on time, since $K$ does not depend on time. The proof of this theorem is given in Kloeden and Platen (2013, p. 142). Finally, We have derived our drift $a$ and diffusion coefficient $b$.

## 2.5 Stochastic integration

In this subsection we will present two different integrals that define the last term of Eq. (2.6). For more detail, please refer to Evans (2010, chapter 6).

### 2.5.1 Itô integral

One common interpretation of the last term of Eq. (2.6) is called Itô integral. This interpretation is equivalent to Riemann sum when we always evaluates the integrand on the left-hand side of each sub-intervals $[t_q, t_{q+1}]$. This means that we are always using the information we have right now. Based on this property, Itô integrals are also frequently used in economics and finance.

$$\int_0^T b(Z_{\tilde{t}})dW_{\tilde{t}} := \lim_{|P^n|\to 0} \sum_{q=0}^{n-1} b(Z(t_q))(W(t_{q+1}) - W(t_q)) \tag{2.20}$$

where $P^n := \{0 = t_0 < ... < t_n = T\}$ is just the partition of $[0, T]$ and $n$ is number of discrete time point of this partition. $|P^n|$ is the number of sub interval.

### 2.5.2 Stratonovich integral

Another common interpretation of the last term of Eq. (2.6) is called the Stratonovich integral. Stratonovich integral is essentially the Riemann sum by evaluating the half of the left-hand side and half of the right-hand side and sum them up.

$$\int_0^T b(Z_{\tilde{t}}) \circ dW_{\tilde{t}} := \lim_{|P^n|\to 0} \sum_{q=0}^{n-1} \frac{b(Z(t_{q+1})) + b(Z(t_q))}{2}(W(t_{q+1}) - W(t_q)) \tag{2.21}$$

where $P^n$ is defined just like for the Itô integral. The $\circ$ symbol helps us to distinguish Itô integral and Stratonovich integral. However, there is a problem. We do not have $Z(t^n_{q+1})$ before we have evaluated the integral. Therefore, to obtain a reasonably estimated value, we use

$$\frac{b(Z(t_{q+1})) + b(Z(t_q))}{2} \approx b\left(Z(t_q) + \frac{1}{2}a(Z) \cdot (t_{q+1} - t_q)\right). \tag{2.22}$$

Recall that $b(z)$ is a function of $z$. Here we use the $\cdot$ symbol to indicate that the $(t^n_{q+1} - t^n_q)$ is not an input argument of $a(Z)$, but it is just a multiplication. At this point, you may wonder how we quantify the performance of numerical schemes? This is done by the concept of convergence, like for ODEs, which will be discussed in the next section.

## 2.6 Weak and Strong convergence

In the ordinary differential equation (ODE), a meaningful numerical approximation should converge to the true solution as we are using a smaller and smaller timestep. Different ODE schemes may converge to a true solution in a different order. The same philosophy also applies to SDEs, but there are two different types of convergence. The first one is called weak convergence and the second one is called strong convergence. The two different type of convergence arise due to the stochastic property of SDEs. A numerical scheme is said to have a weak/strong convergence of $\gamma$ if there exists a constant $\Lambda$ such that

$$\begin{aligned} |\langle p(Z_n)\rangle - \langle p(Z(\tau))\rangle| &\leq \Lambda\ \delta t^\gamma : \text{weak} \\ \langle |Z_n - Z(\tau)|\rangle &\leq \Lambda\ \delta t^\gamma : \text{strong} \end{aligned} \tag{2.23}$$

for any fixed $\tau = n\delta t \in [0, T]$ and $\delta t$ sufficiently small (Gräwe, 2011). $Z_n$ represents the exact solution and $Z(\tau)$ represents the approximation generated by the numerical scheme. The $p(\cdot)$ is an arbitrary function with polynomial growth. This mean that $p(\cdot)$ is also a linear combination of the moments of the distribution since the class of functions $p(\cdot)$ includes $z$, $z^2$, $z^3$, etc. Hence, weak convergence implies that all the moments of the distribution converge.

All the moments are approximated with the desired accuracy. The meaning of weak convergence is to measure the overall difference in the distribution, whereas the strong convergence is to measure the difference in individual trajectories, which means that for a given realisation of a Wiener process $(W(t))$ and use it to solve the SDE (Eq. (2.3)). We will test these properties with different numerical schemes in appendix B.

An interesting thing worth to mention is that if $Z_n$ and $Z(\tau)$ are entirely deterministic, which means the probability density function (PDF) is a delta function, the weak convergence and strong convergence will be equal to each other. This makes sense because changing the PDF to delta function is just changing the SDE to ODE. An informal proof is given in appendix C.

## 2.7 Numerical schemes

Here we will introduce four numerical schemes for solving SDEs that we will use. All the below numerical schemes involve $\Delta W_n$. This is a random variable with a normal distribution with $\langle \Delta W_n \rangle = 0$ and $\mathrm{Std}(\Delta W_n) = \sqrt{\Delta t}$.

### • The Euler-Maruyama (E1)

The Euler-Maruyama scheme (Maruyama, 1955) is quite straightforward. We just exchange the $dt$ to $\Delta t$ and $dw$ to $\Delta W$ in Eq. (2.5):

$$Z_{n+1} = Z_n + a\ \Delta t + b\ \Delta W_n. \tag{2.24}$$

As we can see from Eq. (2.24), this is defined in the Itô sense. The Euler scheme has order $\frac{1}{2}$ in the sense of strong convergence. It also has order 1 in weak convergence sense. We will show these two properties numerically in appendix B.

### • The Visser scheme (V1)

What does the Eq. (2.6) look like if it is defined in the Stratonovich sense? This is the Visser scheme as shown below

$$Z_{n+1} = Z_n + a\ \Delta t + b(\widetilde{Z})\ \Delta W_n \tag{2.25}$$

with $\widetilde{Z} = Z_n + \frac{1}{2} a\ \Delta t$. The $\Delta W_n$ is defined just like for the Euler scheme. This scheme was proposed by Visser (1997). It has the same weak convergence order as E1. But the strong convergence, in Gräwe (2011) is not defined on Stratonovich integral, therefore, it is not defined on the V1 scheme.

### • The Milstein scheme (M1)

The Milstein scheme (Milshtein, 1975) obtains a higher accuracy by not treating $b(Z)$ as constant. The idea is to use Taylor expansion to get higher order terms for the integral, see e.g. Gräwe (2011).

$$Z_{n+1} = Z_n + a\ \Delta t + b\ \Delta W_n + \frac{1}{2} b b' \cdot (\Delta W_n^2 - \Delta t). \tag{2.26}$$

The apostrophe sign (for example: $b'$) means spatial derivative operator (i.e. $\frac{\partial}{\partial z}b$). We can also rewrite Eq. (2.26) in this way:

$$Z_{n+1} = Z_n + \omega\Delta t + \partial_z K\frac{1}{2}(\Delta W_n^2 + \Delta t) + \sqrt{2K}\Delta W_n. \tag{2.27}$$

As we will see later, M1 scheme has both weak and strong convergence order 1.

- **The Milstein second order scheme (M2)**

This scheme is actually derived from another scheme called the 1.5 order strong Taylor scheme, which has 1.5th order strong convergence and 2nd order weak convergence, see Gräwe (2011). But if we are only interested in simulating the particle concentration field, then we only need the weak convergence approximation. The scheme becomes:

$$\begin{aligned} Z_{n+1} = Z_n + a\ \Delta t + b\ \Delta W_n + \frac{1}{2}bb'\cdot(\Delta W_n^2 - \Delta t) \\ + \frac{1}{2}\left((ab)' + \frac{1}{2}b''b^2\right)\Delta W_n\Delta t + \frac{1}{2}(aa' + \frac{1}{2}a''b^2)\Delta t^2 \end{aligned} \tag{2.28}$$

This scheme was also proposed by Milshtein (1979). This scheme does not have a defined strong convergence.

## 2.8 Choice of $\Delta t$

The reader may wonder what $\Delta t$ should we use. It usually depends on how much error we tolerate. But for at least physically meaningful results, we need

$$\Delta t << \min\left(\frac{1}{|\partial_{zz}K(z)|}\right) \quad \text{for} \quad \forall z \in \mathbb{D}, \tag{2.29}$$

where $\partial_{zz}K(z)$ is the double derivative of diffusivity with respect to space, $\mathbb{D}$ is the domain and $|\cdot|$ is the absolute value operator. This limit was originally proposed by Visser (1997), and it is widely accepted. We can see that if $K(z)$ is linear or constant, Eq. (2.29) becomes $\Delta t << \infty$. However, when we have a boundary condition, Eq. (2.29) may not be the only criterion. Some numerical boundary schemes may cause a upper bound of $\Delta t$, which we will see in Chapter 3.

## 2.9 Choice of a SDE scheme

Which one scheme should we choose for particle tracking? The answer is, as usual, it depends. In Section 2.6, we have mention that the strong convergence is to measure the average difference in the individual trajectories for a given Wiener process. Weak convergence is to measure the average overall difference in the distribution for all different Wiener process, if we are interested in simulating the trajectories of individual particles, then we has to use a SDE scheme which have strong order. If we are only interested in concentration distribution, then we can choose a SDE that has both strong and weak convergence or just weak convergence. This that gives us some freedom to choose the SDE scheme with desired order.

## 2.10 Well-mixed condition

Strong convergence and weak convergence are the performance indices of numerical solutions of SDEs in general purpose. In special purpose of use, there may be other criteria that need to be satisfied. In particle tracking, a good SDE scheme has to satisfy four main criteria. Thomson (1987) formulated these four criteria, and states that the first criterion is sufficient to ensure that all other criteria are satisfied. Therefore, we will only address the first one in the following.

*Well-mixed condition*

If the particles of a passive tracer are initially well-mixed, they will remain well-mixed.

I personally found it difficult to understand the proof of this criterion in the article because some mathematical notations and technical terms are not elaborated. However, as Wilson and Flesch (1993) also stated, the well-mixed condition (WMC) is just a statement with respect to entropy. Its satisfaction by a model prohibits the spurious evolution of order from disorder. Of course, the advection $\omega$ is equal to zero in this case.

Instead of proving the WMC, we demonstrate an diffusion process that the WMC has to satisfied. Let us consider a general diffusivity $K(z)$ which is smooth, non-constant and $K(z) > 0$ for $z \in [0, H]$. We write down the

diffusion equation in one dimension:

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial z} j_z \tag{2.30}$$

where

$$j_z = -K \frac{\partial \rho}{\partial z} \qquad \text{(Fick's Law)} \tag{2.31}$$

with boundary conditions

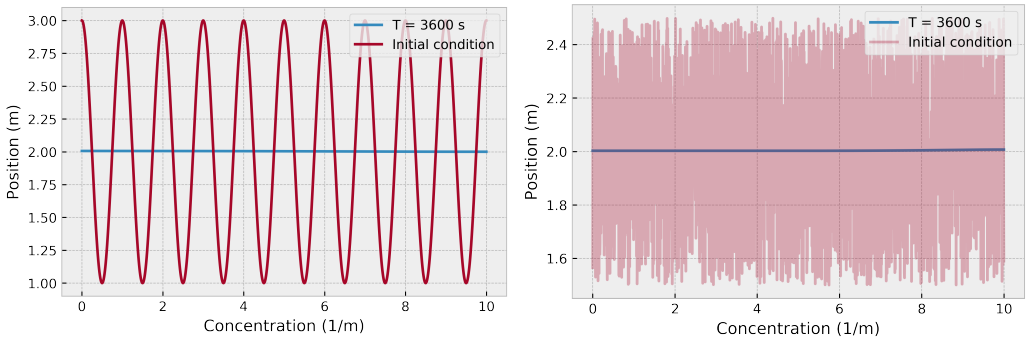
$$\left.\frac{\partial \rho}{\partial z}\right|_{z=0} = \left.\frac{\partial \rho}{\partial z}\right|_{z=H} = 0 \qquad \text{(Neumann boundary conditions).} \tag{2.32}$$

Let us start with $\rho(z, 0) = \sin(2\pi z) + 2$ and apply a commonly used PDE numerical scheme, called the Forward time central space (FTCS) scheme to solve Eq. (2.30) with boundary condition Eq. (2.32) and a diffusivity Eq. (3.1b). For the formulation of FTCS, refer to Sauer (2011). We use model time $T_{end} = 3600\,\text{s}$ and we obtain the result and shown in Fig. 2.2a. The red line is the initial concentration, and the blue line is the concentration at $t = 3600\,\text{s}$. We can easily see that the blue line behaves like a constant over the domain. If we try another initial condition, like the red line in Fig. 2.2b, we still get the same behavior as in Fig. 2.2a. In fact, for diffusion problem with Eq. (2.30), Eq. (2.31) and Eq. (2.32), no matter what initial concentration we use, we will always obtain a constant concentration when $t \to \infty$.

Here we will not give a rigours mathematical proof, but Eq. (2.31) shows the net flux is always from high concentration to low concentration since $K(z) > 0$ for all position $z$. The process continues until the gradient of the concentration is zero ($\frac{\partial \rho}{\partial z} = 0$). In this case, $\frac{\partial \rho}{\partial t}$ in Eq. (2.30) also goes to zero. Therefore, we reach steady concentration (equilibrium) at $t = \infty$. Of course, if the particles are initially (perfectly) well mixed, which mean $\frac{\partial \rho}{\partial z} = 0$, they should always stay well mixed. Remember that we derived the drift $a(z, t)$ and diffusion $b(z, t)$ from this PDE in Section 2.4. The weak convergence of any numerical SDE scheme should converge to the normalised $\rho$. Consequently, the WMC should also hold for numerical SDE schemes.

## 2.11 Entropy demystified

This section is not essential but relevant to our topic. I still encourage the reader take a look at this section because it is quite interesting and controver-




(a) (b)

Figure 2.2: Using FTCS scheme to solve the Eq. (2.30) with two different initial concentration (red line). The blue line is the concentration at $t = 3600\,\mathrm{s}$, and they look like constant comparing to the initial concentration.

sial. But the reader can skip this section if he/she/you want. Many people have been taught form school that entropy ($S$) is disorder. However, there are many questions remaining, such as how the unit joule per kelvin (J/K) associate to disorder, or what $\Delta S = 1\,\mathrm{J/K}$ means. I spent my last Christmas time with an excellent book called *Entropy demystified* (Ben-Naim, 2008). It turns out it is helpful to my thesis. In this section, I will present the book in a very short way; we first argue that the disorder is a highly subjective concept by showing two figures. And then, we try to reformulate the entropy and explain what entropy means. Next, we will see the relation between our new reformulated entropy and the widely adapted definition $S = K_b \log_e(W)$, where $K_b$ is the Boltzmann constant and $W$ is the number of micro-states. Finally, we will discuss the second law of thermodynamics and why do we care in this thesis.

Even though we do not have a precise definition of order/disorder, we still use these terms a lot in our daily life. Most of the time, it does not cause a disagreement, but sometimes it does. For example, my mother always complains about how messy my room is, but I honestly think that my room is very orderly. Everything is exactly where it should be. However, most people, including me, my mother and the reader (you) will agree that the upper image in Fig. 2.3 is very ordered, and the lower image is disordered. These two images have the same number of dice. Each die has a probability to be place in a given position and show a face. Mathematically speaking,

these two images have equal probability to occur. But our intuition tells us that the ordered image has much less probability than the disordered image to occur. It is simply because the upper image means something to us, and the lower image means nothing. Now, take a close look at the disordered image. If we only focus on those dice with side number 1, you may discover something interesting. Yes, it is the author's name (Arieh). For a much clearer look, refer to Fig. A.2. Suddenly, the lower figure means something for us too. Is it still disordered for you? Maybe not. Perhaps we would never know until we find out. At this stage, I hope the reader may agree that order/disorder is highly subjective and unreliable for scientific purposes.

So, what is entropy? I am going to skip all the detail of Ben-Naim (2008) and jump right in to the conclusion. **Entropy is missing information**. More precisely, entropy could be and should be defined as the **Number of optimal binary questions we need to ask, to get the specific configuration of the system.** The reader may wonder what optimal binary questions meas. This is a question that only accepts "yes" or "no" as an answer, and both "yes" or "no" give us equal (or almost equal) amount of information. For example, in Fig. 2.4, there is a coin sitting inside the $4 \times 4$ boxes. Assuming we do not know position of coin and have to guess. There is someone who will only answer "yes" or "no" to our guess or question. The most straightforward way is to ask for every position. Is it in A1? Is it in A2? ... Is it in B1?...Is it in D1?...Is it in D4? Of course, we could be very lucky if our first question is "Is it in C3?" But this happens by chance. The expectation of number of questions we need to ask in this way to obtain the specific position of the coin is 7.5. Now, how about we change the strategy of asking questions. Now, we ask "Is it in column 1 or 2?", "Is it in row A or B?", "Is it in 3?", "Is it in D?". In this way, we always need to ask 4 questions, no less no more. Unlike the previous strategy, this one gives us always equal amount information. Number of question is no longer random and is 4. 4 is less than 7.5, which is expectation of number of question in previous strategy. Therefore, we call the second strategy as optimal binary question. Now, we understand what optimal binary question meas. Let's see another example. We initiate $N$ distinguishable ideal particles in the left box in Fig. 2.5. There is a wall to prevent particles moving to the right box. Here, we define what information means in this context. Information of a particle is either "left" or "right". At the beginning, we do not have information loss because we initiate them in this way. Now, we remove the wall. Particles will diffuse from left to right. Assume we can freeze a time

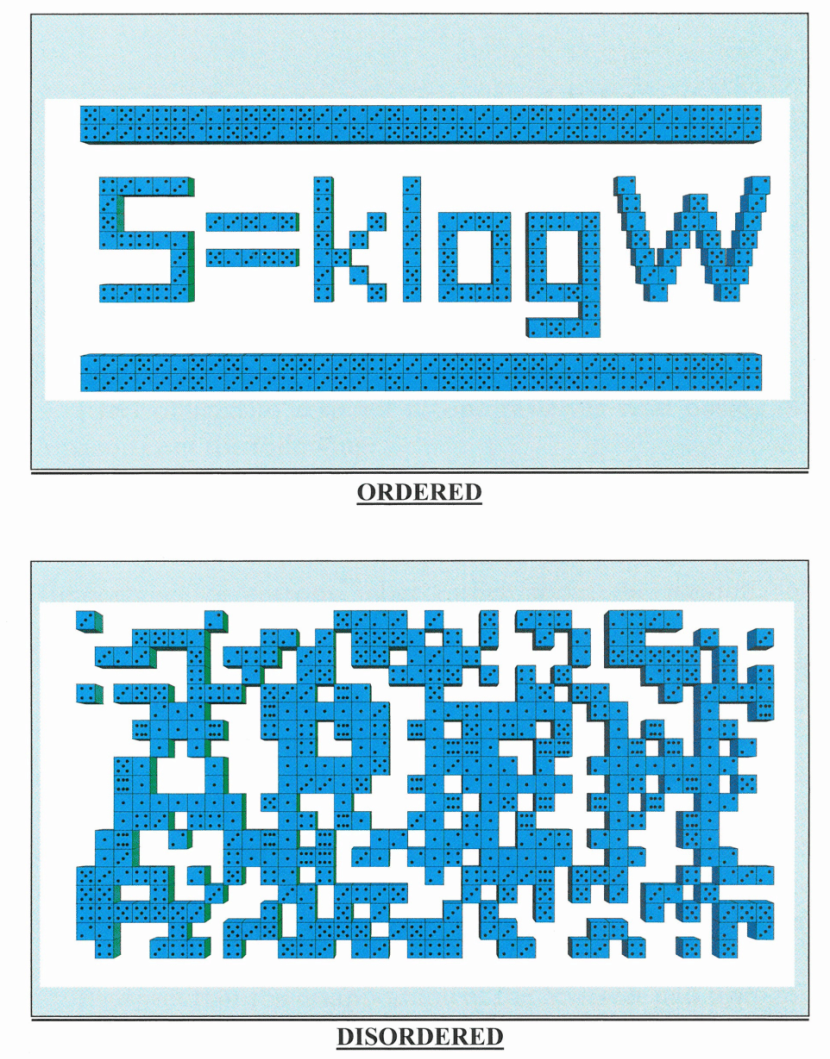



Figure 2.3: Even if we do not have a precise definition of order/disorder, I think the reader may still agree that upper image is very ordered, and the lower image is disordered.
Permission granted by author and World Scientific. Source: *Entropy demystified* second edition, Arieh Ben-Naim, Copyright @ 2016 by World Scientific Publishing Co. Ptw. Ltd

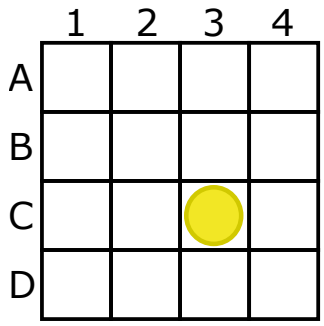



Figure 2.4: A $4 \times 4$ box, one of the square contains a coin. We can ask binary questions to obtain the information of the position of the coin. We want to find a strategy to ask the minimum number of questions.

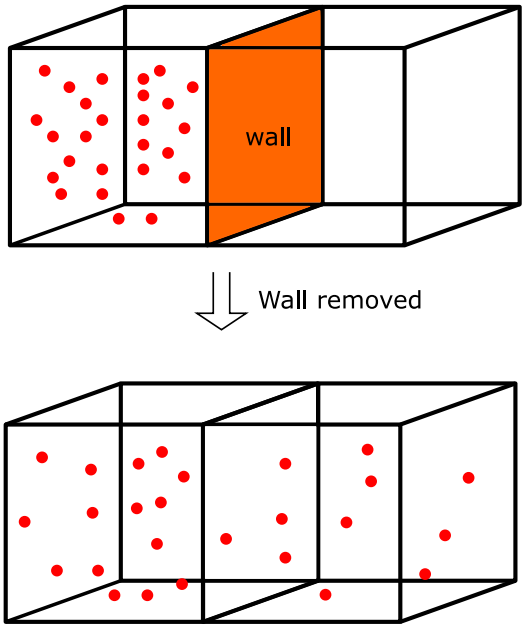



Figure 2.5: Two boxes with a wall in between, we initiate all the distinguishable ideal gas particle in the left box. After we break the wall, we should expect that some particles will diffuse into the right box. However, it is still possible for all particles to remain inside the left box, but extremely unlikely.

frame later on. How many optimal binary questions do we need to ask to get the specific configuration of all particles? Well, for each particle, we need only ask one question. Therefore, we need to ask $N$ questions. We define our entropy as the following:

$$\tilde{S} = \log_2(W^N),$$

where $W$ is the number of states of a particle, and $N$ is the number of particles. Before we remove the wall, there is only one configuration, which is all the particles in the left box. This means $W = 1$. We obtain

$$\tilde{S} = \log_2(1^N) = 0.$$

After the wall has be removed, $W = 2$ because there are left box and right box to be our states. We obtain

$$\tilde{S} = \log_2(2^N) = N \log_2(2) = N.$$

This means that the change of entropy in our definition is $N$. Increasing entropy means missing information and nothing more. Now, we connect our definition of entropy to the general definition of entropy.

$$\tilde{S} = \frac{K_b \log_e(W^N)}{K_b \log_e(2)} = \frac{S}{K_b \log_e(2)}.$$

The denominator is just a constant. This does not change any physics. $\tilde{S}$ is dimensionless but $S$ is (J/K). The origin of the confusion, perhaps, comes form what we call "temperature". In the history of physics, temperature was not fully understood, until statistical mechanics showed that temperature is proportional to the average kinetic energy of the substance. Therefore, we need a constant to convert temperature to energy, which is $K_b$. We could define our new temperature as $\tilde{T} = T \cdot K_b$, then $\tilde{T}$ has unit (J). This makes entropy dimensionless.

Ben-Naim (2008) also shows that the second law of thermodynamics is actually not a law of physics in the conventional sense. Exception do occur although this is extremely rare. Here, I have skipped a lot of detail. For example, what happens to our definition of entropy if we have continues states. However, this is out of our scope. For more detail, please refer to Ben-Naim (2008).

Why do we care about entropy in this thesis? Recall that the WMC is just a statement with respect to entropy. Later on, we will define a WMC test and assume extremely rare events will not occur. For example, all the particles in Fig. 2.5 will not all go to one side of these two boxes. This is still possible, but extremely rare.

# Chapter 3

# Boundary artifact in numerical SDE

In the Section 3.1, we define three diffusivity profiles that will be used later, and explain the difference between thermal diffusion and turbulent diffusion. In Section 3.2, we introduce a reflective scheme for SDE which is widely used in many different research fields. In Section 3.3, we demonstrate the boundary artifact caused by the reflective scheme from Section 3.2. In Section 3.4, we first analyze what is wrong with the reflective scheme. Then, we reach a key result from the analysis. In Section 3.5, we use the result from Section 3.4 to make two quick suggestions to test E1 and verify our conclusion from Section 3.4. In Section 3.7, we discuss whether the suggestions change the physical reality.

## 3.1 Diffusivity profiles

There are usually two kind of diffusion process:

*1. Thermal diffusion*

The first one is called thermal diffusion, also known as molecular diffusion. In this case the diffusivity depends on temperature (internal energy). Let us do an experiment. We start with two glasses of water; one is 95 °C and another one is 10 °C. If we drop an equal amount of color ink into both glasses and observe how they diffuse in two differ-

ent temperatures, we will find out that the ink in the high temperature glass diffuses faster than the low temperature because hot water have higher internal kinetic energy and it makes the color ink particle have higher kinetic energy than cold water. In this case we will say the diffusivity of hot water $K_h$ is higher then the diffusivity of cold water.

*2. Turbulent diffusion*

The second one is called turbulent diffusion, also known as eddy diffusion. In this case the diffusivity generally depends on Reynolds number and others factors, but not temperature. Let us do a similar experiment like the previous one. Assuming we have two glasses of water with the same temperature, called glass A and glass B, and we drop an equal amount of color ink in to both of the glasses. Now, we use a spoon to stir glass A. This creates a turbulent environment in glass A. We will find that ink in glass A diffuses faster then in glass B.

Which type of diffusion dominates in the ocean? The answer is turbulent diffusion. As we have applied external force in our latest experiment, there are also many sources to cause turbulent diffusion in the ocean. For example, tidal force, wind, large-scale currents, human activities, sea creature activities, etc. For more source of the diffusivity in the ocean, please refer to Dabiri (2010); Munk and Wunsch (1998). The point is that thermal diffusion becomes trivial on all scales from millimeters and up. Therefore, from now on, we will only consider turbulent diffusion (eddy diffusion) in this thesis, and we assume that the particle's inertia is negligible such that the particle diffusivity is the same as the fluid diffusivity.

Before we progress any further, we want to set up three different diffusivity profiles which will be frequently used later. These are

$$K_a(z) = K_1 + Gze^{-\alpha z} \tag{3.1a}$$
$$K_b(z) = K_2 + Gze^{-\alpha z} \tag{3.1b}$$
$$K_c(z) = K_3 + Gze^{-\alpha z} \tag{3.1c}$$

where

$$K_1 = 0\,\mathrm{m}^2/\mathrm{s},\ K_2 = 10^{-6}\,\mathrm{m}^2/\mathrm{s},\ K_3 = 10^{-3}\,\mathrm{m}^2/\mathrm{s},$$
$$G = 6 \cdot 10^{-3}\,\mathrm{m/s},\ \alpha = 0.5\,1/\mathrm{m} \text{ and } z \in [0, 10].$$

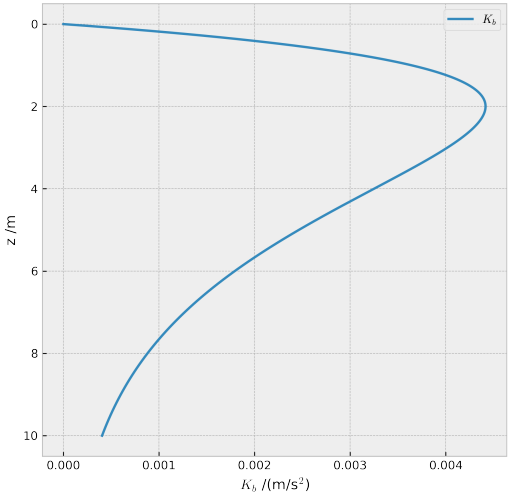



Figure 3.1: The diffusivity of $K_b$.

The Fig. 3.1 shows the diffusivity $K_b$. $K_a$ and $K_c$ are a constant shift from $K_b$. Beware that Eq. (3.1a) has $K_a(0) = 0$. In this case, a process described by Eq. (3.1c) ($K_a$) is technically no longer a diffusion process because the concentration at $z = 0$ will never be changed as followed by the PDE diffusion Eq. (2.30). Although it may not classify as a diffusion problem, we will still include this diffusivity. The reason is that many articles , such as Gräwe et al. (2012); Notter and Sleicher (1971), use diffusivity that goes to zero at the boundary. In this thesis, we will analyze in Section 3.4 the case where diffusivity goes to zero at the boundary, as this will actually cause problems in numerical SDE schemes.

## 3.2 The reflective scheme

Neumann boundary condition (NBC) is the most common reflecting scheme for PDE scheme (Eq. (2.32)) since NBC means there is no diffusion flux. We may wonder what is the corresponding reflective scheme for SDE. Gräwe (2011) suggests a straightforward reflection scheme for SDE to imitate the PDE. This is :

> "When a particle crosses the boundary (due to a too large random displacement), it is simply reflected back into the domain by the

amount it penetrates into the boundary domain." (Gräwe, 2011)

In our case, where the boundaries are located at $z = 0$ and $z = H$, the SDE reflective boundary scheme suggested by Gräwe (2011) is

$$Z_{n+1} \rightarrow \begin{cases} -Z_{n+1}, & \text{if } Z_{n+1} < 0 \\ 2H - Z_{n+1}, & \text{if } Z_{n+1} > H \\ Z_{n+1}, & \text{Otherwise.} \end{cases} \tag{3.2}$$

We name this reflective boundary scheme as the perfect reflection scheme (abbrev. PRS) as in Wilson and Flesch (1993). Beware that with perfect reflection scheme, we effectively change the probability density distribution of $dW$ of Eq. (2.5) near the boundary. We have not taken this into account when we derived the drift coefficient $a$ and diffusion coefficient $b$ in Section 2.4. The derivation of SDE (Eq. (2.5)) in Kloeden and Platen (2013) did not consider the boundary effect either. As we will see in Section 3.3, this causes problems.

Note that the WMC still needs to be satisfied with perfect reflection scheme. Hence, we define our SDE numerical problem set. With eddy diffusivity, domain, numerical scheme, reflective boundary scheme and initial condition will form a numerical problem set.

$$Z(t) \begin{cases} \text{Eddy diffusivity} \\ \text{Domain} \\ \text{SDE scheme} \\ \text{Reflective boundary scheme} \\ \text{Initial condition, } Z(0) \end{cases} \tag{3.3}$$

Moreover, we define our WMC test in the following. We denote the total number of particles as $N_p$ and initialize $N_p$ particles with uniform distribution over the domain. Afterward, we simulate them. For every time-step, we will use a histogram to calculate the concentration at the current time frame. Last, we add together all the concentrations from different time frames and divide it by the number of time frames to obtain time-average concentrations, and normalize the concentrations to 1 every where. As required by the WMC, the time-average concentration should stay even over the domain if it is a "good" particle tracking scheme.

## 3.3 Boundary artifacts arise

We are wondering if our numerical SDE schemes with the PRS (Eq. (3.2)) can satisfy the WMC. Let's run the WMC test. We denote length of our domain as $H$, and $N_b$ is number of bins in the concentration histogram. We test the E1, V1, M1 and M2 schemes with $H = 10\,\mathrm{m}$, $N_p = 24000$, $N_b = 2000$, $\Delta t = 10\,\mathrm{s}$, $T_{end} = \Delta t \cdot 3600$ and both with $K_b(z)$ and constant diffusivity $K = 0.003\,\mathrm{m}^2/\mathrm{s}$.

The result is plotted in Fig. 3.2. We can see that we have a problem near the top boundary for non-constant diffusivity $K_b$. For E1 and V1, the concentration is around 30% lower than 1. M1 and M2 have almost zero concentration and raise up very quickly to around 2 then go back to normal. All of them obviously violate the WMC. But it is interesting that the constant diffusivity profile does not exhibit the same problem. Moreover, the bottom boundary does not show this problem either. Of course, it has to do with our diffusivity profile. Keep in mind that the boundary artifact may be very hard to notice as most practical researcher are more interested in the time-evolution of the concentration field instead of the time-average of the concentration field. For a closer look, we denote $H_h$ as the internal length for the histogram. Let $H_h = 0.2\,\mathrm{m}$ and $N_b = 200$. This means that we are only interested in the domain $z \in [0, 0.2]$. The result is plotted in Fig. 3.3.

The reader may think that we can reduce $\Delta t$ to minimize the problem. This is true, indeed. Now we use the same set up as Fig. 3.2 but with $\Delta t = 0.1\,\mathrm{s}$; Fig. 3.4 shows the result. We can see that the domain that is affected by the artifact is much shorter, but the concentration deviation at the boundary is still very visible. In practical usage, reducing $\Delta t$ by factor 10 will increase the wall-time at least 10 times more. The reader may think that the boundary problem will completely go away if we use small enough $\Delta t$. The answer is **NO** with the perfect reflection scheme. No matter what numerical SDE scheme we used, the boundary artifact is still there. Because there is something **fundamentally wrong**. In the next section, we will analyze why.

## 3.4 Analysis of the boundary artifact

In this section, we will take a look what happens if we extend the domain from $[0, H]$ to $[-H, H]$, and mirror the eddy diffusivity about the point at

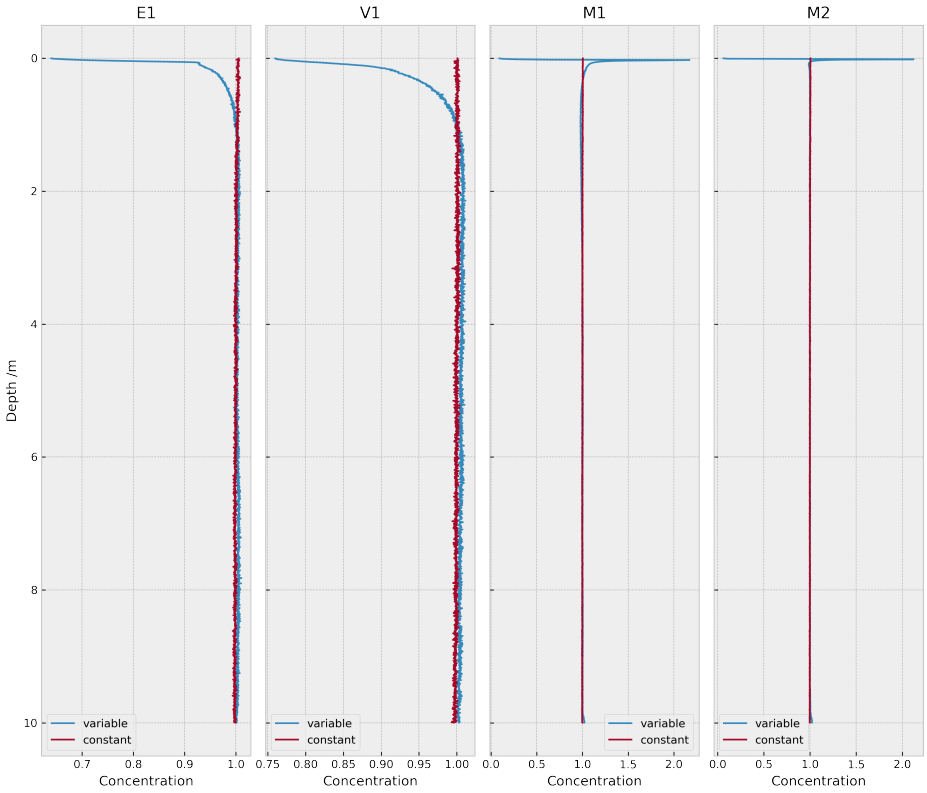



Figure 3.2: The time-average concentration of initially uniformly distributed particles. Here we use $K_b$, $H = 10\,\mathrm{m}$, $N_p = 24000$, $N_b = 2000$, $\Delta t = 10\,\mathrm{s}$ and $T_{end} = \Delta t \cdot 3600$.

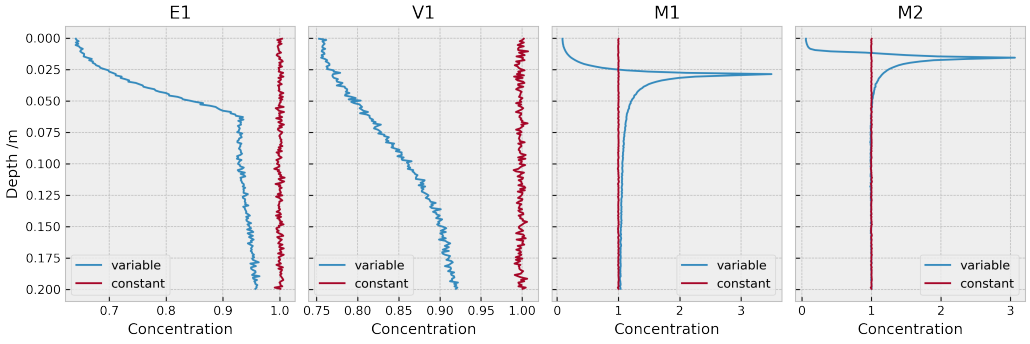



Figure 3.3: The time-average concentration of initially uniformly distributed particles. Here we use $K_b$, $H = 10\,\mathrm{m}$, $N_p = 24000$, $N_b = 200$, $H_h = 0.2\,\mathrm{m}$, $\Delta t = 10\,\mathrm{s}$ and $T_{end} = \Delta t \cdot 3600$.

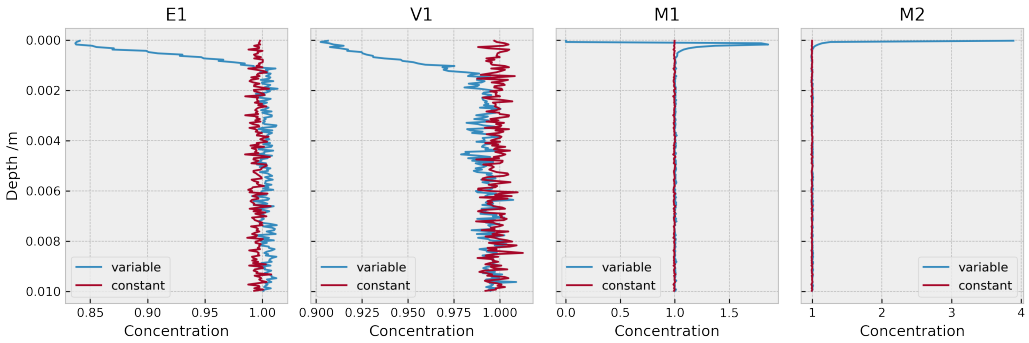
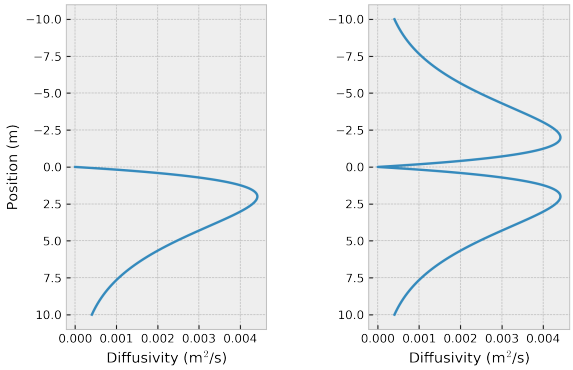



Figure 3.4: The time-average concentration of initially uniformly distributed particles. The set-up are same as Fig. 3.3, but $\Delta t = 0.1\,\mathrm{s}$ and $H_h = 0.01\,\mathrm{m}$.




(a) Diffusivity-original (b) Diffusivity-mirrored

Figure 3.5

$z = 0$ (see Fig. 3.5a. We Fig. 3.5b) and simulate the concentration with diffusivity $K_b$ using both Eulerian PDE and Lagrangian SDE method, which gives us critical information about our PRS.

For an SDE scheme with PRS to be a good particle tracking scheme, the concentration fieldrge has to conveto the PDE concentration field. Therefore, we need the PRS to faithfully imitate the Neumann boundary condition. Instead of performing a weak convergence test, we will mirror the diffusivity $K_b$ in Eq. (3.1b) to see what will happen.

### 3.4.1 Original- Mirrored domain in PDE

Here, we will use $K_b$, which is plotted in Fig. 3.7a, as our diffusivity in the original domain (Non-mirrored domain), and the mirrored version of $K_b$, which is plotted in Fig. 3.7b, as our diffusivity for mirrored domain. First we focus on the Eulerian PDE model on original domain. FTCS is a commonly used finite difference method to solve PDEs. We have used FTCS to solve Eq. (2.13) with $\omega = 0$ in Section 2.10. Here, we will only address the FTCS scheme without deriving it. The reader can easily find it in the literature (see e.g. Sauer (2011)).

$$\Delta\rho_{n,m} = \Delta t\Big[\frac{(K_{n+1} - K_{n-1})}{\Delta z}\frac{(\rho_{n+1,m} - \rho_{n-1,m})}{\Delta z} + K_n\frac{(\rho_{n+1,m} - 2\rho_{n,m} + \rho_{n-1,m})}{\Delta z^2}\Big] \tag{3.4a}$$

$$\rho_{n,m+1} = \rho_{n,m} + \Delta\rho_{n,m}, \tag{3.4b}$$

where $\rho_{n,m} := \rho(n\cdot\Delta z, m\cdot\Delta t)$, $K_n := K(n\cdot\Delta z)$, $\rho(z,t)$ is the concentration at position $z$ and time $t$, $K(z)$ is the eddy diffusivity over position $z \in \mathbb{D}$, $N$ is the number of sub-intervals in $\mathbb{D}$, $\mathbb{D}$ is the spatial closed domain, $\Delta z := \frac{\max\mathbb{D}-\min\mathbb{D}}{N}$; $M$ is the number of sub-intervals in $\mathbb{T}$, $\mathbb{T}$ is the time closed domain, $\Delta t := \frac{\max\mathbb{T}-\min\mathbb{T}}{M}$, $n \in [0,1,...,N]$ and $m \in [0,1,...,M]$. The Neumann boundary condition in this finite difference method may be implemented by creating two pseudo points outside the boundaries. These two points are defined as

$$\rho_{-1,m} := \rho_{0,m} \text{ and } \rho_{N+1,m} := \rho_{N-1,m}. \tag{3.5}$$

When we combine this equation with Eq. (3.4a) then the boundary points for the original domain will be:

$$\Delta\rho_{0,m} = \Delta t\Big[\underbrace{\frac{(K_1 - K_{-1})}{\Delta z}\cdot 0}_{=0} + K_0\frac{(2\rho_{1,m} - 2\rho_{0,m})}{\Delta z^2}\Big] \tag{3.6a}$$

$$\Delta\rho_{N,m} = \Delta t\Big[\underbrace{\frac{(K_{N+1} - K_{N-1})}{\Delta z}\cdot 0}_{=0} + K_N\frac{(2\rho_{N-1,m} - 2\rho_{N,m})}{\Delta z^2}\Big]. \tag{3.6b}$$

Even though the $K_{-1}$ and $K_{N+1}$ are not defined, it does not matter since they will be multiplied with 0.

Similarly, for the mirrored domain, the formula is the same as Eq. (3.4), but we change the $N$ to $2N-1$ and $n \in [0, 1, ..., N]$ to $\tilde{n} \in [-N, -N+1, ..., 0, ...N-1, N]$. The Neumann boundary condition for the mirrored domain will be

$$\rho_{-N-1,m} := \rho_{-N,m} \text{ and } \rho_{N+1,m} := \rho_{N-1,m}, \tag{3.7}$$

and the boundary points for the mirrored domain will be

$$\Delta\rho_{-N,m} = \Delta t\Big[\underbrace{\frac{(K_{-N+1} - K_{-N-1})}{\Delta z} \cdot 0}_{=0} + K_{-N}\frac{(2\rho_{-N+1,m} - 2\rho_{0,m})}{\Delta z^2}\Big] \tag{3.8a}$$

$$\Delta\rho_{N,m} = \Delta t\Big[\underbrace{\frac{(K_{N+1} - K_{N-1})}{\Delta z} \cdot 0}_{=0} + K_{N}\frac{(2\rho_{N+1,m} - 2\rho_{N,m})}{\Delta z^2}\Big] \tag{3.8b}$$

Just like the original domain case, the $K_{-N-1}$ and $K_{N+1}$ are not defined, but it still does not matter.

Now we simulate the PDE with FTCS Eq. (3.4a) in original and mirrored domain. In the original domain, the initial condition is normal distribution function with mean value 5 and variance equal to 0.001. $N = 2000$, $\mathbb{D} = [0, 10]$, $M = 2000$ and $\mathbb{T} = [0, 1000]$. This means $\Delta t = 0.5\,\text{s}$. In the mirrored domain, we will mirror the initial condition from the original domain. Let $N = 4000$ and $\mathbb{D} = [-10, 10]$. The rest is same as original domain. The results are plotted in Fig. 3.7c and Fig. 3.7d. We can see that the result (blue line) is also mirrored. We should not be surprised because in the mirrored domain, the calculation of $\Delta\rho_{\tilde{n},m}$ is just same as $\Delta\rho_{-\tilde{n},m}$ for every $m \in \mathbb{T}$. The calculation is symmetric for symmetric eddy diffusivity.

### 3.4.2 Original- Mirrored domain in SDE

Now we simulate these two diffusivities in SDE with PRS. Here, we will test SDE with PRS using only the E1 scheme, and we evaluate drift coefficient $a$ and diffusion coefficient $b$ analytically. We launch $N_p = 1440000$ particles in the original domain and launch all the particles at position $z = 5\,\text{m}$ to imitate the initial condition in PDE. We use $N_b = 400$ bins for the histogram. For the mirrored domain, we launch $N_p = 2 \cdot 1440000$ particles, half of them at $z = -5\,\text{m}$ and half of them at $z = 5\,\text{m}$. We use $\Delta t = 5\,\text{s}$ and $\mathbb{T} = [0, 1000]$ for both domains. The results are plotted in Fig. 3.7e and Fig. 3.7f. From

Fig. 3.7f, we can see that the result (blue line) is also (statistically) mirrored about position $z = 0$. From Fig. 3.7e, we can see that the PRS at original domain below position $z = 0$ gives visually the same result as the lower half of the domain in Fig. 3.7f, just like Fig. 3.7c and Fig. 3.7d. Beware that the PRS does not apply at $z = 0\,\mathrm{m}$ in mirrored domain in Fig. 3.7f. **Therefore, we conclude that with the PRS at $z = 0$ we get the same results close to $z = 0$ as we do with the mirrored diffusivity.** This also holds for general diffusivity profiles, not just for $K_b$ in the Eq. (3.1b). The area under the blue line has been normalized. For the original domain, it is be normalized to 1. For mirrored domain, it will be normalized to 2.

Even though the results from SDE are almost the same results from PDE, we can still see the boundary artifact in SDE (Fig. 3.7e and Fig. 3.7f). **It is interesting that the boundary artifact is found even in the non-boundary region at $z = 0\,\mathrm{m}$ in the mirrored domain even though the PRS does not play any role at $z = 0\,\mathrm{m}$.** We realize that the problem we have at $z = 0\,\mathrm{m}$ in Fig. 3.7f looks almost identical that the problem we counter at the boundary $z = 0\,\mathrm{m}$ in the Fig. 3.7e. So what is happening?

### 3.4.3 The root of boundary artifact of PRS

We first look at the PDE case. In the mirrored domain, the diffusivity is symmetric ($K_{-\tilde{n}} = K_{\tilde{n}}$). This causes $\rho_{-\tilde{n},m} = \rho_{\tilde{n},m}$ for all $m$. Therefore, at $z = 0\,\mathrm{m}$, we can have $\rho_{-1,m} = \rho_{1,m}$ even though the Neumann boundary condition is not applied at $z = 0\,\mathrm{m}$. Hence, what the Neumann boundary condition actually does in the original domain is to mirror every thing about the boundary at $z = 0\,\mathrm{m}$. At $z = 0$, the $\Delta\rho_{0,m}$ is:

$$\Delta\rho_{0,m} = \Delta t\Big[\frac{(K_1 - K_{-1})}{\Delta z}\frac{(\rho_{1,m} - \rho_{-1,m})}{\Delta z} + K_n\frac{(\rho_{1,m} - 2\rho_{0,m} + \rho_{-1,m})}{\Delta z^2}\Big]. \quad (3.9)$$

We know that $\frac{(K_{+1}-K_{-1})}{\Delta z} = 0$, hence the reader may think that $\frac{(K_{+1}-K_{-1})}{\Delta z}$ is not important since it will multiply with $\rho_{1,m} - \rho_{-1,m} = 0$. But $\rho_{1,m} - \rho_{-1,m} = 0$ is actually due to the symmetric property $K_{\tilde{n}} - K_{-\tilde{n}} = 0$. Therefore, $\frac{(K_{+1}-K_{-1})}{\Delta z} = 0$ is absolutely crucial for the mirrored domain behaving like Neumann boundary condition at $z = 0\,\mathrm{m}$. This inspires us that in order for PRS in SDE to behave like the Neumann boundary condition in PDE,

the derivative of $K$ may have to be zero at the boundary. This is not a mathematical proof, but an inspiration.

Now we look at SDE case, and we realize that $\frac{\partial K_b}{\partial z}$ is not defined at $z = 0\,\mathrm{m}$ in the mirrored domain, see Fig. 3.7b. It does not satisfy the Lipschitz condition (Eq. (2.18)) because

$$\lim_{x\to 0^+, y\to 0^-} |a(x,t) - a(y,t)| = 2\frac{\partial K_b}{\partial z}|_{z=0^+} \neq 0$$

$$\lim_{x\to 0^+, y\to 0^-} L|x-y| = L \cdot 0.$$

$L$ has to be $\infty$ for Eq. (2.18) to hold, and this is not a constant. This violates the assumption that the drift coefficient $a(x,t)$ and diffusion coefficient $b(x,t)$ have to satisfy for Eq. (2.5) to be a diffusion process (refer to Section 2.4). With other words, SDE theory expects that there exists a smooth transition from $\frac{\partial K_b}{\partial z}|_{z=0^+}$ to $\frac{\partial K_b}{\partial z}|_{z=0^-}$, but the derivative of $K_b(z)$ is not continuous at $z = 0\,\mathrm{m}$ in the mirrored domain. This causes a problem in SDE scheme. The reader may wonder why the problem does not appear in the PDE scheme. Following Eq. (3.4a), it turns out it does not matter what $\frac{K_1 - K_{-1}}{\Delta z}$ is since $\rho_{\tilde{n}} = \rho_{-\tilde{n}}$, which makes $\frac{\rho_1 - \rho_{-1}}{\Delta z} = 0$. However, it does not mean that the $\frac{K_1 - K_{-1}}{\Delta z}$ does not have a numerical evaluation, and it is $\frac{K_1 - K_{-1}}{\Delta z} = 0$ in the mirrored domain. **So, it does have a value, and it is zero.** In contrast, the zero can not be sensed by the particles in SDE scheme, first it is not defined, and second the position of a particle is almost surely never $z = 0$. (analytically, numerically the probability is small but finite) At this stage, we have already a lot information. Therefore, we made a mind-map to avoid confusion. The mind-map is showed in Fig. 3.6.

Now we focus on $z = 0^+$ region in both domains and analyze the drift term ($a\Delta t$) and diffusion term ($b\Delta W$) of E1 (Eq. (2.24)) in this case. The sign of the drift coefficient $a = \frac{dK_b}{dz}$ will specify the drift direction since it will multiply to $\Delta t$, and $\Delta t$ is always positive. Consequently, the drift term will tend to move particles to higher diffusivity. In contrast, the diffusion coefficient $b = \sqrt{2K_b}$ is always positive, but the Wiener process $\Delta W$ have 50% chance to be positive and 50% chance to be negative since it is normally distributed with mean 0. Therefore, the sign of the diffusion term will depend on $\Delta W$. Fig. 3.8a shows which term will dominate. When $|\frac{dK_b}{dz}| \gg \sqrt{2K_b}$, the drift term will dominate with high probability; When $|\frac{dK_b}{dz}| \ll \sqrt{2K_b}$, the diffusion term will dominate with high probability. We can see that it forms

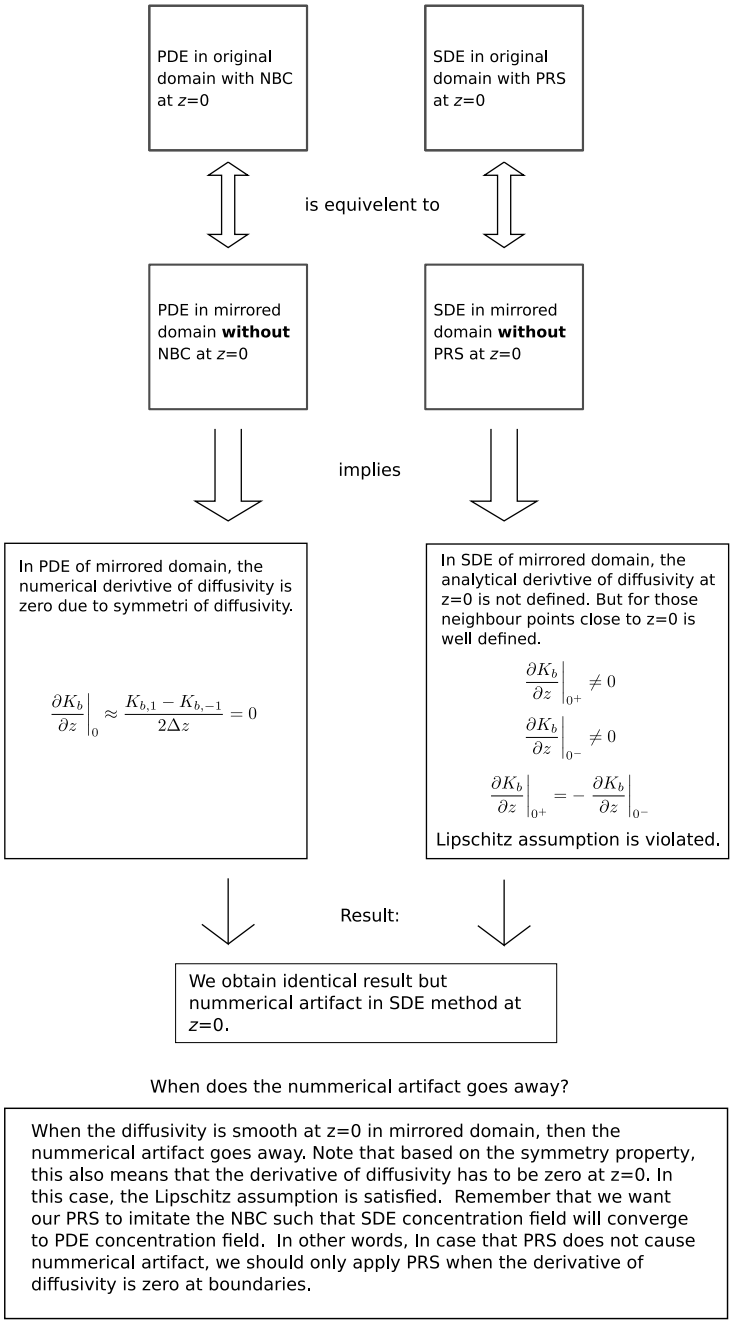
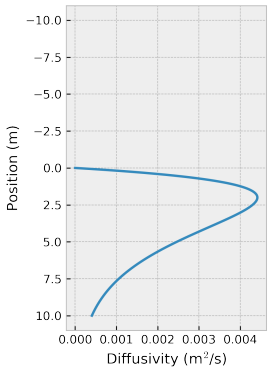
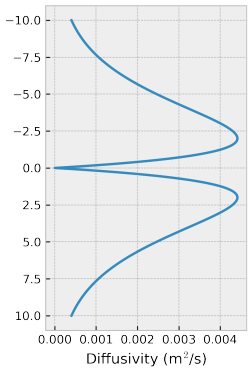


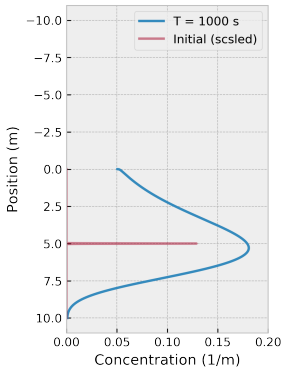

Figure 3.6: A mind map of numerical boundary artifact caused by PRS. The bottom box is actually the conclusion in the Section 3.5.3, which it will come later.



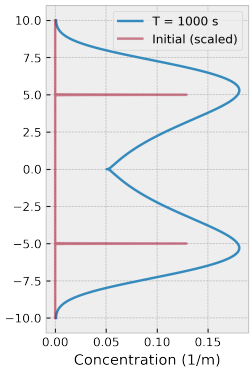

(a) Diffusivity



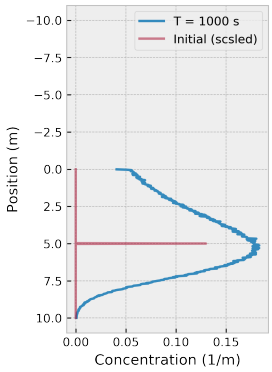

(b) Diffusivity-mirrored



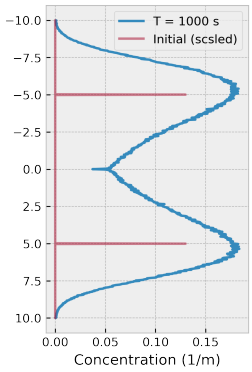

(c) PDE




(d) PDE-mirrored

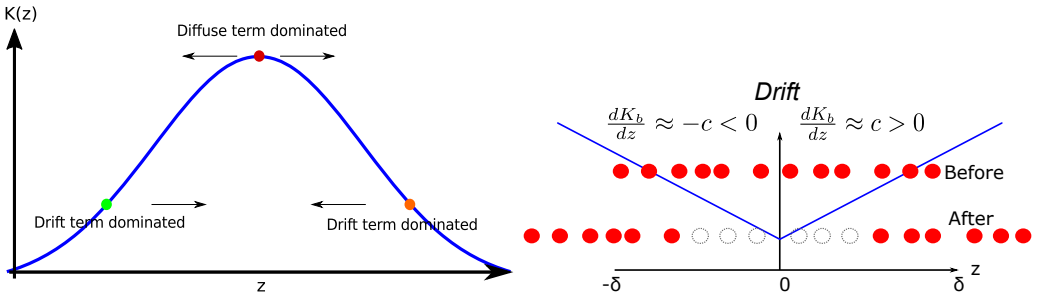



(e) SDE




(f) SDE-mirrored

Figure 3.7




(a) Drift- vs diffuse term (b) Drift effect at $z = 0$

Figure 3.8: Fig. 3.8a shows for a reasonable small diffusivity, where the drift term will dominate and where the diffusion term will dominate. Fig. 3.8b shows that how the drift action takes place in both original and mirrored domain.

a convection pattern to maintain well-mixed states in Fig. 3.8a. Fig. 3.8b shows particles before the drift and after the drift. The circles with dashed line are the particles expected by SDE theory when the Lipschitz condition holds. Obviously, these particles do not exist. Furthermore, the diffusion term can not compensate for this effect. Consequently, this forms lower concentration near the top boundary than at other locations. This obviously does not satisfy the WMC test. We show this in the next paragraph.

In Fig. 3.2, we have performed the WMC test for the original domain. Let us now perform the WMC test for the mirrored domain for the E1 scheme. It has the same configuration as in Fig. 3.7 but uniformly distributed in the initial condition. The result is plotted in Fig. 3.9. As in Fig. 3.7, the original domain and mirrored domain give visually identical results. Solving the problem in the mirrored domain corresponds to solving the PRS boundary condition in the original domain. The green circles indicate the transition point between fluctuation and smooth concentration. Why does this transition exist for the E1 scheme? We will try to answer this question in the next section.

## 3.5 E1 boundary artifact analysis

In this section, we only focus on the original domain, which is $\mathbb{D} = [0, 10]$. The PRS will be applied when position of a particle becomes $Z < 0$ or $Z > H$.

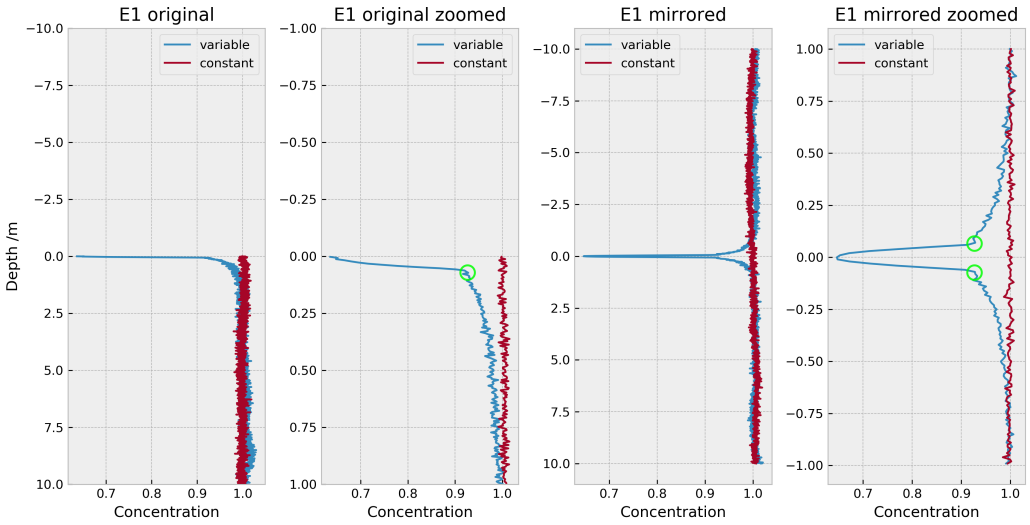



Figure 3.9: WMC test for E1 in original and mirrored domain. It has the configuration like Fig. 3.7 but uniformly distributed in the initial condition. The green circles indicate a sudden drop of concentration.

We are interested in finding out what the maximum value of $\Delta W$ has to be to displace a particle to $Z \leq 0$ near the top boundary. In another word, this is also the minimum value of $\Delta W$ to stay inside the domain ($\mathbb{D} = [0, 10]$). To be precise, we assume that particles near the top boundary will never cross the bottom boundary in a single time-step, since this extreme value of $\Delta W$ is very unlikely. The same argument applies to bottom boundary too. To do this, the E1 scheme Eq. (2.24) becomes

$$Z_{n+1} \leq 0 \implies Z_n + a\Delta t + b\Delta W_n \leq 0$$

and we solve for $\Delta W$

$$\Delta W \leq \frac{-(z + a\Delta t)}{b} = \frac{-(z + \frac{dK}{dz}\Delta t)}{\sqrt{2K(z)}}. \tag{3.10}$$

We define the right hand side of this equation as

$$f(z, \Delta t) := \frac{-(z + a\Delta t)}{b} = \frac{-(z + \frac{dK}{dz}\Delta t)}{\sqrt{2K(z)}} \tag{3.11}$$

where $K$ is a general diffusivity. We call $f(z, \Delta t)$ the "Top boundary quality function". As we will show you later, $f(z, \Delta t)$ will indicate how well

PRS performs at the top boundary to satisfy the WMC. Before we plot this function, we would like to plot the probability density for where a particle crosses from and where a particle lands if the particle crosses over the top boundary. However, it is difficult to calculate such quantities. But we can approach it by performing statistical sampling in all time-steps of simulation. Using the same configuration as Fig. 3.2, but this time, we focus on the concentration between 0 and 0.7 m ($H_h = 0.7\,\mathrm{m}$). I remind the reader that we use $\Delta t = 0\,\mathrm{s}$. The results are plotted in Fig. 3.10. The red line is for constant $K = 0.001\,\mathrm{m}^2/\mathrm{s}$; The blue line is for $K = K_b(z)$. Let's first study the constant diffusivity $K = 0.001\,\mathrm{m}^2/\mathrm{s}$.

### 3.5.1 Boundary artifact of constant diffusivity of E1 scheme

In the constant diffusivity (red line) in Fig. 3.10, the PRS does not cause any boundary artifact, the PDF of crossing and landing are identical if we ignore the random fluctuations (note different scaling in $x$-axis). This means that given a particle has crossed the top boundary, the probability where the particle came from has same probability where the particle will land to. Therefore, for large enough number of particles, the concentration will remain well-mixed. On the other hand, $f(z, 10)$ in this case is

$$\Delta W \leq f(z, 10) = -\frac{z}{\sqrt{2 \cdot 0.001}}.$$

since $\frac{dK}{dz} = 0$. Here $f(0, 10) = 0$. Remember that $\Delta W$ is normally distributed with mean 0 and variance $\Delta t$, in this case is 10 s. For a particle sitting in $z = 0^+$, it has 50% probability of crossing the top boundary. For a particle sitting in $z = 0.1\,\mathrm{m}$, the $f(0.1, 10) = -1.29$, and it has 34% probability of crossing the top boundary. The closer the particle to the boundary, the higher chance to cross the boundary for constant eddy diffusivity.

### 3.5.2 Boundary artifact of non-constant diffusivity of E1 scheme

In the eddy diffusivity, we use $K_b(z)$ in Fig. 3.10 (blue line). The green line is where the $\frac{df(z,10)}{dz} = 0$. From the concentration shown in Fig. 3.10 (blue line, left most panel), we see that the PRS causes a boundary artifact. The PDF

for crossing (second graph) is almost the same as in the constant diffusivity case, but deviates strongly when it gets close to the top boundary. We can see that the probability is actually dropping to zero when it is getting close to top boundary. This happens because the drift term $\frac{dK_b}{dz}\Delta t$ drifts the particles to downwards, and the diffusion coefficient $\sqrt{2K}$ is too small near the boundary. It makes $\frac{df(z,10)}{dz}$ (Eq. (3.11)) very small. For example $f(0, 10) = -41$, and it is rare for $\Delta W \sim \mathcal{N}(0, 10)$ to be lower than -41. The fifth figure shows the derivative of $f(z, 10)$. Here we can see that at $z = 0.05959$ (the green line), this position corresponds to the maximum of $f(z, 10)$ in fourth figure. It is $f(0.05050, 10) = -4.4$. This is unlike for the constant diffusivity, where $f(z, 10)$ increases to zero at the boundary. We can also see that everything seems to go wrong above the green line. This means the region where $\frac{df(z,10)}{dz} > 0$. The green line in the first figure of Fig. 3.10 corresponds a sudden concentration drop for the non-constant diffusivity. The green line in the second figure of Fig. 3.10 indicates where the strong deviation between the red line and the blue line begins. This is somehow an indication that $\frac{df(z,10)}{dz}$ should have a maximum of zero near the top boundary for PRS. Beware that this is not a mathematical proof, but an inspiration.

### 3.5.3 Condition for PRS satisfy WMC

Now, we connect everything that we have learned. Remember that the anomaly of the mirrored domain has the same problem as in the original domain with PRS. The Lipschitz condition from SDE theory expects that we should have smooth transitions of $a(z, t)$ and $b(z, t)$ in the mirrored boundary. Obviously, we do not satisfy this condition as can be seen from Fig. 3.7b. However, if we can alter the diffusivity around $z = 0$, smoothen it so that $\frac{dK_b}{dz} = 0$ (of course, this may change the actual physical problem), the anomaly of the mirrored domain should go away. This means that only the diffusion term will dominate around $z = 0$. We come back to original domain with PRS. We smoothen the diffusivity to get $\frac{dK_b}{dz}|_{z=0} = 0$. This also means that only the diffusion term should dominate. We can rewrite $f(z, \Delta t)$ in the following way

$$f(z, \Delta t) = -\frac{z}{\sqrt{2K}} - \frac{\frac{dK}{dz}\Delta t}{\sqrt{2K}}. \tag{3.12}$$

Recall that $f(z, \Delta t)$ tells us the minimum value $\Delta W$ can take, if the particle is to remain inside the domain. Any $\Delta W < f(z, \Delta t)$ will cross the top

boundary, and any $\Delta W > f(z, \Delta t)$ will stay inside the domain (extreme value of $\Delta W$ is assumed impossible ). If the diffusion term is the dominating term of this process, a particle, which is sitting at the boundary $z = 0$, should have 50% chance to cross the boundary because $\Delta W \leq f(0, \Delta t) = 0$ has 50% chance. This requires the second term of Eq. (3.12) to approach zero near the boundary. But the PDF of $\Delta W$ depends also on $\Delta t$ as well as this therm. We do not want an extra criterion for $\Delta t$ in addition to Eq. (2.29) in Section 2.8. Therefore, we require

$$-\frac{\frac{dK}{dz}}{\sqrt{2K(z)}} \to 0 \text{ when } z \to 0. \tag{3.13}$$

In other words,, $|\frac{dK}{dz}| \ll |\sqrt{2K(z)}|$ when $z \to 0$. We can archive Eq. (3.13) by altering $\frac{dK}{dz}$ to zero near the top boundary, or approach it by using a diffusivity with a higher value of near the boundary. First we test for higher diffusivity, then we test the smoothened diffusivity.

### 3.5.4 Stronger diffusivity $K_c$

Here we use a stronger diffusivity, $K_c(z)$, from Eq. (3.1c) and run the same WMC test. The result is shown in Fig. 3.11. The boundary artifact seems much less significant than in Fig. 3.10, but still there. The right term of Eq. (3.12) is $-\frac{\frac{dK}{dz}}{\sqrt{2K(z)}} = -0.134$ for $z = 0$. For a particle on the position $z = 0$, the probability for $\Delta W \leq -0.134$ is 48.94%. This is still not close enough to 50%. Therefore, the boundary artifact is still visible there. If we add the $K_c$ to a bigger constant, for example $0.1\,\mathrm{m}^2/\mathrm{s}$, the boundary artifact will not be visible in this $H_h$. But it is still there, just hard to notice. We are not going to plot this here. This explains why the boundary artifact at the bottom boundary does not seems very obvious compared to the top boundary in Fig. 3.9, because near the bottom boundary, the drift coefficient $a = \frac{d}{dz}K$ is smaller than at the top boundary, and the diffusion coefficient $b = \sqrt{2K}$ is bigger than at the top boundary. We can define a "bottom boundary quality function" as $f_H(z, \Delta t)$ and see how these two factors make

$$f_H(H, \Delta t)$$

close to zero. The $f_H(z, \Delta t)$ is defined in the similar way as $f$;

$$f_H(z, \Delta t) := -\frac{H - z}{\sqrt{2K}} - \frac{\frac{dK}{dz}\Delta t}{\sqrt{2K}}.$$

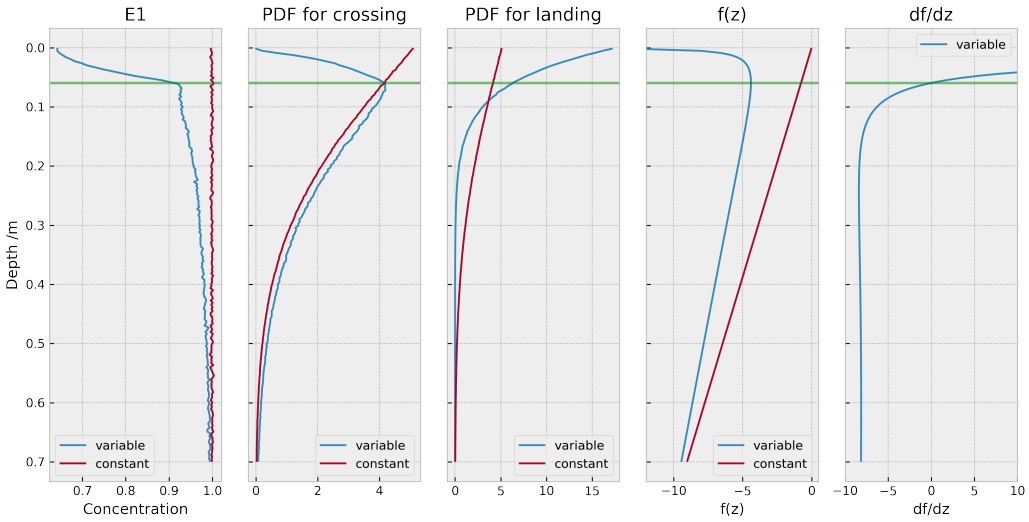



Figure 3.10: Analysis of E1 for $K_b(z)$ and $K = 0.001$ near the top boundary. The first figure (left most) is the concentration of particles. The second figure is the probability density function that given a particle jumps over top boundary, what is the position it jumps from. The third figure is the probability density function that given a particle jumps over top boundary, what is the position it lands on. The fourth figure shows $f(z, 10)$ of Eq. (3.11) and the fifth figure shows $\frac{df(z,10)}{dz}$. The green line is the position $z$ where $\frac{d}{dz} f(z, \delta t = 10) = 0$.

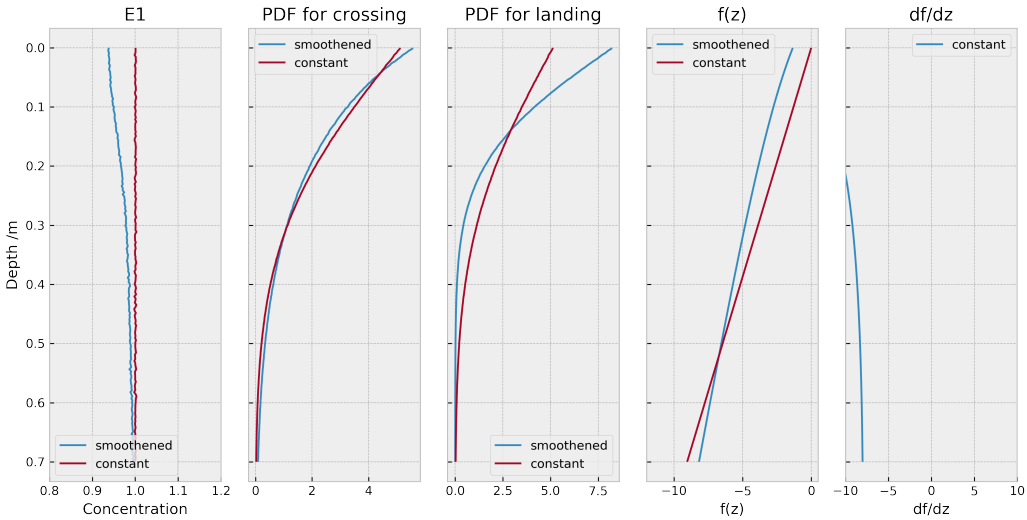



Figure 3.11: Same configuration as Fig. 3.10, but using $K_c$ instead. This time, there is not green line in this figure since the $\frac{d}{dz}f(z, \Delta t = 10) = 0$ has not solution in the domain $[0, 10]$.

The Fig. 3.12 shows the PDF of normal distribution for $\Delta t = 10$ and $\Delta t = 1$.

### 3.5.5 Using smoothened diffusivity $K_b$

Here we try to smoothen $K_b$ near the top boundary such that $\frac{d}{dz}K_b = 0$. This will help us to satisfy the Lipschitz assumption in SDE theory. Therefore, we expect the boundary artifact caused PRS to completely go away. To do this, we define

$$\tilde{K}_b := K_b + \frac{G}{C}(1 - \tanh Cz) \tag{3.14}$$

as the smoothened diffusivity of $K_b$, where the $G$ is the same constant as our diffusivity profile in Eq. (3.1b), and $C$ is a truncation factor to indicate how fast $\tilde{K}_b$ turn to $K_b$. The Fig. 3.13 shows the difference between $\tilde{K}_b$ and $K_b$ near the top boundary for $C = 10\,\mathrm{m}^{-1}$, the rest of the domain is (almost) identical. We run the same test as Fig. 3.10, but we can no longer use $\Delta t = 10\,\mathrm{s}$ due to the criterion in Section 2.8 because for $C = 10\,\mathrm{m}^{-1}$, the

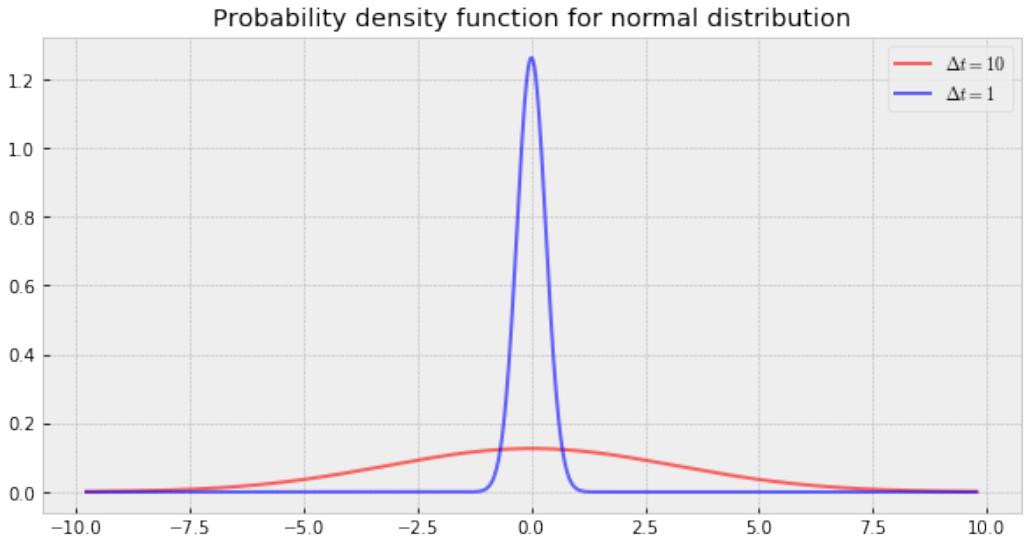



Figure 3.12: Probability density function of normal distribution $\mathcal{N}(0, \Delta t)$ with $\Delta t = 10$ and $\Delta t = 1$.

term $\frac{G}{C}(1 - \tanh Cz)$ dominates in

$$\min\left(\frac{1}{\frac{d^2}{dz^2}\tilde{K}_b}\right) \quad \forall z \in \mathbb{D}.$$

In this case we have

$$\min\left(\frac{1}{\frac{d^2}{dz^2}\tilde{K}_b}\right) = 24.$$

Therefore, we choose $\Delta t = 10\,\text{s}$. We run for same configuration as in Fig. 3.10, except $\Delta t$. The result is shown in Fig. 3.14. We can see that the boundary artifact is no longer visible now. $f(z, \Delta t) = -\frac{\frac{d\tilde{K}_b}{dz}}{\sqrt{2\tilde{K}_b(z)}} = 0$ for $z = 0$. This means that a particle has 50% chance to cross the boundary when it is at position $z = 0$. This can not be achieved by using the same $\Delta t = 0.1$ in $K_b$ in Eq. (3.1b), and this is shown in Fig. 3.16. The red line concentration used smoothened diffusivity ($\tilde{K}_b$), and the blue line concentration used diffusivity ($K_b$). We have used 20000 particles ($Np = 20000$), focused on ($[0, 0.01]$) with 100 bins, $T_{end} = 8604\,\text{s}$ and $\Delta t = 0.1\,\text{s}$. The boundary artifact is very obvious in non-modified diffusivity in all four different schemes even using $\Delta t = 0.1\,\text{s}$.

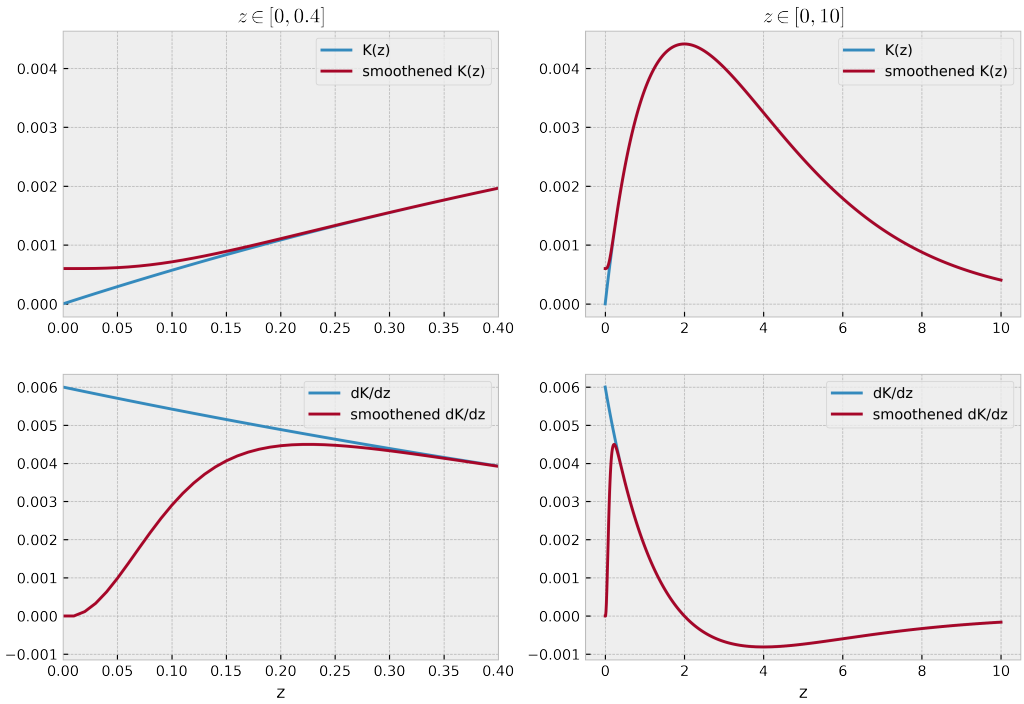



Figure 3.13: These figures show how the diffusivity has been modified such that $\frac{d}{dz}K = 0$ at the boundary (See Eq. (3.14))

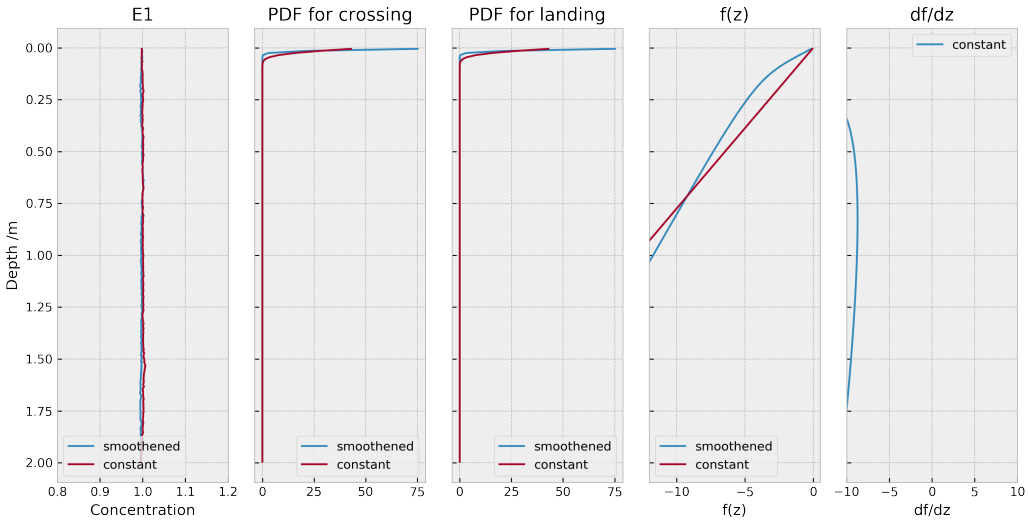



Figure 3.14: Same configuration as Fig. 3.10, but using smoothened diffusivity in Fig. 3.13, $\Delta t = 0.1$ and $N_b = 0.7\,\mathrm{m}$ in stead.

At this stage, we are able to understand what PRS is actually doing in the particle tracking scheme. In a diffusion dominated process, the PRS in the original domain gives the same behaviour as at $z = 0$ in the mirrored domain without PRS, with the same probability. Figure 3.15 illustrates the idea. For the red particle to move from 1 to 3, has the same probability as the gray particle moving from 4 to 5. This effectively mirrors the red particle at 1 in the original domain to 4 in the mirrored domain. If it is correct, the smoothened diffusivity should not only solve the boundary artifact in E1 but also in V1, M1 and M2 since the problem was the Lipschitz condition not being satisfied, and smoothen the diffusivity at boundaries will satisfy it. Fig. 3.16 shows the result. All these four schemes do not exhibit the boundary artifact, and they satisfy WMC as we have expected.

## 3.6 Two different boundary artifacts

From Fig. 3.10, which used diffusivity $K_b$, and Fig. 3.11, which used diffusivity $K_c$, the green line does not exist in Fig. 3.11 because $\frac{d}{dz}f(z, \Delta t = 10) = 0$ does not have a solution in our domain. Even though Fig. 3.11 still exhibits boundary artifact, it is much less significant than in Fig. 3.10. Looking at

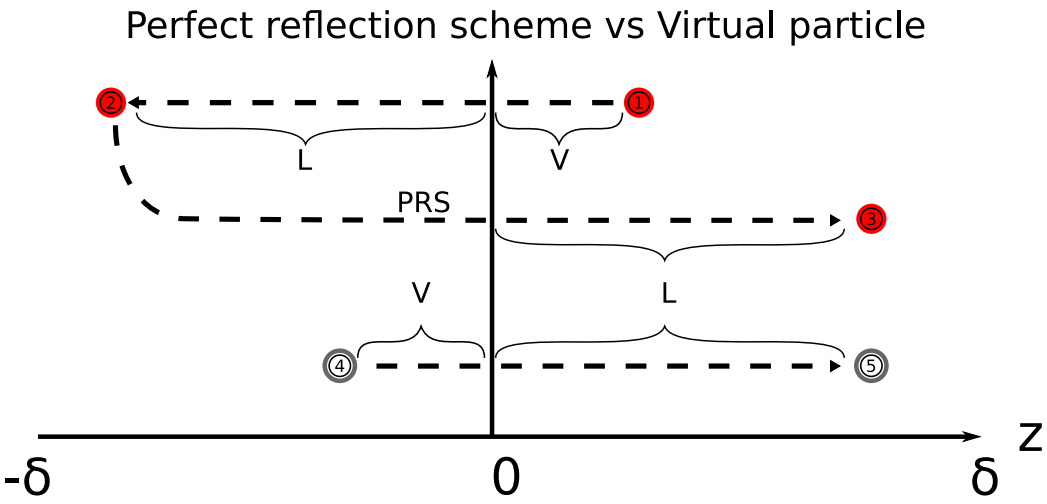



Figure 3.15: In a diffusion dominated process, the PRS in the original domain gives the same behavior as in the mirrored domain without PRS with the same probability.

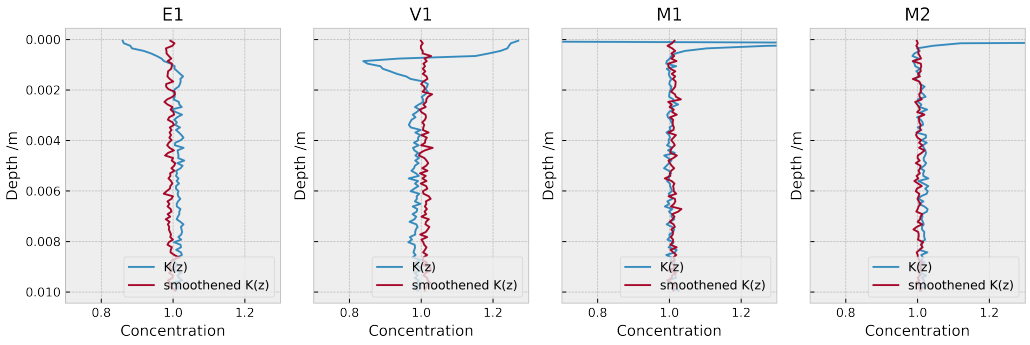



Figure 3.16: These four figures shows how the smoothened diffusivity performs with four different schemes.

the PDF for crossing in Fig. 3.11, there is no strong deviation between $K_c$ and constant diffusivity. In contrast, looking at the PDF for crossing in Fig. 3.10, there is a strong deviation between $K_b$ and constant diffusivity. Particles actually have less chance to cross the top boundary when they are very close to top boundary. This happens above the green line, which indicates $\frac{d}{dz}f(z, \Delta t = 10) > 0$. Perhaps, the green line, which is the derivative of top boundary quality function, can be our indicator to classify two different boundary artifacts in E1 scheme with PRS. Here we classify them in the following way:

*Boundary artifact type A*

Particles have less chance to cross the boundary when they are getting close to the boundary.

*Boundary artifact type B*

Particles **do not** have less chance to cross the boundary when they are getting close to the boundary.

For the top boundary, boundary artifact A happens when $\frac{d}{dz}f(z, \Delta t = 10) > 0$. Recall the physical meaning of top boundary quality function ($f(z, \Delta t)$) is the minimum of $\Delta W$ needed to cross the top boundary at a given position $z$. Any $\Delta W$ below $f(z, \Delta t)$ will cross the top boundary. If $\frac{d}{dz}f(z, \Delta t) > 0$ and $f(z, \Delta t) < 0$ at near top boundary region, $f(z, \Delta t)$ is more and more negative when $z \to 0$. Recall that $\Delta W$ is a Wiener process, which is stochastic. Therefore, boundary artifact A occur. We know that type B has less significant boundary effect than type A, and the different between $K_b$ and $K_c$ is only a constant. Therefore, we are looking for the constant where transition from type A to type B occurs at position $z$. Here, we rewrite $f$ in the following way:

$$f(z, \Delta t) = -\frac{(z + a\Delta t)}{b}.$$

The $\frac{d}{dz}f(z, \Delta t)$ is

$$\frac{d}{dz}f(z, \Delta t) = -\frac{b - zb'}{b^2} - \frac{a'b - ab'}{b^2}\Delta t,$$

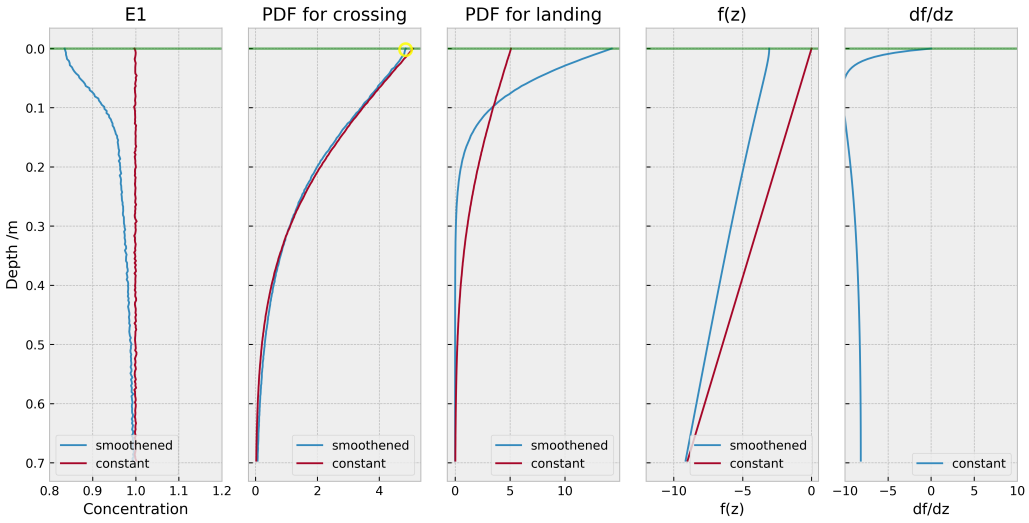



Figure 3.17

where $a'$ and $b'$ mean $\frac{\partial}{\partial z}a$ and $\frac{\partial}{\partial z}b$. Based on the definition of $a := \frac{d}{dz}K(z)$ and $b := \sqrt{2K(z)}$, we found $b' = -\frac{a}{b}$. For $\frac{d}{dz}f(z, \Delta t) = 0$, we obtain

$$b^2 = \frac{a(z + a\Delta t)}{a'\Delta t + 1}.$$

We know that $b^2 = 2K$ and let $K = C + g(z)$. In our diffusivity, $g(z) = 6 \cdot 10^{-3} z e^{-\frac{z}{2}} \mathrm{m}^2/\mathrm{s}$. Therefore, we obtain

$$C = \frac{a(z + a\Delta t)}{2(a'\Delta t + 1)} - g(z).$$

Now, we have the formula for the transition constant $C$. We want the transition to happen at $z = 0$. In our case, $C = 0.000\,19\,\mathrm{m}^2/\mathrm{s}$. We can see that $C$ lays between $K_3$ and $K_2$ (Eq. (3.1)). Let us run the the same WMC analysis as before for this constant. Fig. 3.17 shows the result. If we look at the PDF for crossing, we can see that the derivative of the blue line, indicated by a yellow circle, is close to zero at the boundary. The boundary artifact is still there, but only type B, not type A. When $f(z = 0, \Delta t) = 0$, the boundary artifact will go away as we have discussed previously. Here, we may notice the importance of $f$. A similar function can be define at the bottom boundary or on another scheme. For example, for the M1 scheme $f$ will become

an implicit function due to the $\Delta W$ in the drift coefficient. Here, we skip V1, M1 and M2 scheme analysis since we have enough information to give a suggested solution.

## 3.7 Legitimization of changing diffusivity near boundary

Can we change the eddy diffusivity near boundary without changing the physical reality? The logical answer is NO, obviously. But in practice, it depends how we obtain the eddy diffusivity profile, and how well we know the statistics at the boundary. Following Wilson and Flesch (1993), in atmospheric theory, normally the statistics of diffusion near the ground is unknown. They call it "unresolved basal layer" (UBL). Therefore, one is free to design the statistics of eddy diffusion near the boundary. Here, we are not sure we can use UBL as a reason to change the diffusivity at the top boundary, even though the height of sea surface is usually fluctuating. It is difficult to obtain exact height and eddy diffusivity. We may argue that the measurement of eddy diffusivity is usually macroscopic and time-average. A slight change, we guess, should not alter the physical reality too much, especially when we treat the non-stationary surface as a stationary hard wall. In ocean turbulence environment, the height of sea surface can not be stationary. However, the PRS has been used for many years and in many research fields due to its simplicity. More practical information of measurements is given in the Chapter 5. Finally, if we can assume that we have the right to design the top boundary diffusivity, the problem is solved here. However, we are still not sure if it is legitimate to do this, even though Wilson and Flesch (1993) indicate it is normal in atmospheric theory. For the rest of the thesis, we will try solve the boundary issue from three different perspectives. First is a mathematical solution, which is trying to find SDE theory that can have PRS embedded. Second is the engineering solution, where we assume the diffusivity is coming from measurements, and approach the problem by using cubic-spline interpolation with "clamped" boundary condition. Third is the numerical solutions, where we will try approach the problem without changing the diffusivity near boundary.

## 3.8 Summary

We give a summary to this chapter because this chapter is crucial to our later work. We have shown that the Lipschitz condition does not be satisfied in the mirrored domain and where it cause numerical artifact . We have also shown that the original domain with PRS has identical result in mirrored domain. If we change the diffusivity about the mirror position in the mirrored domain such that the Lipschitz condition has be satisfied, then numerical particle goes away in mirrored domain, and the PRS will no longer cause boundary artifact in original domain with this changed diffusivity. Therefore, for PRS satisfy the WMC, we need the drift coefficient $a$ equal to zero and diffusion coefficient $b$ can not be zero at the boundary.

# Part II

# Suggested solutions

# Chapter 4

# Mathematical approach

In this chapter, we will look for any reflection scheme to satisfy the WMC without any restrictions. We first argue that if we assign a particle to the domain $[0, H]$ with uniform distribution $\mathcal{U}[0, H]$, the probability for the position of the particle should stay uniform distribution $\mathcal{U}[0, H]$. Then, Wilson and Flesch (1993) use this idea to show there does not exist such a reflection scheme.

## 4.1 Existence and uniqueness

In the Chapter 3, we did not introduce any rigorous way to decide whether the PRS with an SDE scheme satisfies the WMC, except looking at the histogram and make decision. This is generally good enough. However, there are many ways to perform a rigorous and scientific measurement to detect whether WMC has be satisfied. For example, we define $\sigma$ is such way,

$$\sigma^2 = \sum_{i=1}^{N_b} (C_i - 1)^2,$$

where $C_i$ is the concentration of our $i$ sub-interval in the histogram, $N_b$ is number of sub-interval of interesting of domain, as usual, $\sigma^2$ will be our indicator parameter how the concentration deviate from 1. We can define a threshold $\sigma_t^2$ such that if $\sigma^2 < \sigma_t^2$, WMC is satisfied. If $\sigma^2 > \sigma_t^2$, WMC is no satisfied. The downside of the method is troublesome to define $\sigma_t^2$. A graphical result is much better, as we have done so long. Even though this is usually good enough, we still introduce a even more rigorous way to do this.

Let us consider that we launch a large enough number of particles with uniform distribution $\mathcal{U}[0, H]$ over the domain $[0, H]$ and normalize the histogram. We should obtain a uniform distribution $\mathcal{U}[0, H]$ if the number of particles we launched $N \to \infty$. This is same as, we launch a single particle with $\mathcal{U}[0, H]$. The two cases are the same. Since the two cases are the same, we focus on one single particle. The position of the particle is denoted $Z(t)$ and it is subject to SDE schemes. Obviously, $Z(t)$ is random variable. Let us denote $\rho(z, t)$ as the probability density function of $Z(t)$. Here, $\rho(z, 0)$ is $\mathcal{U}[0, H]$. For the WMC has to be satisfied, the $\rho(z, t)$ has to equal $\mathcal{U}[0, H]$ for all $t$. Therefore, it also has to hold for one small time increment $\Delta t$. Wilson and Flesch (1993) utilize this idea, combined with the Chapman-Kolmogorov equation, to test the PRS and other reflection schemes. In their article, they consider an SDE for the velocity instead of position. But this is very similar to our case. Because for small $\Delta t$, the change of velocity (denoted $\Delta V$) can be considered as linear respect to $t$. Therefore, a displacement ($\Delta Z$) is just

$$\Delta Z = \int_{\Delta t} V dt \approx \Delta V \frac{\Delta t}{2} + V \Delta t.$$

This means that for a given $\Delta V$ and $V$, we can always convert it to $\Delta Z$. Therefore, their results also apply to our case. The reader may think there exists a reflection scheme that in theory, can satisfy WMC for general eddy diffusivity $K$. Wilson and Flesch (1993) had the same idea. The answer they obtained is surprising. They state:

> "We have proven that perfect reflection is in restricted circumstances rigorous, but that where it fails, no alternative and rigorous reflection scheme exists. This is not to say there are not schemes that are *better* than perfect reflection, and that for suitably limited time steps might be acceptable in practice." (Wilson & Flesch, 1993)

The "restricted circumstances" in the quotation is a similar situation to what we have discussed in Chapter 3, which is $\frac{d}{dz}K = 0$ at boundaries. On onther words, the Lipschitz condition is satisfied. Surprisingly, according to Wilson and Flesch (1993), there does not exist a mathematically rigorous reflection scheme for general eddy diffusivity. Therefore, we can give up to find an analytically reflection scheme for general $K$. But, perhaps, does this mean that a correct diffusivity should have $\frac{d}{dz}K = 0$ at boundaries in hard-wall

condition? This remains a puzzle for me. Beware that all we have done in Chapter 3 is just saying that in order to make sure PRS with SDE performs like Neumann boundary condition with PDE, the derivative of diffusivity at boundaries has to be zero ($\frac{d}{dz}K = 0$). This does not mean a correct diffusivity has $\frac{d}{dz}K = 0$ at boundaries. I hope someone can investigate it further.

# Chapter 5

# Engineering solution

In Section 3.7, we have argued that the height of the sea surface is fluctuating. This means that this boundary does not have an exact height in turbulent environments. It may first appear to be surprising that many researchers treat the sea surface as hard wall. Obviously, a fluctuating sea surface is not a hard wall, But it is still commonly used. The first reason is that the hard wall model is simple and still gives accurate enough result. The second reason is that even if the height of the sea surface fluctuate, the expectation value of the height is still stable in the time domain of interest. However, this is much beyond our scope. Eddy diffusivity can sometimes be derived from turbulence theory, or it can be measured. Let us discuss measurements. In practice, we can not measure diffusivity directly because the diffusivity is a statistical concept. Here, we take constant diffusivity as an example. In this case, the drift term is zero, but the diffusion is not zero and the direction is random. This means a constant diffusivity does not have a direction. In measurements, we can only measure the velocities of the fluid flow. In this case, diffusivity can be obtained by serial of measurements of the velocity field at a stationary point. We repeat it with many points in vertical water column. Afterwards, we can calculate the diffusivity by transport theory. Therefore, we obtain discrete value of $K$. A typical ocean velocimeter apparatus can be, for example, Vectrino from Nortek [1]. In practice, we can not measure the velocity exactly on sea surface. Usually, we measure the velocity a little bit under the sea surface where the apparatus is submerged in the ocean during the fluctuating height of the sea surface.

[1] https://www.nortekgroup.com/products/vectrino

## 5.1 Smoothened diffusivity method

We will here consider the smoothened diffusivity. This method is already presented in Section 3.5.5. For the sake of completeness, we will briefly present it again, since it is apparently a good solution.

We expect the boundary artifact caused by PRS will completely go away if $\frac{d}{dz}K = 0$ at the boundary. To do this, we define the $\tilde{K_b} := K_b + \frac{G}{C}(1 - \tanh Cz)$ as the smoothened diffusivity of $K_b$, where the $G$ is the same constant as our diffusivity profile in Eq. (3.1b), and $C$ is the truncation factor to indicate how fast $\tilde{K_b}$ turns to $K_b$. Fig. 3.13 shows the difference between $\tilde{K_b}$ and $K_b$ near the top boundary for $C = 10$, the rest of the domain is (almost) identical. We run the same test as Fig. 3.10,but with a shorter time step $\Delta t = 0.1\,\mathrm{s}$ (see discussion in Section 3.5.5). The result is shown in Fig. 3.14. We can see that the boundary artifact is no longer visible now. $f(z, \Delta t) = -\frac{\frac{d\tilde{K_b}}{dz}}{\sqrt{2\tilde{K_b}(z)}} = 0$ for $z = 0$. This means that the particle has 50% chance to cross the boundary when it is at position $z = 0$. This can not be achieved by using the same $\Delta t = 0.1$ in $K_b$ in Eq. (3.1b), and this is shown in Fig. 3.16. The red line concentration used smoothened diffusivity ($\tilde{K_b}$), and the blue line concentration used diffusivity ($K_b$). We have used 20000 particles ($Np = 20000$), focused on ($[0, 0.01]$) with 100 bins, $T_{end} = 8604\,\mathrm{s}$ and $\Delta t = 0.1\,\mathrm{s}$. The boundary artifact is very obvious in non-modified diffusivity in all four different schemes even using $\Delta t = 0.1\,\mathrm{s}$.

The drawback of this method is that the diffusivity near the top boundary is completely altered, as we can see in Fig. 3.13. Even though we have mentioned that the "Unresolved basal layer" gives us some freedom to alter the diffusivity near the boundary, someone may still not want to change (or change too much) the diffusivity profile around the boundary as we did in the smoothened diffusivity. Therefore, in the next section, we will try to keep the diffusivity mostly unchanged around the top boundary (at least keep the $K$ at the boundary), while we can alter the derivative of the $K(z)$.

## 5.2 Cubic spline interpolation method

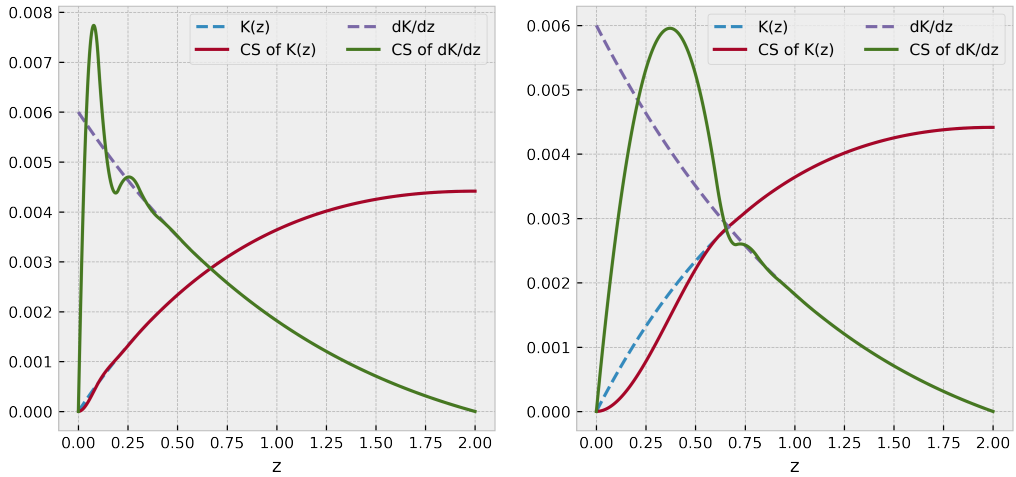
In this section, we assume the we have obtained the diffusivity from measurements. The diffusivity that we have obtained is given at discrete points. Let's denote it $K_n$, where $n \in [0, 1, 2, ..., N]$ is the index of the discrete posi-

tion, and $N+1$ is the total number of discrete points. With discrete data for $K$, there is no problem for the Eulerian PDE method since it is essentially finite difference method. But in the Lagrangian SDE method, the position of a particle in the next time frame depends on a random process $\Delta W$. This means that to calculate the position of next time frame, we need the $K$ and $\frac{dK}{dz}$ continuously since particles can have any position between 0 and $H$. Therefore, we need to interpolate our discrete data, $K_n$. We require the interpolation to have the following properties.

- The interpolation scheme should avoid oscillation between data points because the the SDE schemes depend on $\frac{dK}{dz}$.
- Easy to decide the derivative of the start point and end point (boundaries points). In another words, we can easily decide $\frac{dK}{dz}$ at the boundary points.

Here, **the cubic spline interpolation with clamped boundary condition** satisfies our needs. The Python package called Scipy, has a class for cubic spline interpolation. Here, we will utilize it. For the formulation of clamped cubic spline, please refer to Sauer (2011, chapter 3.4). Following the analysis we have given in Section 3.4, the clamped condition should eliminate the boundary artifact. The clamped condition means that derivative of the start point and end point is zero. Analogically, we clamp something on the end of the table, then it is parallel to the table.

Let us test the idea. Of cause, first, we discretize the diffusivity $K_b$ with 100 points. We obtain $K_{n,b}$, where $b$ indicates the we are using $K_b$. We perform the clamped cubic spline on the data set $K_{n,b}$. It produces a function of $z$. We denote it $K_{cs}(z)$, where $cs$ indicates it comes from cubic spline. Fig. 5.1a shows the difference between $K_b$ and $K_{cs}$ and the derivatives of $K_b$ and $K_{cs}$ near the boundary condition. However, remember that the size of $\Delta t$ depends on the second derivative of $K(z)$. If we do not remove any data, the derivative of $\frac{dK_{cs}}{dz}$ will change dramatically between the first two points which is $K_{0,b}$ and $K_{1,b}$. Therefore, we try to remove some points around the boundary but keep the boundary point. Let us try to remove $n \in [1, 2, 3, 4, 5]$, then we hold $n \in [0, 6, 7, ...99]$. This means that we still keep the starting point. The plot is shown in Fig. 5.1b. Looking at the red line compared to the dashed blue line, we hope the reader may agree that we have not changed too much of the diffusivity. Beware that we deleted the five data points near the boundary only because we want to avoid using

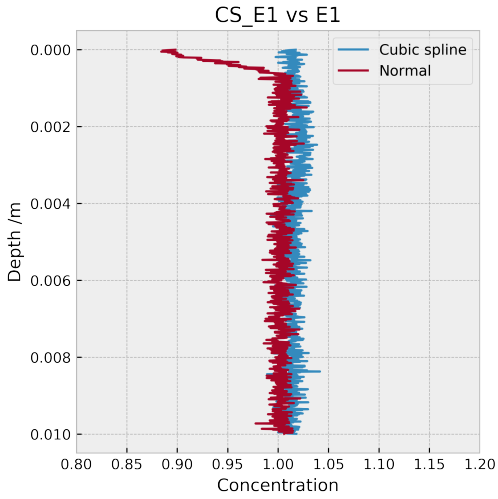



(a) Cubic spline with no removed points (b) Cubis spline with 5 removed points

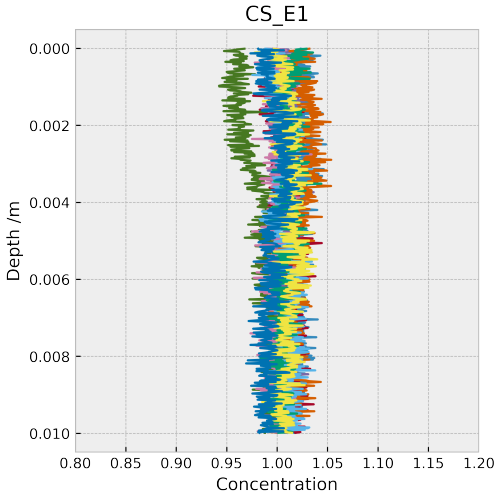
Figure 5.1

very small $\Delta t$. The reader can keep all the data points if using very small $\Delta t$ is not a problem. For example, using a super computer. Before we run the simulation based on cubic spline interpolated diffusivity, we will only focus on domain $z \in \mathbb{D} = [0, 2]$ because $\frac{d}{dz}K_b = 0$ at $z = 2$. This means that there should not be any boundary effect around the boundary at $z = 2\,\text{m}$.

This time we use $H = 2\,\text{m}$, $\Delta t = 0.05\,\text{s}$, $N_p = 20000$, $H_h = 0.01\,\text{m}$, $N_b = 1000$ and $T_{end} = 10\,800\,\text{s}$. $K(z)$ and $\frac{dK}{dz}$ in the E1 scheme we use is from cubic spline interpolation. The result is plotted into Fig. 5.2a. The blue line is using the cubic spline interpolation of the diffusivity as in Fig. 5.1b. The red line is using the unchanged diffusivity $K_b$ as in Fig. 5.1b. The results shown by the blue line are better than the red line as we expected. However, the reader may want to point out that there are some fluctuations for the blue line near the boundary in Fig. 5.2a. This is caused by low drift term and low diffusion term. It takes much longer time to diminish the fluctuations than for the unchanged $K_b$. So, there is not anything wrong with cubic spline interpolation, we just have to use much bigger $T_{end}$ to diminish the fluctuation. To verify this, we actually do not have to run very long $T_{end}$, all we need is to just repeatedly simulate a few diffusion processes under the same conditions with cubic spline interpolation. We use the same configuration as previously, but $T_{end} = 3600\,\text{s}$. We simulate it 10 times. Fig. 5.2b shows the




(a) Cubic spline interpolation diffusivity vs Normal diffusivity

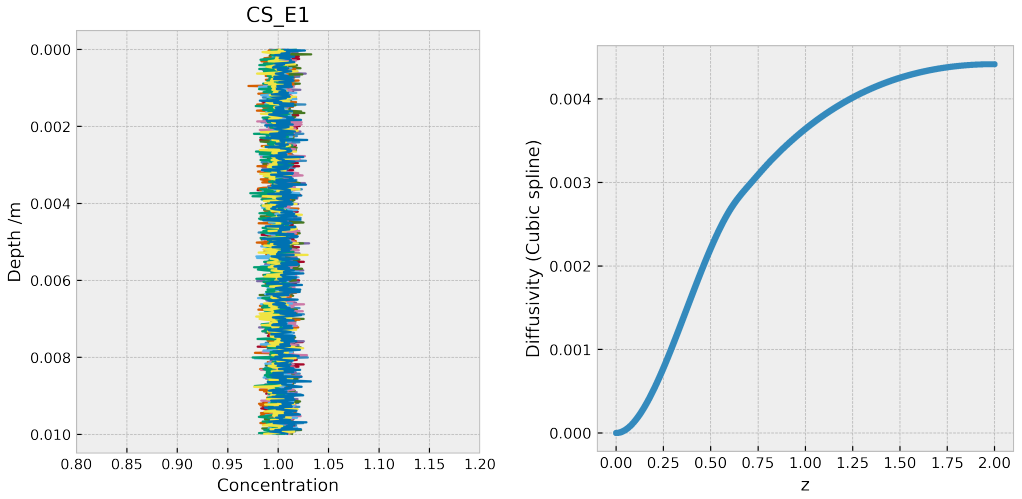



(b) A cubic spline simulation under the same configuration as Fig. 5.2a, but simulate 10 times and using $T_{end} = 3600\,\mathrm{s}$ and $K_c$

Figure 5.2

result. We can see that they fluctuate around concentration 1. This means if we simulate it long enough, the fluctuations will go away. Since we claim the fluctuation is due to both drift and diffusion coefficients being small at the boundary comparing to maximum in the water domain. This should imply that if we use a stronger diffusivity with cubic spline, the fluctuation should goes away. Let us perform the same test as in Fig. 5.2b but using $K_c$ (Eq. (3.1c)) with cubic spline applied. Remember that $K_c = K_b + 0.001\,\mathrm{m}^2/\mathrm{s}$. The result shows in Fig. 5.3a. Obviously, the fluctuation goes a way. There is still a problem, how do we know that there is not oscillation in the cubic spline model? Fig. 5.1b may not have good enough resolution of $z$ axis to show the oscillation. To show the cubic spline does not have oscillation, we distribute 20000000 particles with uniform random distribution and use these large amount random position to plot out the cubic spline diffusivity. If there is any oscillation, there is very good chance that we will observe some some discontinuous poits. Fig. 5.3b shows the result. The figure suggests that there is not oscillation in diffusivity generated by Cubic spline.

This does not just work on the E1 scheme but also on V1, M1 and M2.




(a) A cubic spline simulation under the same configuration as Fig. 5.2a, but simulate 10 times and using $T_{end} = 3600\,\mathrm{s}$ and $K_c$

(b) 20000000 particles uniformly random distributed in water domain and use these large amount random positions to plot the cubic spline diffusivity with $K_b$. Figure seems very smooth, not jumping points.

Figure 5.3

The plot is in appendix Fig. A.1. We can see that the cubic spline interpolation method performs very well in the others three schemes as well. We believe cubic spline interpolation method should also perform well in other SDE schemes with PRS, since the root of the problem is the violation of the Lipschitz condition for the drift coefficient $a(z)$. When the Lipschitz condition at boundary is satisfied, there should not be any reason for PRS to cause boundary artifacts.

The reader may wonder how the cubic spline approach performs if $K = K_c$, i.e. for a diffusivity with a larger constant term. The answer is "not worth it". Surely, there will be a boundary artifact around the boundary if we just use the non-modified diffusivity, but since the denominator in Eq. (3.12) $|\sqrt{2K_c(0)}|$ is much bigger than $|\frac{dK_c}{dz}|_{z=0}$, the $f(z, \Delta t)$ in Eq. (3.12) is close to zero for the diffusivity given by $K_c$. This means that the boundary artifact will be extremely small. This can probably be ignored for most applied research areas. If we use cubic spline in this case, We have to use much smaller $\Delta t$. It is usually not worth to do so.

# Chapter 6

# Numerical solution

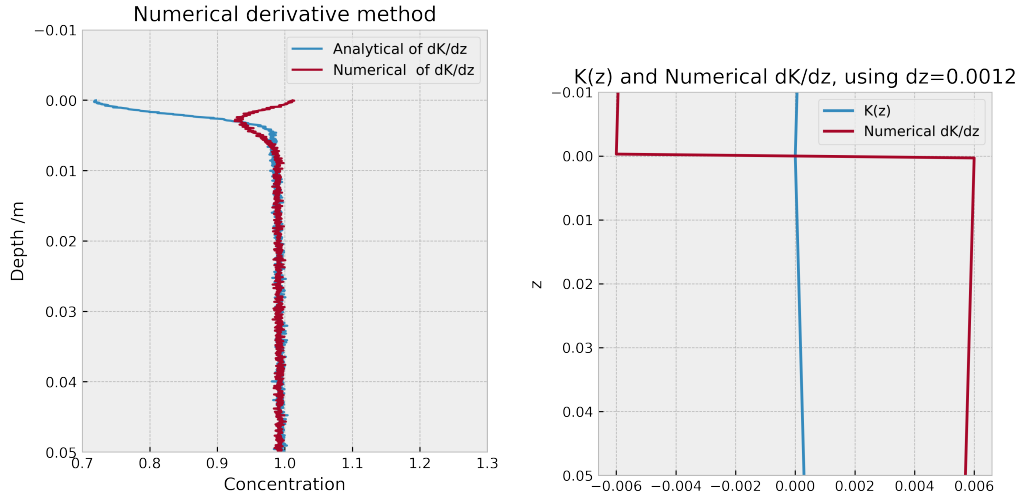
In this section, we will not change the diffusivity, but instead look for another numerical ways to minimize the boundary artifact. However, none of the methods below work better than those presented in Chapter 5.

## 6.1 Numerical derivative method

In this method, we still utilise what we have learnt from Chapter 3. We use a numerical derivative to define the drift coefficient $a(z)$ instead of analytical as before. The drift term $a$ in the E1 scheme becomes the following:

$$a(z, dz) = \frac{K(z + dz) - K(z - dz)}{2 \cdot dz} \tag{6.1}$$

If $K(z)$ needs to be evaluated outside the domain, we simply just mirror it. This may sound counter-intuitive. How may a numerical derivative be a better choice than a analytical derivative. But in this way, we have $a(0, dz) = 0$ which satisfies the condition by conducting a WMC test as before. Let us try to find out how this approach performs in the E1 scheme. We use $H = 2\,\mathrm{m}$, $\Delta t = 0.5\,\mathrm{s}$, $N_p = 20000$, $H_h = 0.05\,\mathrm{s}$, $N_b = 1000$, $dz = 0.0012\,\mathrm{m}$ and $T_{end} = 86\,400\,\mathrm{s}$ . The result is plotted in Fig. 6.1a for $K = K_b$. It does not work perfectly, but it reduces the boundary artifact effect somewhat. The reason it does not work perfectly is obvious in Fig. 6.1b. It indicate that we should use $\Delta t$ extremely small because there exists a point where the double derivative of $K(z)$ is very huge. Not just that, to pick a suitable $dz$ can be challenging. A suitable $dz$ depends on $\Delta t$ and $K(z)$. Generally, this is not a good way to tackle the boundary artifact problem.




(a) Numerical derivative method with $dz = 0.0012$

(b) Diffusivity profile of $K_b$ and numerical $\frac{dK_b}{dz}$

Figure 6.1

## 6.2 Naive time adaptive method

We can minimize the boundary artifact by reducing the time step. However, this will use a lot of computational force. But how about we only reduce the time-step for the particles near the boundary. For any particle in any position, given a reasonable small $\Delta t$, due to normal distribution, there will be a limited interval that we are 99.999% sure that a given particle can move to. Therefore, we separate our water column in two parts. We name the region near the boundary condition as "EDGE", and the region not near the boundary condition as "BODY". Assuming we have water column depth 2 meter, let's denote the EDGE interval as $\mathbb{E} := [0, 0.2] \cup [1.8, 2]$ and the BODY interval is denoted as $\mathbb{B} := (0.2, 1.8)$. We define $\Delta t_{BODY}$ as the time increment for particle position inside the BODY set, which means $Z \in \mathbb{B}$, and define $\Delta t_{EDGE}$ as the time increment for particle position inside the EDGE set, which means $Z \in \mathbb{E}$. Let $n$ be an integer and $n \in \mathbb{N}$. Letting the $\Delta t_{EDGE} = \frac{\Delta t_{BODY}}{n}$ we can have smaller or equal time increment at the EDGE than BODY.

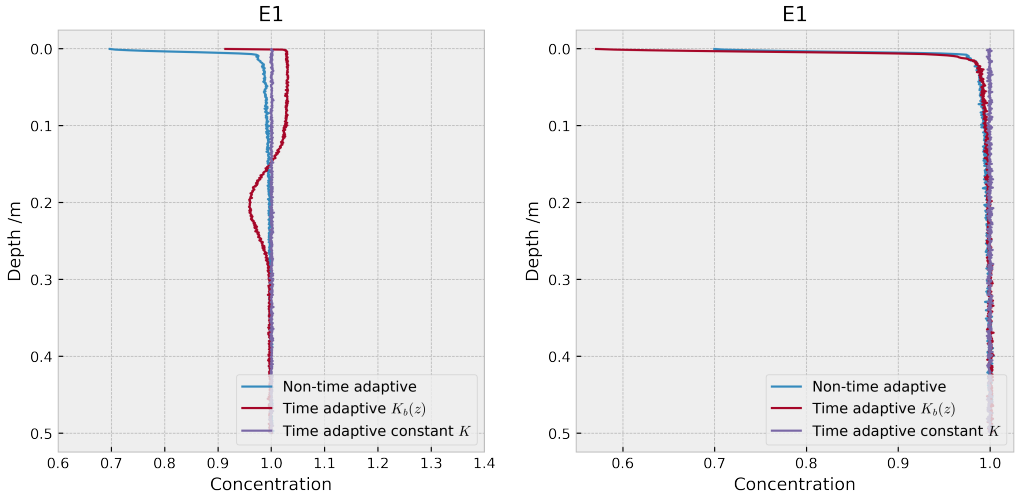
The pseudo code for this scheme is the following:

**Algorithm 1:** This snippet of code shows the idea of how we perform a single time step

**Result:** Perform a single time step
**if** $Z \in \mathbb{B}$ **then**
    /* Update the $Z$ with $\Delta t_{BODY}$ */
    $Z = Z+$ Numerical_scheme$(Z, \Delta t_{BODY})$;
**else**
    /* Update the $Z$ $n$ times with $\Delta t_{EDGE}$ */
    **for** *Looping $n$ times* **do**
        $Z = Z+$ Numerical_scheme$(Z, \Delta t_{EDGE})$;
    **end**
**end**

Let us test it with E1, we use $H = 2\,\text{m}$, $\Delta t = 1\,\text{s}$, $N_p = 80000$, $H_h = 0.5$, $T_{end} = 14\,400\,\text{s}$, $N_b = 600$, $\mathbb{E} = [0, 0.2]$, $\mathbb{B} = (0.2, 2]$, $n = 10$ and $K = K_b(z)$. The result is plotted in Fig. 6.2a. The blue line is ordinary E1; the red line is E1 with naive time adaptive method, and E1 with constant diffusivity using naive time adaptive is the purple line. The naive time adaptive method seems to help little bit, but it creates its own problem. We have a lower concentration than 1 at $z = 0.2\,\text{m}$, and a higher concentration than 1 in most of the $\mathbb{E}$ region. At the boundary we have concentration lower than 1. But this method works perfectly for constant $K$. Intuitively, this should also work for non-constant diffusivity. Because, let us say $n = 2$, we have random variables $X \sim N(0, \sigma_x^2), Y \sim N(0, \sigma_y^2)$. The distribution of $X + Y$ is $N(0, \sigma_x^2 + \sigma_y^2)$. In our case, we have $\sigma_x^2 = \sigma_y^2 = \Delta t/2$. However, this is not the case for the non-constant diffusivity profiles. The drift and diffusion coefficients depend on the previous position. In a constant diffusivity profile, it is just a constant. Therefore, this method did not work well for the non-constant diffusivity. But it works in constant $K$. However, it is trivial.

What happens if we only use the drift and diffusion coefficients at the starting point to evaluate all the sub time-step in Edge. This means:




(a) Naive time adaptive method

(b) Naive time adaptive method, Fixed evaluation for drift term and diffuse term.

Figure 6.2

**Algorithm 2:** This snippet of code shows the idea of how we perform a single time step

**Result:** Perform a single time step
**if** $Z \in \mathbb{B}$ **then**
  `/* Update the` $Z$ `with` $\Delta t_{BODY}$ `*/`
  $Z = Z+$ Numerical_scheme$(Z, \Delta t_{BODY})$;
**else**
  `/* Save the starting position in EDGE */`
  $Z_c = Z$ `/* Update the` $Z$ $n$ `times with` $\Delta t_{EDGE}$`. Only use` $Z_c$ `in the numerical scheme. */`
  **for** *Looping* $n$ *times* **do**
    $Z = Z+$ Numerical_scheme$(Z_c, \Delta t_{EDGE})$;
  **end**
**end**

Using same configuration as previous. The result is plotted as a red line in Fig. 6.2b. The weird behavior at $z = 0.2\,\text{m}$ is gone. This confirms our idea, but the boundary artifact is actually getting worse. However, the reason is

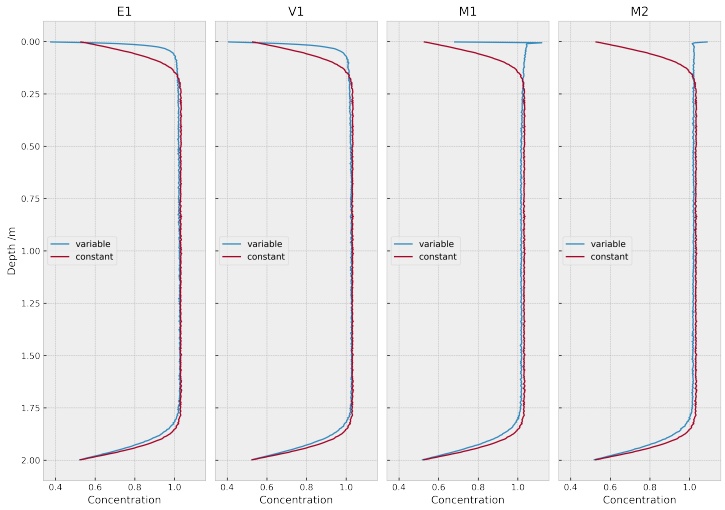



Figure 6.3: Result of bias wiener method. Whenever a particle crosses the boundaries, we re-pick the $\Delta W$ until it does not cross the boundaries.

still not known to us.

## 6.3 Bias Wiener method

Reader may wonder that what happens if we do not use PRS. Instead, whenever a particle cross the boundaries, we re-pick the $\Delta W$ until it does not cross the boundaries. We do a WMC test, using the configuration $N_p = 4000$, $N_b = 500$, $T_{end} = 7200$, $dt = 1$, $\mathbb{D} = [0, 2]$ and $H_h = 2$ on E1, V1, M1, and M2. The result is plotted in Fig. 6.3. The red line is the constant diffusivity $K = 0.001$, and blue line is $K = K_b$. It obviously fails WMC for all four schemes. But what is interesting is that all four schemes behave the same near the bottom boundary $z = 2$. That is, where the drift coefficient $a = 0$. This actually make sense since all four schemes are are very close to each other when $a = 0$. In contrast, the behavior of these four schemes are all different near top boundary. We obtained very different behaviors on top and bottom boundary. Therefore, we conclude that this method does not satisfy WMC in general.

## 6.4 Adaptive methods

Before I have found out that the problem of boundary artifact was essentially caused by the Lipschitz condition no being satisfied. I thought there was just ordinary numerical error. A higher order should minimize the boundary artifact. Therefore, I have tried a time adaptive strong order Runge-Kutta 1.0 and 1.5 from Rackauckas and Nie (2017). The Runge-kutta 1.0 is embedded in Runge-Kutta 1.5, therefore we do not have to calculate one more Runge-kutta 1.0 if we have calculated Runge-kutta 1.5. If the error between Runge-kutta 1.0 and Runge-Kutta 1.5 is too big, use smaller $\Delta t$ instead. The technical problem is that by rejecting an outcome, we have also rejected a random Brownian motion. The solution is to store the Browian motion that we have rejected and construct a Brownian bridge. Here, we do not go in to detail of what Brownain bridge is. The reader can refer to Kloeden and Platen (2013); Rackauckas and Nie (2017).

I have tried the RSwM algorithm 1 from Rackauckas and Nie (2017), and received almost identical result as E1 scheme, but the boundary artifact is even worse. Of course, it turns out the boundary artifact is not a simply numerical error, but caused by violation of Lipschitz condition. The code is written to run on super computer. For more details, see readme.md on my GitHub.

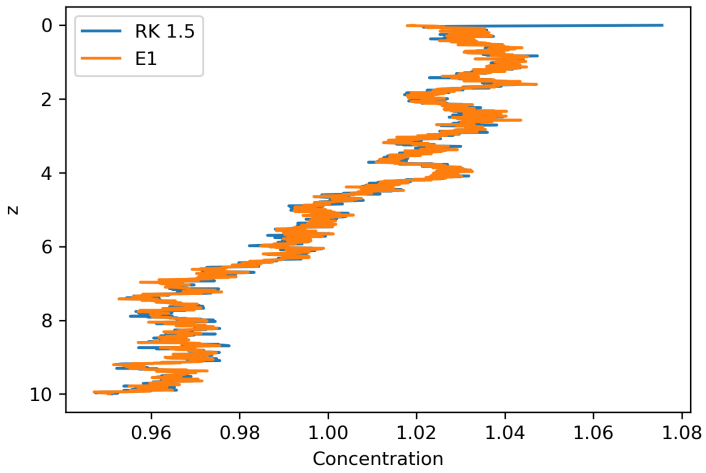



Figure 6.4: Time adaptive Runge-Kutta 1.5 vs E1 scheme. Using maximum of $\Delta t = 0.2\,\mathrm{s}$ for Time adaptive Runge-Kutta 1.5 and store the Brownian motion to be used in E1 scheme.

# Chapter 7

# Conclusion

In this thesis, we have analyzed a numerical a artifact of stochastic differential equations in one-dimensional model of diffusion in the water column, with reflecting boundary at the surface. The numerical artifact occurs near the boundary, therefore we call it the boundary artifact. In many research areas, using stochastic differential equation in a Lagrangian model for particle tracking is very common. A good particle tracking scheme should ensure that a passive tracer stays well mixed when it is originally well mixed. We call this the well-mixed condition (Thomson, 1987). However, most of the SDE particle tracking schemes do not satisfy this condition near reflecting boundaries, and this has been ignored in many research areas. Sometimes, this can lead us to the wrong conclusion.

The perfect reflection scheme (PRS) is a commonly used scheme in many research areas to imitate Neumann boundary condition in Eulerian PDE models. We compared the results obtained with Neumann boundary conditions to the results from an expanded domain, mirrored about the boundary, and with the addition of an image source. Comparing the results in the original domain, to the results in half of the mirrored domain, we find that they are identical even though the Neumann boundary condition was not applied at the mirroring point.

There is no visible numerical artifact in either the original or the mirrored domain in the Eulerian PDE model. In the SDE case with perfect reflection scheme, we have shown that we got the same result as for the PDE (which it should be, based on the argument of weak convergence), except that there is a boundary in the artifact in original domain, and at the same position, the same shape of numerical artifact occurs in mirrored domain.

This is very interesting; an identical numerical artifact occurs in both boundary and non-boundary region. Therefore, we think that there may be something else than just the PRS that causes the problem, but rather something fundamental when one is using SDEs. This turns out to be the Lipschitz condition not being satisfied in the mirrored domain at the mirroring point, and what the PRS does is effectively mirroring the original domain just like the way the Neumann boundary condition does. The Lipschitz condition was one of the conditions in the derivation of an SDE that is equivalent with the advection-diffusion equation (see Section 2.4).

For the Lipschitz condition to be satisfied when using the PRS, the derivative of diffusivity ($\frac{d}{dz}K$) at the boundary should be zero. This is a problem for the SDE. For the PDE, the numerical derivative of diffusivity at the boundary is already zero.

In Section 3.5, we analyze the boundary artifact for the Euler-Maruyama scheme. We defined $f(z, \Delta t)$, which gives the minimum value $\Delta W$ can take, if the particle is to remain inside the domain. Any $\Delta W < f(z, \Delta t)$ will cross the top boundary, and any $\Delta W > f(z, \Delta t)$ will stay inside the domain (extreme values of $\Delta W$ are assumed impossible such that a particle close to the top boundary will not cross the bottom boundary). We have used $f(z, \Delta t)$ to show that in order for the perfect reflection scheme to perfectly imitate the Neumann scheme, $f(z, \Delta t)$ in Eq. (3.13) should go to zero when there is a boundary at $z = 0$ in order to satisfy the Lipschitz condition in SDE theory, as well as the WMC.

We have also used $\frac{\partial}{\partial z} f(z, \Delta t)|_{z=0}$ to help us distinguish two different types of boundary artifact. We call these Type A, which is defined as "Particles have less chance to cross the boundary when they are getting close to the boundary", and Type B, which is "Particles **do not** have less chance to cross the boundary when they are getting close to the boundary" see Section 3.6. Type A is more severe, and leads to larger errors than Type B. Depending on the application, Type B may give an acceptable error.

Personally, I think Type A is not acceptable for most applications. In fact, this is a dangerous boundary artifact, because most researchers simulate only the time evolution of concentration, not the time-average concentration. In this case, the researcher can always be misled to think that the change in concentration near the boundary happens by chance, not the tendency. In this case, the researcher may not realize that the boundary artifact occurs in the simulation. The user should always perform a WMC test and decide by his/her own. However, the boundary artifact will completely go away when

the Lipschitz condition is satisfied.

We has investigated the problem with three different diffusivity profiles, which differed only in an added constant (see Eq. (3.1)). We did not provide a solution for diffusivity equal to zero at the boundaries ($K_a(z)$). As this is no longer a diffusion problem, it is not guaranteed to satisfy well-mixed the condition. We gave a few different suggested solutions to the boundary artifact problem for low diffusivity at the boundaries. Of course, these can also be used for high diffusivity boundaries. However, as the boundary quality function points out ($f(z, \Delta t)$ in Eq. (3.13)), if the diffusion term is dominating, the boundary artifact that is caused by PRS can be negligible for most applications.

In practical situations, the velocity statistics near the boundary is unknown. Wilson and Flesch (1993) called this "Unresolved basal layer". This can give us some freedom to design our own diffusivity near the boundaries. If one obtains diffusivity from analytical expressions, and it has low diffusivity at the boundary , and one is not willing to change the diffusivity profile too much, then Cubic spline interpolation with clamped condition may be a good choice.

The smoothened-diffusivity method (Section 3.5.5 and Section 5.1) will always satisfy the WMC with sufficiently small $\Delta t$, but includes changing diffusivity. We have discussed the legitimization of changing the diffusivity near boundary. Wilson and Flesch (1993) argue that the we simply cannot know the details of the turbulent motion near boundaries, As mentioned, this gives us some freedom near the boundaries. Furthermore, Most researchers often treat the fluctuating sea surface as a hard-wall. This is already changing part of the physical reality. Therefore, smoothing the diffusivity very close to the boundary should not alter too much of the result.

We mostly used a diffusivity profile where the drift term dominates over the diffusion term near the boundary ($K_b$, Eq. (3.1b)) to perform the WMC tests in this thesis. In general practical ocean applications the diffusivity will often be much bigger. For example, $K_c$ (Eq. (3.1c)) is from Visser (1997) and it is almost 1000 times stronger than $K_b$. We choose to use low diffusivity because we want to highlight the boundary artifact. For higher diffusivity, smoothing the diffusivity near the boundary has very small impact on the results. Therefore, we conclude that smoothened diffusivity method is a generally good method to satisfy WMC for most applications.

# References


Arnold, L. (1974). Stochastic differential equations. *New York*.

Ben-Naim, A. (2008). *Entropy demystified: The second law reduced to plain common sense*. World Scientific.

Dabiri, J. O. (2010). Role of vertical migration in biogenic ocean mixing. *Geophysical research letters*, *37*(11).

Evans, L. C. (2010). Partial differential equations.

Gräwe, U. (2011). Implementation of high-order particle-tracking schemes in a water column model. *Ocean Modelling*, *36*(1-2), 80–89.

Gräwe, U., Deleersnijder, E., Shah, S. H. A. M., & Heemink, A. W. (2012). Why the euler scheme in particle tracking is not enough: the shallow-sea pycnocline test case. *Ocean Dynamics*, *62*(4), 501–514.

Kloeden, P. E., & Platen, E. (2013). *Numerical solution of stochastic differential equations* (Vol. 23). Springer Science & Business Media.

Maruyama, G. (1955). Continuous markov processes and stochastic equations. *Rendiconti del Circolo Matematico di Palermo*, *4*(1), 48.

Milshtein, G. (1975). Approximate integration of stochastic differential equations. *Theory of Probability & Its Applications*, *19*(3), 557-562. doi: 10.1137/1119062

Milshtein, G. (1979). A method of secord-order accuracy integration of stochastic differential equations. *Theory of Probability & Its Applications*, *23*(2), 396-401. doi: 10.1137/1123045

Munk, W., & Wunsch, C. (1998). Abyssal recipes ii: energetics of tidal and wind mixing. *Deep Sea Research Part I: Oceanographic Research Papers*, *45*(12), 1977 - 2010. Retrieved from http://www.sciencedirect.com/science/article/pii/S0967063798000703 doi: https://doi.org/10.1016/S0967-0637(98)00070-3

Notter, R. H., & Sleicher, C. (1971). The eddy diffusivity in the turbulent boundary layer near a wall. *Chemical Engineering Science*, *26*(1), 161–

171.

Øksendal, B. (2003). Stochastic differential equations. In *Stochastic differential equations* (pp. 65–84). Springer.

Rackauckas, C., & Nie, Q. (2017). Adaptive methods for stochastic differential equations via natural embeddings and rejection sampling with memory. *Discrete and continuous dynamical systems. Series B*, *22*(7), 2731.

Risken, H. (1996). Fokker-planck equation. In *The fokker-planck equation* (pp. 63–95). Springer.

Sauer, T. (2011). *Numerical analysis* (2nd ed.). USA: Addison-Wesley Publishing Company.

Singer, J., Atkinson, J., Manley, P., & McLaren, P. (2008). Understanding sediment dynamics using geological and engineering approaches: a case study of the buffalo river area of concern, buffalo, new york. *International Journal of River Basin Management*, *6*(1), 31–40.

Thomson, D. (1987). Criteria for the selection of stochastic models of particle trajectories in turbulent flows. *Journal of Fluid Mechanics*, *180*, 529–556.

van Sebille, E., Griffies, S. M., Abernathey, R., Adams, T. P., Berloff, P., Biastoch, A., ... Zika, J. D. (2018). Lagrangian ocean analysis: Fundamentals and practices. *Ocean Modelling*, *121*, 49 - 75. Retrieved from http://www.sciencedirect.com/science/article/pii/S1463500317301853 doi: https://doi.org/10.1016/j.ocemod.2017.11.008

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*, 275–281.

Wilson, J. D., & Flesch, T. K. (1993). Flow boundaries in random-flight dispersion models: enforcing the well-mixed condition. *Journal of Applied Meteorology*, *32*(11), 1695–1707.

Yunus, A. C., & Cimbala, J. M. (2006). Fluid mechanics fundamentals and applications. *International Edition, McGraw Hill Publication*, *185201*.

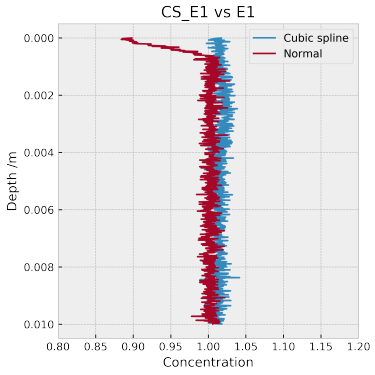
# Appendices

# Appendix A

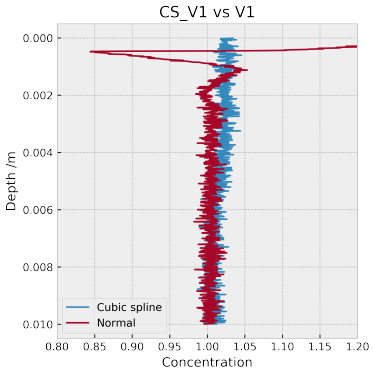
# Figures

## A.1 Cubic spline for E1, V1, M1 and M2

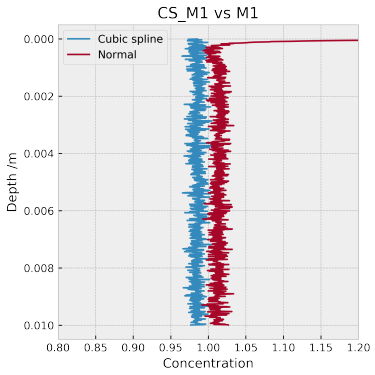
## A.2 Figure from "Entropy demystified"



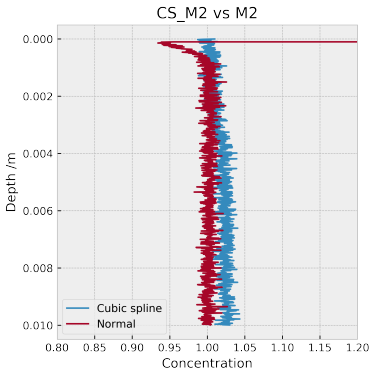

(a) E1




(b) V1




(c) M1




(d) M2

Figure A.1: Cubic spline interpolation method on 4 different SDE scheme

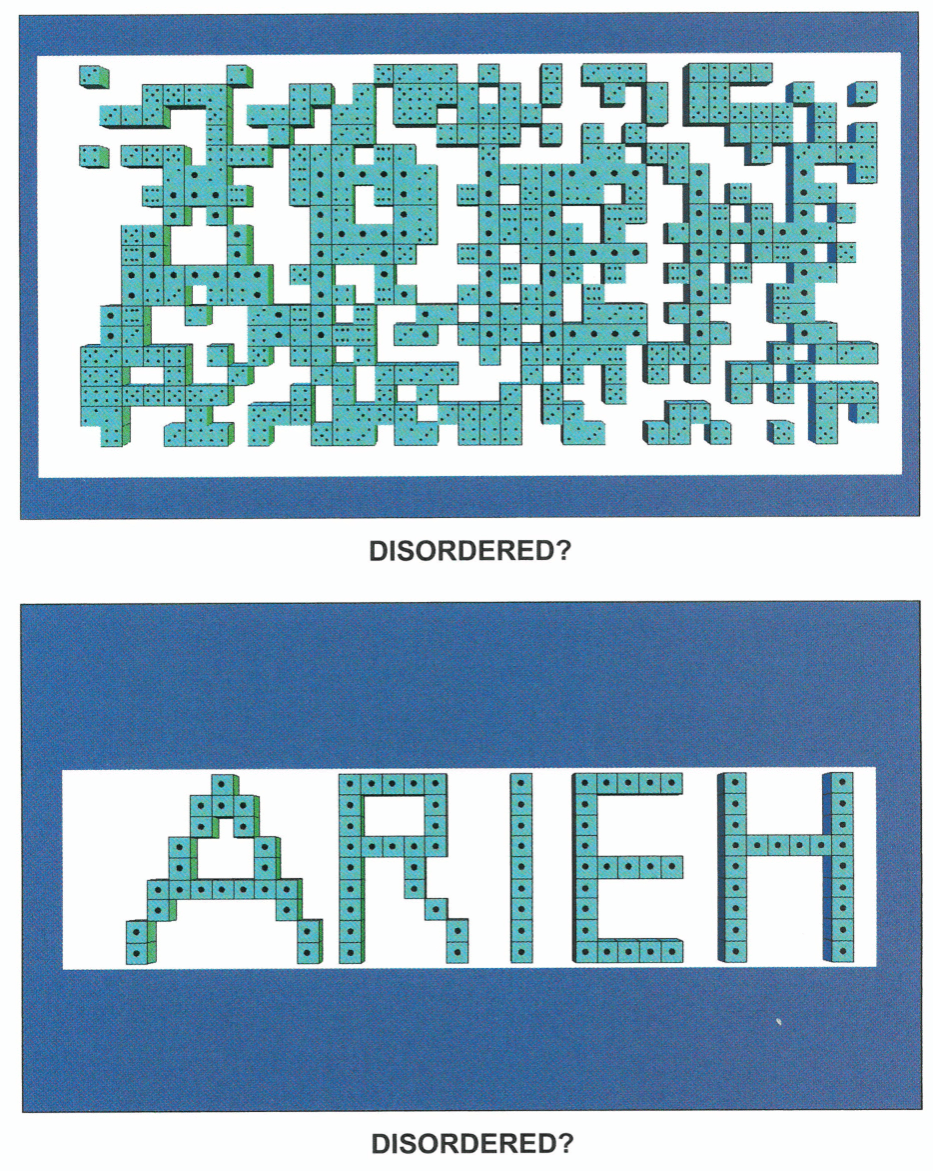



Figure A.2: Permission granted by author and World Scientific. Source: Entropy demystified second edition, Arieh Ben-Naim, Copyright @ 2016 by World Scientific Publishing Co. Ptw. Ltd

# Appendix B

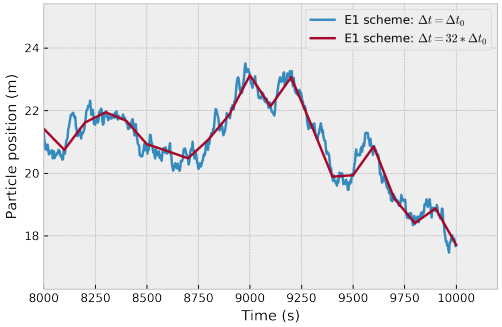
# Weak and Strong convergence

## B.1 Strong convergence

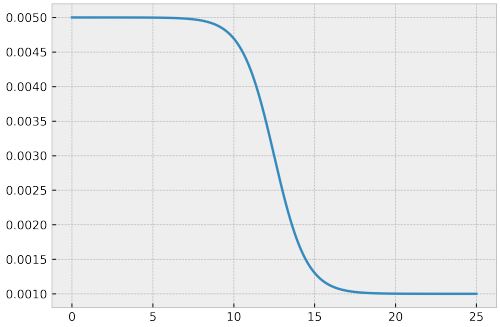
We only have the strong convergence that is defined in the Itô sense, not in the Stratonovich sense. Therefore, Visser scheme and Milstein 2nd-order scheme is not defined in this case since the former is defined in the Stratonovich sense and the later is to ignore some term from 1.5 order Taylor scheme. Therefore, we will test strong convergence of E1 and M1 schemes.

To illustrate this effect, we use a diffusivity profile with a smooth transition between two constants. The idea is to show that how fast a numerical scheme converges to a true solution with difference $\Delta t$ under the same Wiener process. Therefore, we first define minimal time-step and end-time, denoted $\Delta t_{min}$ and $T_{end}$, and we use this minimal time-step to simulate $T_{end}/\Delta t_{min}$ of $\Delta W \sim N(0, \Delta t_{min})$. Now we obtain sequence of $\Delta W$, denoted $(\Delta W)_n$, and our Wiener process will become $W_i = W(t_i) = \sum_{n=0}^{i} \Delta W_n$. Using this $W_i$ we can simulate Lagrangian model with 6 different timesteps under the same Wiener process. They are $\Delta t = \Delta t_{min}$, $\Delta t = 2 \cdot \Delta t_{min}$, $\Delta t = 4 \cdot \Delta t_{min}$, $\Delta t = 8 \cdot \Delta t_{min}$, $\Delta t = 16 \cdot \Delta t_{min}$ and $\Delta t = 32 \cdot \Delta t_{min}$. Of cause, we do not have the exact solution. Therefore, we will treat the densest time interval, which $\Delta t_{min}$, as an exact solution and find the distance difference between our "exact" solution and the current time interval of the scheme. Figure B.1a is an example between the densest time interval ($\Delta t_{min}$) and $\Delta t = 32 \cdot \Delta t_{min}$ under the same realized Wiener process.

To get a statistical average, we repeat the process with different Wiener process. Because we have to repeat in many wiener processes, it is better




(a) Euler schemes with two different $\Delta t$ under a same realized Wiener process. $T_{end} = 10000\,\mathrm{s}$




(b) Diffusivity profile.

Figure B.1: Two figures to illustrate the idea of strong convergence.

that we reduce the $T_{end}$ and $N$.

### B.1.1 Testing environment and parameters set-up:

The depth of container is $H = 25\,\mathrm{m}$. The diffusivity profile is the following:

$$K(z) = K_0 - (K_1 - K_0)\left(1 - \frac{1}{1 + e^{-\frac{m-z}{\eta}}}\right) \tag{B.1}$$

where $K_0 = 5{\cdot}10^{-3}\,\frac{\mathrm{m}^2}{\mathrm{s}}$, $K_1 = 10^{-3}\,\frac{\mathrm{m}^2}{\mathrm{s}}$, $m = \frac{H}{2}\,\mathrm{m}$ and $\eta = 1\,\mathrm{m}$. The diffusivity profile, which describes the turbulence, is plotted in figure B.1b

We have also $k \in \{1, 2, 4, 8, 16, 32\}$, where $\Delta t = k \cdot \Delta t_{min}$, $T_{end} = 100\,\mathrm{s}$ and $N = 100$. We run 2000 different Wiener process for each $k$ and find the average error between this timestep and the "exact" solution at $T_{end}$.

### B.1.2 Result

The result is shown in figure B.2. The blue line scales as $\Delta t^{\frac{1}{2}}$. The red line scales as $\Delta t^1$. We can see that the E1 have strong convergence around order $\Delta t^{\frac{1}{2}}$ and M1 have order $\Delta t^1$. Figure B.2 tell us not just M1 converge faster than E1 in the strong sense, it also tells us that M1 have more accuracy in the strong sense.

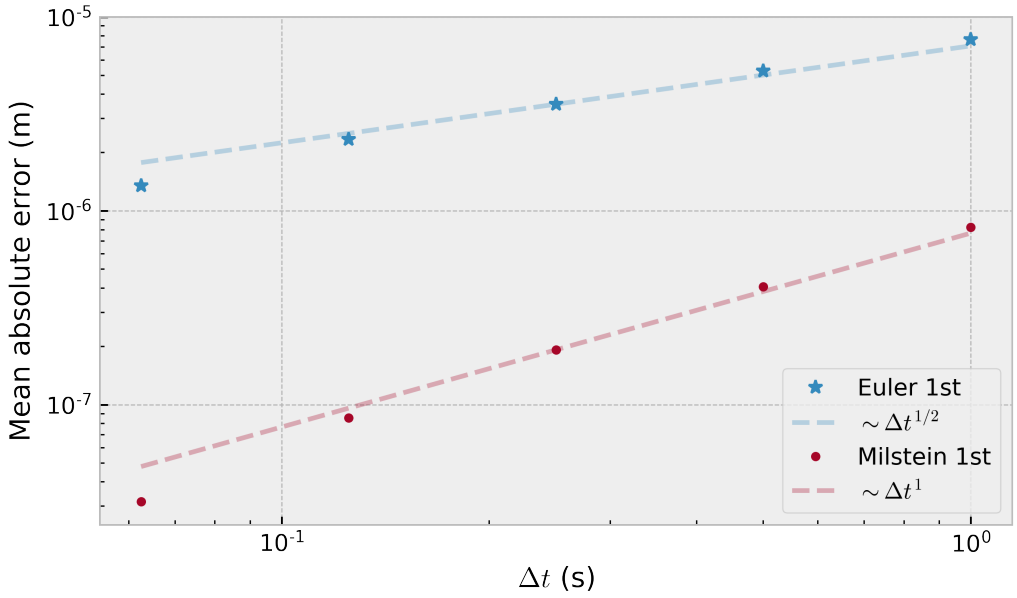



Figure B.2: Strong convergence of Euler-Maruyama and Milstein 1st order schemes for 2000 difference Wiener process. The blue line scales as $\Delta t^{\frac{1}{2}}$. The red line scales as $\Delta t^1$.

## B.2 Weak convergence

In weak convergence, we will study how fast and how well the numerical scheme converges to the distribution of the true solution. In another word, if we have infinite many particles, the true solution is the probability density function from the Fokker-Planck equation. However, we still do not have the true solution with a finite particle. This time we will use the Eulerian model as a true solution, and the numerical integration on this model will be Crank-Nicolson scheme to solve the diffusion equation. This is because the concentration of the true solution of SDE with infinite particles should converge to Eulerian model. The precision of Eulerian model is very important. It must be more accurate to the Lagrangian model. At least one order more accurate. Otherwise the error of Eulerian model will dominate the weak convergence. Crank-Nicolson may be a good choice since it is second order scheme of time.

Figure B.3 is an example of concentration in the Lagrangian model has

been calculated by Milstein 2nd-order scheme. We can see that the Lagrangian model follows almost the same shape of the true solution. This encourages us that the SDE can converge to Eulerian. However, we need to calculate $|\langle p(Z_n)\rangle - \langle p(Z(\tau))\rangle|$ which from equation (2.23). In this test, we choose

$$P(x) = x^n,$$

where $n = 1$ in our case. This means that the first moment of the SDE is calculated as the sum of all particles positions and divide by the number of particles. The first moment of the Eulerian model is

$$\int_0^H C(z) z^n dz,$$

where $n = 1$. A numerical scheme can efficiently compute this. As the name supposed, weak convergence converges weakly in number of particles. The more the particles is used, the more accurate the result will be. Therefore, we need many particles to perform the simulation. Here we use 20000000 particles for all four different SDE schemes. The problem turns to computational power that we have. Because we are already using huge number of particles. If we still use small $\Delta t$, it will take a very long time to compute. Using bigger $\Delta t$ may be our alternative option. But does our schemes still accurate enough or stable? One may wonder. Visser (1997) proposed a restriction on the $\Delta t$ for the integration:

$$\Delta t << \min\left(\frac{1}{|\partial_{zz} K(z)|}\right) \quad \text{for} \quad \forall z \in D \tag{B.2}$$

where $\partial_{zz} K(z)$ is the double derivative of diffusivity respect to space, $D$ is the domain and $||$ is the absolute value operator. This restriction is widely accepted. From the Eq. (B.2), we can always use bigger $\Delta t$ as long as our diffusivity changes weak enough. Therefore, all we have to do is just using a diffusivity that changes weakly respect to space. In this way, we make our restriction much bigger than our $\Delta t$, instead of using a smaller $\Delta t$. We can, therefore, use large $\Delta t$ to compare the weak convergence between all four different schemes when number of particle is large enough.

### B.2.1 Testing environment and parameters set-up:

The diffusivity profile is exactly the same as strong convergence case except $K_0 = 1 \cdot 10^{-4} \frac{\mathrm{m}^2}{\mathrm{s}}$ and $K_1 = 1 \cdot 10^{-6} \frac{\mathrm{m}^2}{\mathrm{s}}$. It gives us the right-hand

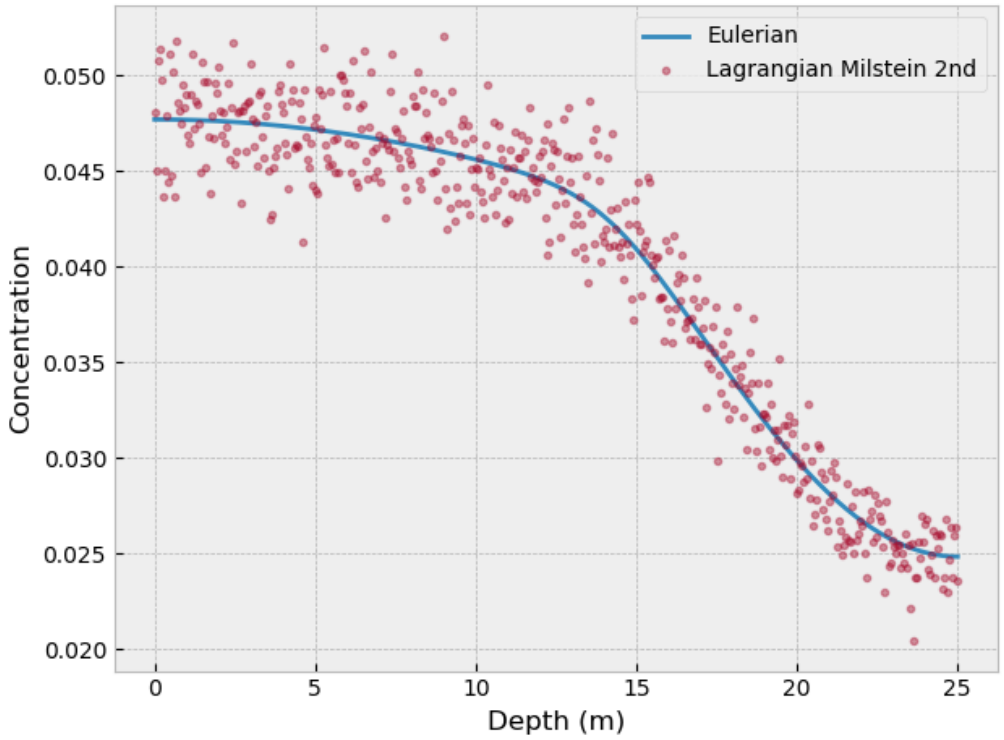



Figure B.3: Weak convergence of Eulerian and Milstein 2nd-order schemes. For Milstein 2nd-order schemes, we use for 2000000 particles, $\Delta t = 1\,\mathrm{s}$ and $T_{end} = 12 \cdot 3600\,\mathrm{s}$, and number of bins is 1000.

side of the Eq.(B.2) equals to 104973 s. This means that we can have $\Delta t \in \{3600, 4320, 4800, 5400, 7200, 8640, 10800, 14400, 21600\}$. Furthermore, we use for 20000000 particles and the $T_{end} = 12 \cdot 3600\,\mathrm{s}$ and number of bins is 1000.

### B.2.2 Result

The result is plotted in figure B.4. We can see that the E1, V1 and M1 schemes scale well with $\Delta t$. M2 scales well with $\Delta t^2$ with a large time step. However, for a smaller time step, it shows that the 20000000 particles are not enough to exhibit the weak divergence property. Furthermore M1 have slightly higher error than E1 and V1.

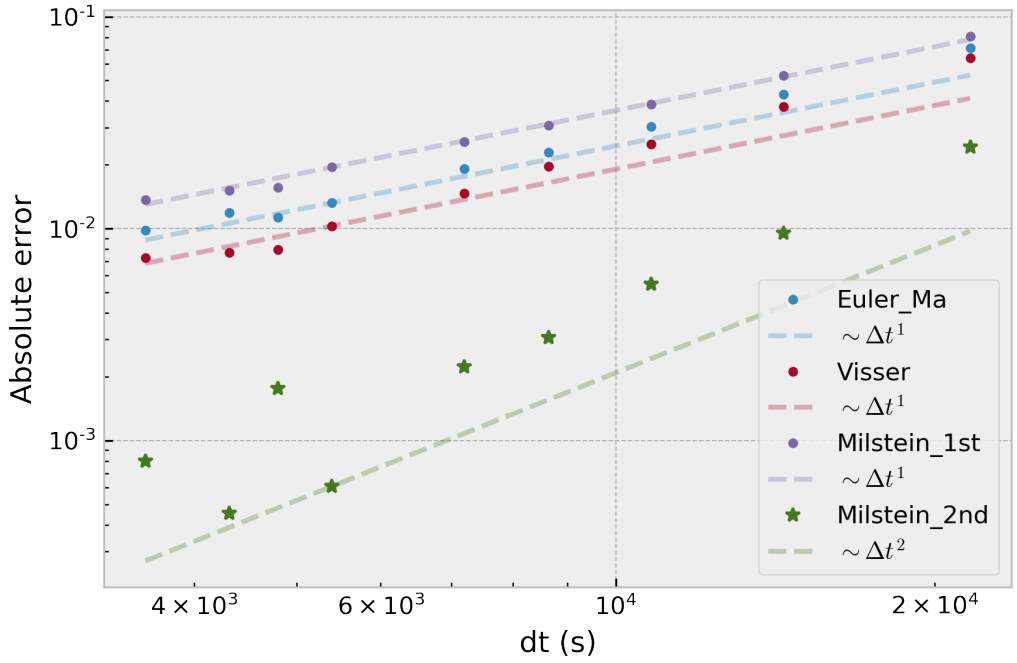



Figure B.4: Weak convergence: The blue line, red line and purple line scales as $\Delta t^1$. The green line scales as $\Delta t^2$

# Appendix C

# Proof of section 2.6

*Proof.* Let $Z_n \sim N(\mu = Z_{\delta,n}, \sigma = 0)$ and $Z(\tau) \sim N(\mu = Z_\delta(\tau), \sigma = 0)$. This mean that $Z_n$ and $Z(\tau)$ have PDF like delta function with peak $Z_{\delta,n}$ and $Z_\delta(\tau)$ respectfully.

For weak convergence and $P(x) = x$, we have

$$|\langle p(Z_n)\rangle - \langle p(Z(\tau))\rangle| = |\langle Z_n\rangle - \langle Z(\tau)\rangle|,$$

recall that

$$\langle Z_n\rangle = \int_{-\infty}^{\infty} Z_n \delta(Z_n - Z_{\delta,n}) dZ_n = Z_{\delta,n},$$

and

$$\langle Z(\tau)\rangle = \int_{-\infty}^{\infty} Z(\tau)\delta(Z(\tau) - Z_\delta(\tau)) dZ(\tau) = Z_\delta(\tau).$$

we obtain

$$|\langle Z_n\rangle - \langle Z(\tau)\rangle| = |Z_{\delta,n} - Z_\delta(\tau)| : \text{weak}.$$

For strong convergence, we have

$$\langle |Z_n - Z(\tau)| \rangle = \langle \sqrt{(Z_n - Z(\tau))^2}\, \rangle$$

and it is equal

$$\langle \sqrt{(Z_n - Z(\tau))^2}\, \rangle = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \sqrt{(Z_n - Z(\tau))^2}\delta(Z_n - Z_{\delta,n}) \cdot \delta(Z(\tau) - Z_\delta(\tau)) dZ_n dZ(\tau)$$

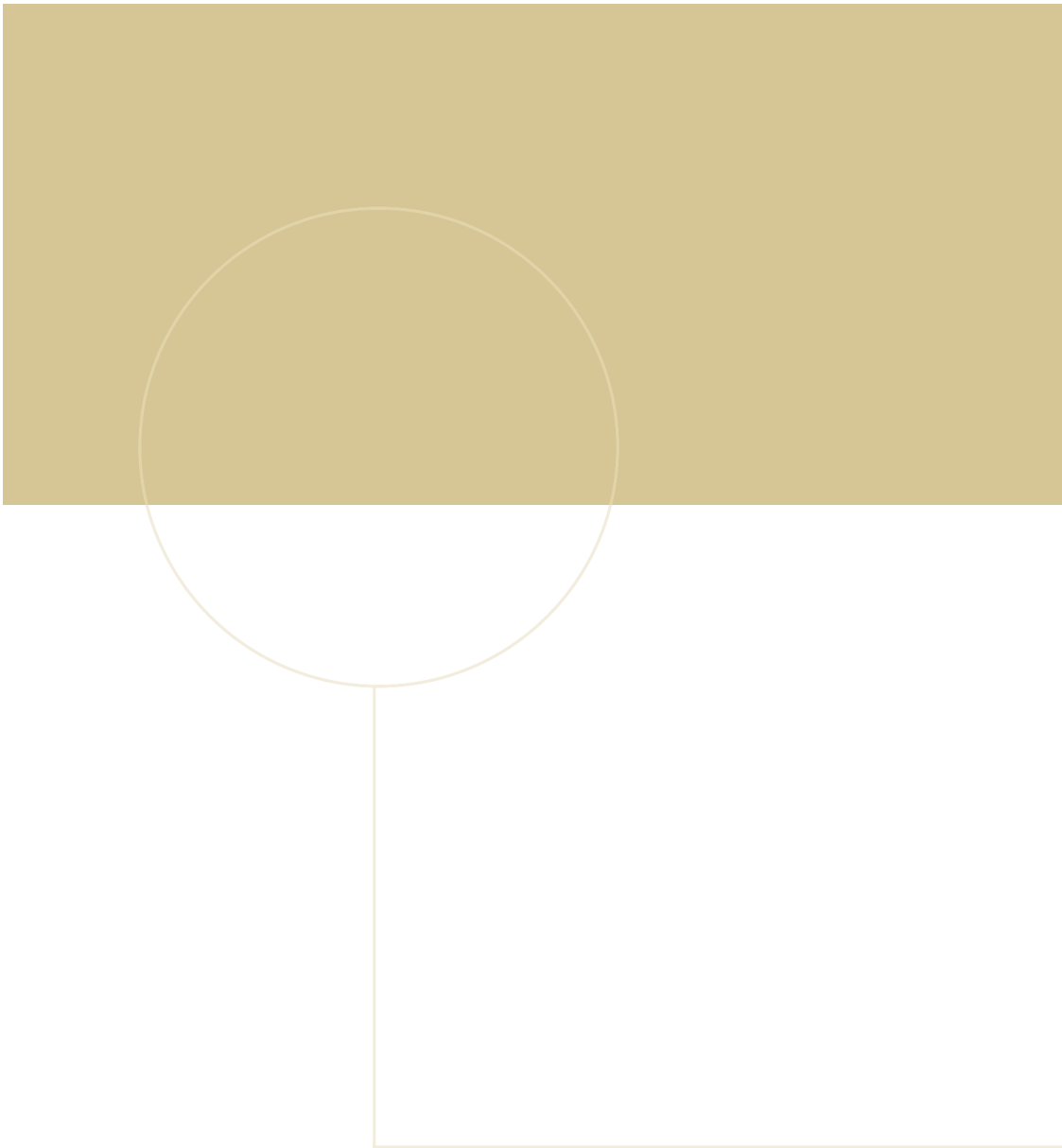
$$= \sqrt{(Z_{\delta,n} - Z_\delta(\tau))^2} = |(Z_{\delta,n} - Z_\delta(\tau)| : \text{strong}$$

This show that the weak convergence and strong convergence is same when the SDE with delta function as PDF. It essentially transform the SDE to ODE. ■

# Appendix D

# Code

This thesis is already almost hundred pages, to avoid this thesis too many pages. I uploaded all code to the Github. Most figure has its own Jupyter file. All the Jupyter involve WMC test has been parallelized. The reader can specific number of threads to be used in the simulation. The default value is four. I suggestion user to use number of process unit in the CPU. All the configuration of simulation should be very clear for the reader because I always specific my configuration before simulation. The code can also run on Windows, Linux, iOS platform. If reader want to have a look. Here is the link: https://github.com/Mauhing/MasterThesis.git. For more information, see the readme file in Github.



NTNU
Norwegian University of
Science and Technology