

Pedestrian Flow in the Mean Field Limit

Thesis by
Abdul Lateef Haji Ali, B.Sc.

In Partial Fulfillment of the Requirements

For the Degree of

Masters of Science

King Abdullah University of Science and Technology, Thuwal
Kingdom of Saudi Arabia

November, 2012

The thesis of Abdul Lateef Haji Ali is approved by the examination committee

Committee Chairperson: Dr. Raúl Tempone

Committee Member: Dr. Aslan Kasimov

Committee Member: Dr. David Ketcheson

Committee Member: Dr. David Keyes

Copyright ©2012

Abdul Lateef Haji Ali

All Rights Reserved

ABSTRACT

Pedestrian Flow in the Mean Field Limit

Abdul Lateef Haji Ali

We study the mean-field limit of a particle-based system modelling the behaviour of many indistinguishable pedestrians as their number increases. The base model is a modified version of Helbing's social force model. In the mean-field limit, the time-dependent density of two-dimensional pedestrians satisfies a four-dimensional integro-differential Fokker-Planck equation. To approximate the solution of the Fokker-Planck equation we use a time-splitting approach and solve the diffusion part using a Crank-Nicholson method. The advection part is solved using a Lax-Wendroff-Leveque method or an upwind Backward Euler method depending on the advection speed. Moreover, we use multilevel Monte Carlo to estimate observables from the particle-based system. We discuss these numerical methods, and present numerical results showing the convergence of observables that were calculated using the particle-based model as the number of pedestrians increases to those calculated using the probability density function satisfying the Fokker-Planck equation.

ACKNOWLEDGMENTS

I would like to deeply thank my advisor Raúl Tempone for his support and invaluable comments and contributions to this thesis. We also thank Prof. David Ketcheson, Mohammad Motamed, Matteo Parsani, Erik von Schwerin and Prof. Hamidou Tembine for their help and support in this project. We would also like to thank Prof. Dirk Helbing and Illes Farkas for generously providing the code of their social-force model presented in their work [1].

TABLE OF CONTENTS

Examination Committee Approval	2
Copyright	3
Abstract	4
Acknowledgments	5
List of Figures	8
List of Tables	9
I Introduction	10
II Theory	13
II.1 Background	13
II.1.1 Itô Stochastic Differential Equations	13
II.1.2 Mean-Field Limit	18
II.2 Model Formulation	21
II.2.1 Helbing Particle-Based Model	22
II.2.2 Modified Particle-Based Model	25
II.2.3 Mean Field Limit	30
III Numerical Simulations and Solutions	32
III.1 Computational Methods	32
III.1.1 Numerical Schemes for Solving PDEs	32
III.1.2 Numerical Schemes for Solving SDEs	39
III.1.3 Monte Carlo Methods	40
III.1.4 Multilevel Monte Carlo	42
III.2 Implementation	43

III.2.1 Particle-Based Simulation	44
III.2.2 Numerical Solution of the Continuous Model	45
IV Numerical Results	50
IV.1 Observables and Parameters	50
IV.2 Particle-Based Model	52
IV.3 Continuous Model	55
IV.4 Comparison	60
IV.5 Computational Challenges in <i>CaseII</i>	60
V Particle Multilevel Monte Carlo	65
VI Conclusions	71
References	73

LIST OF FIGURES

II.1	Typical environment in <i>CaseII</i>	22
II.2	Signed distance between two pedestrians	25
II.3	Cut-off function φ for different values of the parameter ζ	28
II.4	δ_v illustration	29
IV.1	Sample realizations of <i>CaseI</i>	56
IV.2	Sample realizations of <i>CaseII</i>	57
IV.3	Convergence of particle-based model.	58
IV.4	L^1 convergence of approximate solution to continuous model in <i>CaseI</i>	62
IV.5	L^1 convergence of approximate solution to continuous model in <i>CaseII</i>	62
IV.6	Convergence of vector of observables computed using the approximate solutions of the continuous model in <i>CaseI</i>	63
IV.7	Convergence of vector of observables computed using the approximate solutions of the continuous model in <i>CaseII</i>	63
IV.8	Convergence of particle-based model to continuous model	64
V.1	MLMC statistics	70

LIST OF TABLES

IV.1 Parameters used in the simulations for <i>CaseI</i> and <i>CaseII</i>	53
IV.2 Values of some observables at final time.	54

Chapter I

Introduction

Developing an efficient design of an environment where humans are involved requires a deep understanding of human behaviour, whether this design is for a hospital building, a shopping mall, an escape route or even a computer game. This is why crowd modelling and simulation are important tools for analysts to run mock experiments, to plan real ones and to observe how humans may behave in a crowd.

Crowd models can be divided into three broad categories [2]: Particle-based, cellular automata-based and continuous. Particle-based models are the oldest, most obvious and most popular [3–6]. In these discrete models, each individual is modelled by a particle whose dynamics are coupled to the dynamics of all other individuals (particles) in the system. Cellular automata-based models [7, 8], also discrete, model the space instead as a lattice of cells, where each cell can attain different states based on interaction with other neighbouring cells.

On the other hand, continuous models are at the other end of the modelling spectrum. In addition to being able to model a larger number of pedestrians, continuous models allow for the use of the already established analytical techniques of partial differential equations (PDEs) to understand the behaviour of the model without actually solving it.

Hughes [9] models the movement of pedestrians as a fluid, as it was observed that displacement of pedestrians resembles a flow of fluid to a great extent [10]. This model was also used in traffic problems with great success and is in fact very similar to the model derived in this project. However, this model was postulated based on some intuition of the crowd dynamics and not motivated from other, more basic models. In contrast, the work [11] uses techniques from game theory to find the path that each pedestrian takes by optimizing a cost function that assigns a high cost to penetrating obstacles or other pedestrians and a high payoff to reaching the individual's goal. A drawback of the model in [11] is that it assumes that pedestrians have global knowledge of the environment and the state of all other pedestrians in order to find an optimal path to their goal.

The basic idea of this project is to take an already established and relatively basic discrete crowd model, and investigate its limit as the number of pedestrians increases, called a mean-field limit. This is similar to what Bolley et. al [12] did with Cucker-Smale [13] models of animal flocking. As the underlying discrete model we chose the social force model introduced by Helbing [1]. This is a particle-based model that describes the dynamics of each individual by a stochastic differential equation (SDE) that depends on a local neighborhood of that individual and the individual's goal. Moreover, the SDE of each individual is coupled to the SDEs of other individuals. We chose this model because it was observed to reproduce many phenomena of crowd flow as it was compared against data from the real world [1, 3, 4, 14].

We will be looking at two cases. ***Case I*** is a one-dimensional test case in which pedestrians move clockwise on a circle with no obstacles. On the other hand, ***Case II*** is a more realistic two-dimensional case in which pedestrians move towards a fixed target in a fixed environment with obstacles.

This thesis first presents a brief overview of the theory of SDEs, the Fokker-Planck equation and mean-field limits in Chapter II. In that chapter we also review Helbing's

model and propose necessary modifications to it. The chapter concludes by looking at the model formulation in the mean-field limit. Chapter III then proceeds to discuss the different computational methods that are used to find approximate solutions to both the modified Helbing's particle-based model and the continuous mean-field model. Chapter IV presents convergence studies of the numerical solutions of both models. In that chapter we present plots that show convergence of quantities of interest or observables approximated using the particle-based model to those approximated using the continuous, mean-field limit. We also discuss the computational difficulties in *CaseII*. In Chapter V we look at a possible extension of the multilevel Monte Carlo method that utilizes the convergence of the particle-system to a mean-field to efficiently compute observables from the mean-field. Finally in Chapter VI we conclude by summarizing and discussing future work.

Chapter II

Theory

II.1 Background

In this section we give a brief overview and definitions of basic theoretical concepts that we will use in this project.

II.1.1 Itô Stochastic Differential Equations

The following review is taken from [15, 16], where more details and proofs can be found.

Wiener Process

The one-dimensional Wiener process $W(t)$ is a random process defined to have the following properties:

1. With probability 1, the mapping $t \rightarrow W(t)$ is continuous and $W(0) = 0$.
2. For any final time T and time discretization $0 = t_0 < t_1 < \dots < t_K = T$, the increments

$$W(t_K) - W(t_{K-1}), \dots, W(t_2) - W(t_1), W(t_1) - W(t_0)$$

are independent.

3. For all $t > s$ the increment $W(t) - W(s) \sim \mathcal{N}(0, t - s)$. That is, the increment $W(t) - W(s)$ is a random variable that has a normal distribution with zero mean and variance $t - s$.

Such a process can be proven to exist [15]. Usually we are interested in sampling the Wiener process on a finite set of points. The previous properties imply that this can be achieved by sampling random independent increments from a normal distribution with the correct mean and variance.

Strong and weak convergence

We say that a sequence of d -dimensional random variables $\{\mathbf{x}_n\}_{n=0}^{\infty}$ converges weakly or in law to \mathbf{x} and write $\mathbf{x}_n \xrightarrow{\mathcal{L}} \mathbf{x}$ if

$$\lim_{n \rightarrow \infty} \mathbb{E} \{g(\mathbf{x}_n)\} = \mathbb{E} \{g(\mathbf{x})\}, \quad (\text{II.1})$$

for all continuous and bounded scalar functions g , where $\mathbb{E} \{\cdot\}$ denotes an expectation.

On the other hand, we say that $\{\mathbf{x}_n\}_{n=0}^{\infty}$ converges strongly or in mean square to \mathbf{x} and write $\mathbf{x}_n \xrightarrow{L^2} \mathbf{x}$ if

$$\lim_{n \rightarrow \infty} \mathbb{E} \{\|\mathbf{x}_n - \mathbf{x}\|^2\} = 0. \quad (\text{II.2})$$

Note that strong convergence implies weak convergence but the converse is not true in general.

Itô Stochastic Integrals

Given $h > 0$ define a time discretization $0 = t_{h,0} < t_{h,1} < \cdots < t_{h,K} = T$ such that $\max_n(t_{h,n+1} - t_{h,n}) < h$. Assume that a function f is Lipschitz, i.e., f satisfies

$$|f(t) - f(s)| \leq C|t - s|, \quad (\text{II.3})$$

for some positive constant C and all $t, s \geq 0$. Also let $\Delta W_{h,i} = W(t_{h,i+1}) - W(t_{h,i})$. Then the limit of the following forward Euler discretization

$$\hat{I}_h(T) = \sum_{i=0}^K f(t_{h,i}) \Delta W_{h,i}, \quad (\text{II.4})$$

as $h \rightarrow 0$ is defined to be the Itô stochastic integral and is written as

$$I(T) = \int_0^T f(t) \, dW(t), \quad (\text{II.5})$$

It can be proven that such a limit exists (cf. [15]) and the sequence of forward Euler discretizations $\hat{I}_h(T)$ converges strongly to $I(T)$ as $h \rightarrow 0$.

Itô Stochastic Differential Equations

Assume that the functions $a(x, t)$ and $b(x, t)$ satisfy, for all $x, y \in \mathbb{R}$ and $t, s \in [0, T]$ for some final time T , the following conditions:

$$|a(x, t) - a(y, t)| \leq C|x - y|, \quad (\text{II.6a})$$

$$|b(x, t) - b(y, t)| \leq C|x - y|, \quad (\text{II.6b})$$

$$|a(x, t) - a(x, s)| + |b(x, t) - b(x, s)| \leq C(1 + |x|)\sqrt{|t - s|}, \quad (\text{II.6c})$$

for some positive constant C . Also, let $X(t)$ be a stochastic process that satisfies the following equation for all $t \in [0, T]$

$$X(t) = \mathring{X} + \int_0^t a(X(s), s) \, ds + \int_0^t b(X(s), s) \, dW(s), \quad (\text{II.7})$$

Then we write this equation in differential form as

$$dX(t) = a(X(t), t) \, dt + b(X(t), t) \, dW(t), \quad (\text{II.8a})$$

$$X(0) = \mathring{X}. \quad (\text{II.8b})$$

and call it a stochastic differential equation (SDE). Moreover, X satisfies

$$\max_{t>0} \mathbb{E} \{ (X(t))^2 \} < \infty. \quad (\text{II.9})$$

For brevity, we might sometimes drop the argument t from $X(t)$ and $dW(t)$ when it is apparent from the context. We refer to the dt term as the advection or drift term, while the dW term is referred to as the noise or diffusion term.

Multidimensional SDEs

Let $\mathbf{X}(t)$ be a d -dimensional stochastic process and assume that \mathbf{a} and \mathbf{b} are also d -dimensional and satisfy (II.6). Then we extend the previous definition of a one-dimensional SDE to higher dimensions by writing

$$d\mathbf{X}(t) = \mathbf{a}(\mathbf{X}(t), t) \, dt + \text{diag}(\mathbf{b}(\mathbf{X}(t), t)) \, d\mathbf{W}(t), \quad (\text{II.10a})$$

$$\mathbf{X}(0) = \mathring{\mathbf{X}}. \quad (\text{II.10b})$$

Here $\mathbf{W} \in \mathbb{R}^d$ is a vector of independent Wiener processes and $\text{diag}(\mathbf{x})$ is the square matrix whose diagonal is the vector \mathbf{x} .

Fokker-Planck Equation

Given a time t , let $\rho(t, \mathbf{x})$ be the probability distribution function (pdf) of the d -dimensional random variable $\mathbf{X}(t)$ taken from the random process that satisfies the SDE (II.10). Then, assuming \mathbf{a} and \mathbf{b} are sufficiently regular, the function $\rho(t, \mathbf{x})$ satisfies the Fokker-Planck initial-value problem (IVP) [15]

$$\partial_t \rho + \nabla_{\mathbf{x}} \cdot (\mathbf{a} \rho) - \frac{1}{2} \nabla_{\mathbf{x}}^2 (\mathbf{b}^T \mathbf{b} \rho) = 0, \quad t \geq 0, \quad \mathbf{x} \in \mathbb{R}^d, \quad (\text{II.11a})$$

$$\rho(0, \mathbf{x}) = \mathring{\rho}(\mathbf{x}), \quad (\text{II.11b})$$

where $\mathring{\rho}$ is the pdf of the random variable $\mathring{\mathbf{X}}$, the initial condition of (II.10). Here $\nabla_{\mathbf{x}} \cdot \mathbf{a}$ is the divergence of the vector \mathbf{a} . On the other hand, $\nabla_{\mathbf{x}}^2(s)$ is the Laplacian of the scalar s . That is

$$\nabla_{\mathbf{x}} \cdot \mathbf{a} = \sum_i \partial_{x_i} a_i, \quad (\text{II.12a})$$

$$\nabla_{\mathbf{x}}^2 s = \sum_i \partial_{x_i}^2 s. \quad (\text{II.12b})$$

where $\partial_x(\cdot)$ and $\partial_x^2(\cdot)$ denote the first and second partial derivative with respect to \mathbf{x} , respectively.

II.1.2 Mean-Field Limit

Consider an SDE system of $2P$ coupled equations in \mathbb{R}^d describing P interacting particles, where for each particle i we have

$$d\mathbf{X}_i(t) = \mathbf{V}_i(t) dt, \quad (\text{II.13a})$$

$$d\mathbf{V}_i(t) = \left\{ \mathbf{F}(\mathbf{X}_i(t), \mathbf{V}_i(t)) + \frac{1}{P} \sum_{\substack{j=1 \\ j \neq i}}^P \mathbf{H}(\mathbf{X}_i(t) - \mathbf{X}_j(t), \mathbf{V}_i(t) - \mathbf{V}_j(t)) \right\} dt \\ + \sigma(\mathbf{X}_i(t), \mathbf{V}_i(t)) d\mathbf{W}_i(t), \quad (\text{II.13b})$$

Here σ is a scalar and $\mathbf{W}_i(t)$ is a vector of d independent Wiener processes. Moreover, \mathbf{X}_i and \mathbf{V}_i are the position and velocity of particle i , respectively. We call the \mathbf{F} term a self-drift term, while we refer to the \mathbf{H} term as a coupling term. Also, the initial states of each particle $(\mathbf{X}_i(0), \mathbf{V}_i(0))$ are independent and identically distributed with a common law $\dot{\rho}(\mathbf{x}, \mathbf{v})$. In other words, the particles are indistinguishable and exchangeable because of the symmetry of initial states and of the evolution of these states. Formally (cf. [17]), a sequence of P random variables $\{x_i\}_{i=1}^P$ is said to be exchangeable if the joint law of $(x_{\pi(0)}, x_{\pi(1)}, \dots, x_{\pi(P)})$ is the same as the joint law of (x_0, x_1, \dots, x_P) for any permutation π of the indices $\{1, 2, \dots, P\}$. A trivial example is if $\{x_i\}_{i=1}^P$ are independent and identically distributed then the sequence is exchangeable.

Next if we define the empirical random measure ρ^P as

$$\rho^P(t, \mathbf{x}, \mathbf{v}) = \frac{1}{P} \sum_{i=1}^P \delta(\mathbf{x} - \mathbf{X}_i(t), \mathbf{v} - \mathbf{V}_i(t)), \quad (\text{II.14})$$

where δ is a Dirac delta function. Then we can write the coupling term as

$$\begin{aligned}
& \frac{1}{P} \sum_{j=1}^P \mathbf{H}(\mathbf{X}_i(t) - \mathbf{X}_j(t), \mathbf{V}_i(t) - \mathbf{V}_j(t)) \\
&= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbf{H}(\mathbf{X}_i(t) - \mathbf{x}, \mathbf{V}_i(t) - \mathbf{v}) \rho^P(t, \mathbf{x}, \mathbf{v}) \, d\mathbf{x} \, d\mathbf{v} \\
&= (\mathbf{H} * \rho^P(t, \cdot, \cdot))(\mathbf{X}_i(t), \mathbf{V}_i(t)), \tag{II.15}
\end{aligned}$$

where $\mathbf{H} * \rho^P$ is the convolution between \mathbf{H} and ρ^P .

Assuming sufficient conditions on the initial condition, and on the functions \mathbf{F}, \mathbf{H} and σ and given the indistinguishability of particles it is reasonable to expect that as the number of particles increases the system reaches a limit that we call a mean-field limit. Moreover, it is reasonable to expect that the empirical measure ρ^P will converge to a pdf $\rho(t, \mathbf{x}, \mathbf{v})$ of a generic particle being in a specific position \mathbf{x} and velocity \mathbf{v} . The convergence of ρ^P to ρ is weak convergence in the sense that for all continuous and bounded functions g the following holds

$$\lim_{P \rightarrow \infty} \left| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} g(\mathbf{x}, \mathbf{v}) (\rho^P(t, \mathbf{x}, \mathbf{v}) - \rho(t, \mathbf{x}, \mathbf{v})) \, d\mathbf{x} \, d\mathbf{v} \right| = 0. \tag{II.16}$$

Given such a pdf ρ we look at the system of P particles each described by the following SDE system

$$d\bar{\mathbf{X}}(t) = \bar{\mathbf{V}}(t) \, dt, \tag{II.17a}$$

$$\begin{aligned}
d\bar{\mathbf{V}}(t) = & \{ \mathbf{F}(\bar{\mathbf{X}}(t), \bar{\mathbf{V}}(t)) + (\mathbf{H} * \rho(t, \cdot, \cdot))(\bar{\mathbf{X}}(t), \bar{\mathbf{V}}(t)) \} \, dt \\
& + \sigma(\bar{\mathbf{X}}(t), \bar{\mathbf{V}}(t)) \, d\mathbf{W}(t), \tag{II.17b}
\end{aligned}$$

Observe that all particles satisfy the same SDE system, however the initial state and noise path driving each particle are independent of the initial states and noise paths driving all other particles. Most importantly, observe that while the dimensionality

of (II.13) is $2P$, the dimensionality of the system (II.17) is only 2. This is a huge reduction of dimensionality that justifies approximating the solution of the system (II.13) by the solution to (II.17). Moreover, ρ satisfies the nonlinear integro-PDE, Fokker-Planck equation (cf. Section II.2.3)

$$\partial_t \rho = -\mathbf{v} \cdot \nabla_{\mathbf{x}} \rho - \nabla_{\mathbf{v}} \cdot ((\mathbf{F} + \mathbf{H} * \rho) \rho) + \nabla_{\mathbf{v}}^2 (\sigma \rho), \quad (\text{II.18a})$$

$$\rho(0, \mathbf{x}, \mathbf{v}) = \mathring{\rho}(\mathbf{x}, \mathbf{v}), \quad (\text{II.18b})$$

$$t \geq 0, \quad \mathbf{x} \in \mathbb{R}^d, \quad \mathbf{v} \in \mathbb{R}^d, \quad (\text{II.18c})$$

For example, Bolley et al. [12] proved the existence and uniqueness of such a limit for the specific case when $\sigma(\mathbf{x}, \mathbf{v}) = \sqrt{2}$ and the following boundedness and locally-Lipschitz assumptions

$$\mathbf{v} \cdot \mathbf{F}(\mathbf{x}, \mathbf{v}) \leq C(1 + |\mathbf{v}|^2), \quad (\text{II.19a})$$

$$(\mathbf{v} - \mathbf{w}) \cdot (\mathbf{F}(\mathbf{x}, \mathbf{v}) - \mathbf{F}(\mathbf{x}, \mathbf{w})) \leq L|\mathbf{v} - \mathbf{w}|^2(1 + |\mathbf{v}|^p + |\mathbf{w}|^p), \quad (\text{II.19b})$$

$$|\mathbf{F}(\mathbf{x}, \mathbf{v}) - \mathbf{F}(\mathbf{y}, \mathbf{v})| \leq L|\mathbf{x} - \mathbf{y}|(1 + |\mathbf{v}|^p), \quad (\text{II.19c})$$

$$|\mathbf{H}(\mathbf{x}, \mathbf{v})| \leq C(1 + |\mathbf{v}|), \quad (\text{II.19d})$$

$$|\mathbf{H}(\mathbf{x}, \mathbf{v}) - \mathbf{H}(\mathbf{y}, \mathbf{w})| \leq L(|\mathbf{x} - \mathbf{y}| + |\mathbf{v} - \mathbf{w}|)(1 + |\mathbf{v}|^p + |\mathbf{w}|^p), \quad (\text{II.19e})$$

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (|\mathbf{x}|^2 + e^{|\mathbf{v}|^{p'}}) \mathring{\rho}(\mathbf{x}, \mathbf{v}) \, d\mathbf{x} \, d\mathbf{v} < +\infty, \quad (\text{II.19f})$$

hold for all $\mathbf{x}, \mathbf{y}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^d$ and some positive constants C, L and $0 < p \leq 2$ and $p \leq p'$. Moreover, in this case, Bolley et al. proved the existence and uniqueness of a solution to (II.18). They also proved that the rate of strong convergence of the particles from system (II.13) to the corresponding ones in (II.17) is of $\mathcal{O}(P^{\epsilon-1})$ for any $0 < \epsilon < 1$ as $P \rightarrow \infty$. This means

$$\mathbb{E} \{ \|\bar{\mathbf{X}}_i(T) - \mathbf{X}_i(T)\|^2 + \|\bar{\mathbf{V}}_i(T) - \mathbf{V}_i(T)\|^2 \} = \mathcal{O}(P^{\epsilon-1}), \quad (\text{II.20})$$

for $i = 1 \cdots P$ and any final time $T < \infty$. Here it is assumed that the initial condition of both particles is the same

$$\overline{\mathbf{X}}_i(0) = \mathbf{X}_i(0), \quad (\text{II.21a})$$

$$\overline{\mathbf{V}}_i(0) = \mathbf{V}_i(0), \quad (\text{II.21b})$$

and the same driving noise in $\mathbf{V}_i(t)$ and $\overline{\mathbf{V}}_i(t)$ is used. Finally, the expectation in (II.20) is taken with respect to the measure of the initial conditions and the driving noise.

II.2 Model Formulation

This section discusses different models for crowd simulation in one and two dimensions. Recall that **CaseI** is a test case in which pedestrians move clockwise on a circle with no obstacles. In the more realistic **CaseII**, pedestrians move towards a fixed target in a fixed environment (see Figure II.1). In Section II.2.1 we review the general particle-based model introduced by Helbing [1]. This model is based on the concept of psychological forces, i.e., non-physical forces that drive pedestrians away from each other and from obstacles. A few modifications to the general model are suggested and justified in Section II.2.2. Some of these modifications are necessary for the mean-field limit to exist as the number of pedestrians increases. Other modifications are introduced to make the analysis easier. Section II.2.3 formulates and discusses the mean-field limit in both *CaseI* and *CaseII*.

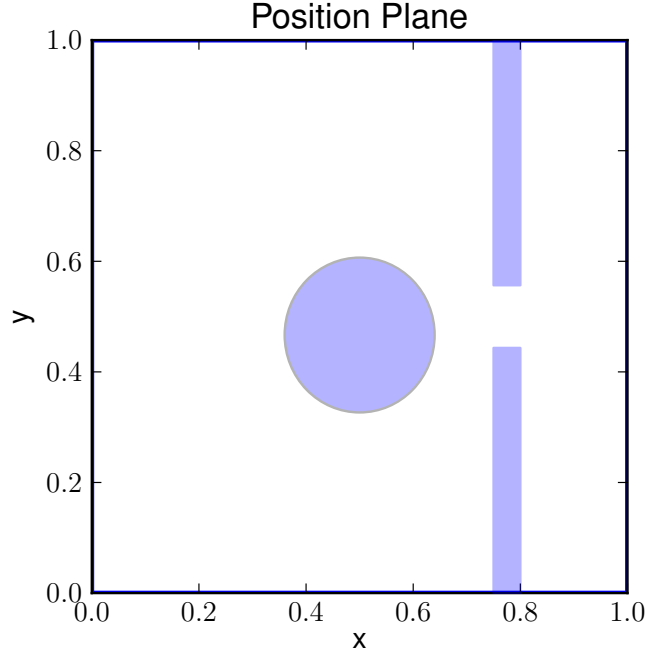


Figure II.1: Typical environment in *CaseII*. The blue areas are obstacles and the exit is through the opening in the right wall.

II.2.1 Helbing Particle-Based Model

In [1], Helbing proposed the following model for pedestrian α

$$d\mathbf{x}_\alpha = \mathbf{v}_\alpha dt, \quad (\text{II.22a})$$

$$d\mathbf{v}_\alpha = \left(\mathbf{F}_\alpha^{\text{ind}}(\mathbf{x}_\alpha, \mathbf{v}_\alpha) + \sum_{\substack{\beta=1 \\ \beta \neq \alpha}}^P J(\mathbf{x}_\beta) \mathbf{F}_{\alpha\beta}^{\text{pair}}(\mathbf{x}_\alpha, \mathbf{v}_\alpha, \mathbf{x}_\beta, \mathbf{v}_\beta) \right) dt + \xi_\alpha(t) dt, \quad (\text{II.22b})$$

where \mathbf{x}_α and \mathbf{v}_α are the position and velocity of pedestrian α , respectively. Here, $\mathbf{F}_\alpha^{\text{ind}}$ is the sum of forces related to a single pedestrian, and $\mathbf{F}_{\alpha\beta}^{\text{pair}}$ is the sum of forces that come from the interactions between two pedestrians α and β . Moreover, $J(\mathbf{x}_\beta)$ is an indicator function that is one if the pedestrian β is interacting with other pedestrians and zero otherwise. For example, in *CaseII*, J is one if the pedestrian has not exited the room yet. In *CaseI*, J is always one. In Helbing model, ξ_α is a random process that adds fluctuations to the movement of pedestrian α . In his works, Helbing only

specifies that the random process ξ_α should be independent of other processes added to movement of other pedestrians. However, in the code that he generously provided, this process is implemented by adding independent increments to each time step in the Forward Euler method. Similar to a Wiener process, these increments are scaled by $\sqrt{\Delta t}$, where Δt is the uniform step size of the Forward Euler approximation. However, they are sampled from a *truncated* Gaussian distribution with parameters $(0, \sigma_\alpha)$ in the interval $[-\sigma_\alpha \eta_\alpha, \sigma_\alpha \eta_\alpha]$ for some constant η_α .

$\mathbf{F}_\alpha^{\text{ind}}$ is split further into forces that come from the interaction of a pedestrian with obstacles and a force $\mathbf{F}_\alpha^{\text{trg}}$ that drives the pedestrian to his target

$$\mathbf{F}_\alpha^{\text{ind}}(\mathbf{x}_\alpha, \mathbf{v}_\alpha) = \mathbf{F}_\alpha^{\text{trg}}(\mathbf{x}_\alpha, \mathbf{v}_\alpha) + \sum_{\mu} I_{\mu}(\mathbf{x}_\alpha) \mathbf{F}_{\alpha\mu}^{\text{obs}}(\mathbf{x}_\alpha, \mathbf{v}_\alpha), \quad (\text{II.23})$$

where $\mathbf{F}_{\alpha\mu}^{\text{obs}}$ is the psychological force of obstacle μ acting on pedestrian α and $I_{\mu}(\mathbf{x})$ is one if a pedestrian at \mathbf{x} is affected by the obstacle μ and zero otherwise. On the other hand, the target force is

$$\mathbf{F}_\alpha^{\text{trg}}(\mathbf{x}_\alpha, \mathbf{v}_\alpha) = \frac{1}{\tau_\alpha} (v_{0,\alpha} \mathbf{e}_\alpha(\mathbf{x}_\alpha) - \mathbf{v}_\alpha). \quad (\text{II.24})$$

Here the parameter $\tau_\alpha > 0$ is the relaxation time. A larger relaxation time means that the pedestrian is less aggressive in going in the direction of his destination \mathbf{e}_α . Also, $v_{0,\alpha} > 0$, assumed constant, is the desired speed that pedestrian α wishes to move with. Observe that the target, as defined by \mathbf{e}_α , does not need to be a fixed point as long as it is a function of the pedestrian's position. This can model, for example, pilgrims circling around the Kaaba, although in that case more work needs to be done to track the number of rounds each pilgrim makes. One way to do this is by adding that information as an additional state variable for each pedestrian and then add more SDEs to describe the evolution of the new state variables. This of course would increase the dimensions of the SDE system.

Each pedestrian α is modelled by a circle with radius r_α . The interactions between two pedestrians $\mathbf{F}_{\alpha\beta}^{\text{pair}}$, or between a pedestrian and an obstacle $\mathbf{F}_{\alpha\mu}^{\text{obs}}$, are described using the same rule, denoted by $\mathbf{F}_{ij}^{\text{int}}$. Let \mathbf{n}_{ij} be a unit normal vector from i to j , and \mathbf{t}_{ij} a unit vector orthogonal to \mathbf{n}_{ij} . Also, let D_{ij} be the *signed* distance between i to j along \mathbf{n}_{ij} (see Figure II.2). This distance is positive if objects i and j do not intersect, and negative otherwise. Finally, let \mathbf{v}_i and \mathbf{v}_j be the velocities of i and j , and let the velocity of a static obstacle be 0. Then we define

$$\mathbf{F}_{ij}^{\text{int}} = \theta(R - D_{ij})\mathbf{F}_{ij}^{\text{psy}} + \theta(-D_{ij}) \left[\mathbf{F}_{ij}^{\text{bdy}} + \mathbf{F}_{ij}^{\text{fric}} \right] \quad (\text{II.25a})$$

$$\mathbf{F}_{ij}^{\text{psy}} = Ae^{-D_{ij}/B} \mathbf{n}_{ij} \quad (\text{II.25b})$$

$$\mathbf{F}_{ij}^{\text{bdy}} = -2CD_{ij} \mathbf{n}_{ij} \quad (\text{II.25c})$$

$$\mathbf{F}_{ij}^{\text{fric}} = \begin{cases} -\gamma [(\mathbf{v}_i - \mathbf{v}_j) \cdot \mathbf{t}_{ij}] \mathbf{t}_{ij} & \text{or,} \\ \kappa D_{ij} [(\mathbf{v}_i - \mathbf{v}_j) \cdot \mathbf{t}_{ij}] \mathbf{t}_{ij} \end{cases} \quad (\text{II.25d})$$

given A, B, C, R, γ and κ are positive constants. Observe that the constants that describe the forces might be different for each of the two cases: pedestrian-pedestrian interaction and pedestrian-obstacle interaction. Moreover, θ here denotes the Heaviside function, that is

$$\theta(x) = \begin{cases} 0 & x < 0, \\ 1 & x \geq 0. \end{cases} \quad (\text{II.26})$$

Observe that $\mathbf{F}_{ij}^{\text{psy}}$ is a psychological force that is activated if the other object is within an interaction radius R . This force pushes pedestrians away from each other and from obstacles and decays exponentially as the distance between them, D_{ij} , increases. On the other hand, $\mathbf{F}_{ij}^{\text{bdy}}$ and $\mathbf{F}_{ij}^{\text{fric}}$ are the body force and the sliding friction force, respectively. They are activated only if D_{ij} is less than or equal to zero; that is, only when a pedestrian makes contact with another pedestrian or with

an obstacle. This makes those forces essential in escape situations, where contact between pedestrians is expected. The body force counteracts body compression in the normal direction, while the friction force opposes relative tangential motion. Note that Helbing suggests two choices for the friction force, one that depends on the distance D_{ij} and one that does not.

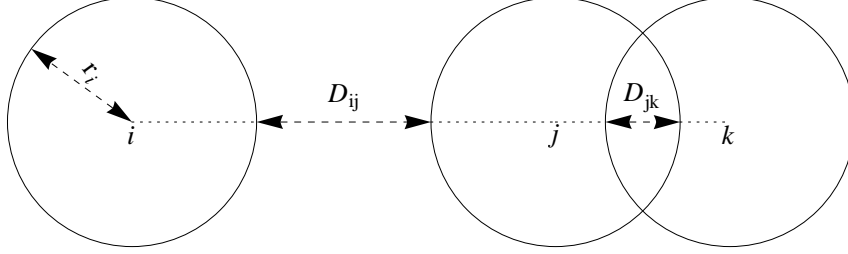


Figure II.2: Signed distance between two circles modelling two pedestrians. Here D_{ij} is positive and D_{jk} is negative.

Finally, the pedestrian model described until now allows for unlimited velocities of pedestrians. Helbing removes this unphysical behaviour by applying a post-processing step to the forward Euler scheme solving (II.22) to cut-off the velocities within a given maximum speed s_{\max} . This is essentially a projected Euler method [18].

II.2.2 Modified Particle-Based Model

In this section we modify the model (II.22) so that it is of the form (II.13). We begin by ensuring the interchangeability of the particles evolution by sampling the initial positions and velocities from the same initial distribution $\mathring{\rho}$ and choosing common parameters for all pedestrians

$$\sigma_\alpha = \sigma, \quad \tau_\alpha = \tau, \quad v_{0,\alpha} = v_0, \quad r_\alpha = r.$$

Similarly, all rules must be symmetric so they will be written as $f_\alpha(\mathbf{x}_\alpha, \mathbf{v}_\alpha) = f(\mathbf{x}_\alpha, \mathbf{v}_\alpha)$. These interchangeability or indistinguishability restrictions may be relaxed by intro-

ducing groups or by including the different parameters in the system state, with dynamics describing the evolution of these parameters. This generalization is beyond the scope of this project.

Moreover, looking at (II.22), the second term in the advection of (II.22b) is the only term coupling different pedestrians. In that term, \mathbf{F}^{pair} has contributions from three forces \mathbf{F}^{psy} , \mathbf{F}^{bdy} and \mathbf{F}^{frc} , all of them discontinuous since they are multiplied by Heaviside functions (see (II.25a)). This makes these forces inherently local and will violate the assumptions in (II.19) even for bounded domains. For these reasons we drop the \mathbf{F}^{bdy} and \mathbf{F}^{frc} terms. We also drop the multiplication of \mathbf{F}^{psy} by a Heaviside function. Instead, the interaction is assumed to decay smoothly and exponentially as the distance increases; as is the case for the current definition of the psychological force (II.25b). This has the same effect of a limited interaction radius. In summary, we propose the simplification $\mathbf{F}_{\alpha\beta}^{\text{pair}} = \mathbf{F}_{\alpha\beta}^{\text{psy}}$.

Another modification to the second term in the advection of (II.22b) is also necessary to ensure that it is of the form (II.13b), namely we need to normalize the coupling term by the number of interacting pedestrian in the system $P_J \leq P$ defined as

$$P_J = \sum_{\beta=1}^P J(x_\beta), \quad (\text{II.27})$$

From a modelling perspective, this can be justified by arguing that pedestrians tend to increase their social circle when there are few pedestrians in the same area. When many pedestrians are present, they tend to accept small social circles.

Also, instead of using the random process ξ_α we will use the standard Wiener process W_α . This will allow us to use the established theory discussed in Sections II.1.1 and II.1.2.

Next we replace the post-processing step of Helbing to bound the velocities by a continuous decay of the advection and diffusion term in (II.22b). This is justified by arguing that the acceleration of pedestrians is decreased even before the maximum

speed is attained. Moreover, this will eliminate the need for any boundary conditions in the model and will greatly simplify the analysis. We achieve the continuous decay by multiplying the advection and diffusion terms in (II.22b) by a scalar function φ that takes values on $[0, 1]$. This is a function of how close the velocity can get to the maximum speed given that the particle is subject to diffusion or advection in a specific direction.

Summing up, the final modified particle-based system becomes

$$d\mathbf{x}_\alpha = \mathbf{v}_\alpha dt, \quad (\text{II.28a})$$

$$d\mathbf{v}_\alpha = \vartheta(\mathbf{X}, \mathbf{V}, \mathbf{x}_\alpha, \mathbf{v}_\alpha) dt + \varsigma(\mathbf{v}_\alpha) d\mathbf{W}_\alpha, \quad (\text{II.28b})$$

$$\vartheta(\mathbf{X}, \mathbf{V}, \mathbf{x}_\alpha, \mathbf{v}_\alpha) = \varphi(\delta_a(\mathbf{v}_\alpha, \check{\mathbf{a}}(\mathbf{X}, \mathbf{V}, \mathbf{x}_\alpha, \mathbf{v}_\alpha))) \mathbf{a}(\mathbf{X}, \mathbf{V}, \mathbf{x}_\alpha, \mathbf{v}_\alpha), \quad (\text{II.28c})$$

$$\varsigma(\mathbf{v}_\alpha) = \sigma\varphi(\delta_d(\mathbf{v}_\alpha)), \quad (\text{II.28d})$$

$$\mathbf{a}(\mathbf{X}, \mathbf{V}, \mathbf{x}_\alpha, \mathbf{v}_\alpha) = \mathbf{F}^{\text{ind}}(\mathbf{x}_\alpha, \mathbf{v}_\alpha) - \frac{1}{P_J} J(\mathbf{x}_\alpha) \mathbf{F}^{\text{psy}}(\mathbf{x}_\alpha, \mathbf{x}_\alpha) + \frac{1}{P_J} \sum_{\beta=1}^{P_J} J(\mathbf{x}_\beta) \mathbf{F}^{\text{psy}}(\mathbf{x}_\alpha, \mathbf{x}_\beta), \quad (\text{II.28e})$$

with the same definition of \mathbf{F}^{ind} , \mathbf{F}^{psy} and J as before and where \mathbf{X}, \mathbf{V} are vectors containing the positions and velocities of all particles, respectively. Here $\check{\mathbf{a}}$ is defined as a normalization of \mathbf{a} , that is $\check{\mathbf{a}} = \frac{\mathbf{a}}{\|\mathbf{a}\|}$. Moreover, ϑ and ς denote the modified drift and noise coefficients, respectively. Note that \mathbf{F}^{psy} depends only on the positions not velocities of pedestrians. We also choose φ for some value of $\zeta > 0$ as follows

$$\varphi(t) = \begin{cases} 0, & t \leq 0, \\ \exp\left(\frac{-\zeta}{t^2}\right), & t > 0; \end{cases} \quad (\text{II.29})$$

see for instance Figure II.3.

The function δ_d measures how close the velocity can get to the maximum imposed speed given that it is subject to diffusion. Since the diffusion term is symmetric, i.e.

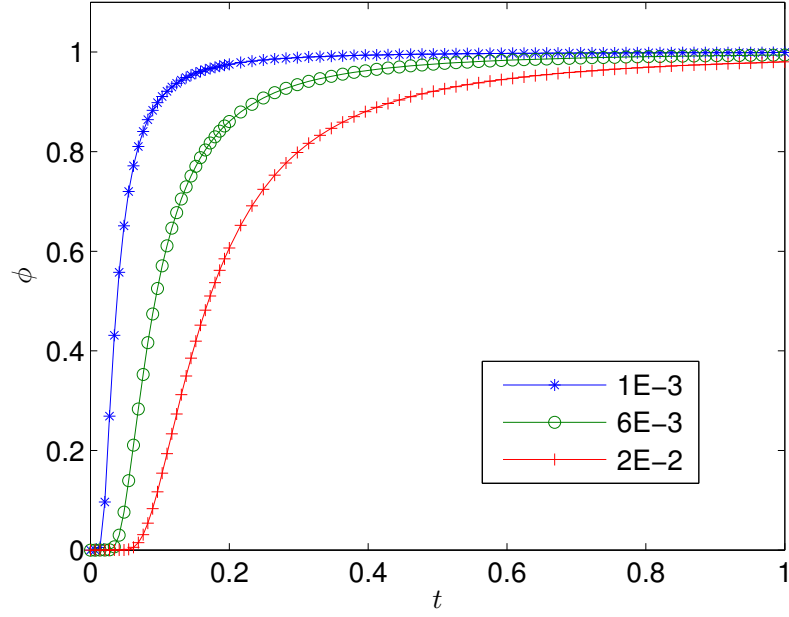


Figure II.3: Cut-off function φ for different values of the parameter ζ .

it is the same in all directions, δ_d is simply taken to be

$$\delta_d(\mathbf{v}) = \max(0, s_{\max} - \|\mathbf{v}\|), \quad (\text{II.30})$$

regardless of the diffusion coefficient. On the other hand δ_a measures how close the velocity can get to the maximum imposed speed given that it is subject to advection in a specific direction $\check{\mathbf{a}}$. In *CaseII*, $\delta_a(\mathbf{v}, \check{\mathbf{a}})$ is defined as the distance along the unit vector $\check{\mathbf{a}}$ between the velocity \mathbf{v} and the circle of radius s_{\max} , see Figure II.4, namely

$$\delta_a(\mathbf{v}, \check{\mathbf{a}}) = \begin{cases} 0 & , \Delta \leq 0, \\ \max(0, -va_1 - wa_2 + \sqrt{\Delta}) & , \Delta > 0, \end{cases} \quad (\text{II.31a})$$

$$\text{with } \Delta = s_{\max}^2 + 2a_1a_2vw - w^2 + (w^2 - v^2)a_2^2, \quad (\text{II.31b})$$

where $\mathbf{v} = [v, w]$ and $\check{\mathbf{a}} = [a_1, a_2]$. Note that the distance is assumed to be zero if the

vector from \mathbf{v} along $\check{\mathbf{a}}$ does not intersect the circle of radius s_{\max} . In Figure II.4, even though \mathbf{v}_2 has a magnitude greater than s_{\max} , the advection will tend to decrease the value of $\|\mathbf{v}_2\|$. Therefore \mathbf{v}_2 is considered to be “far” from the maximum speed s_{\max} and the advection is not reduced. On the other hand, $\delta_a(\mathbf{v}_3, \check{\mathbf{a}}) = 0$ and the advection is reduced to zero, since in this case the unmodified advection will increase the value of $\|\mathbf{v}_3\|$.

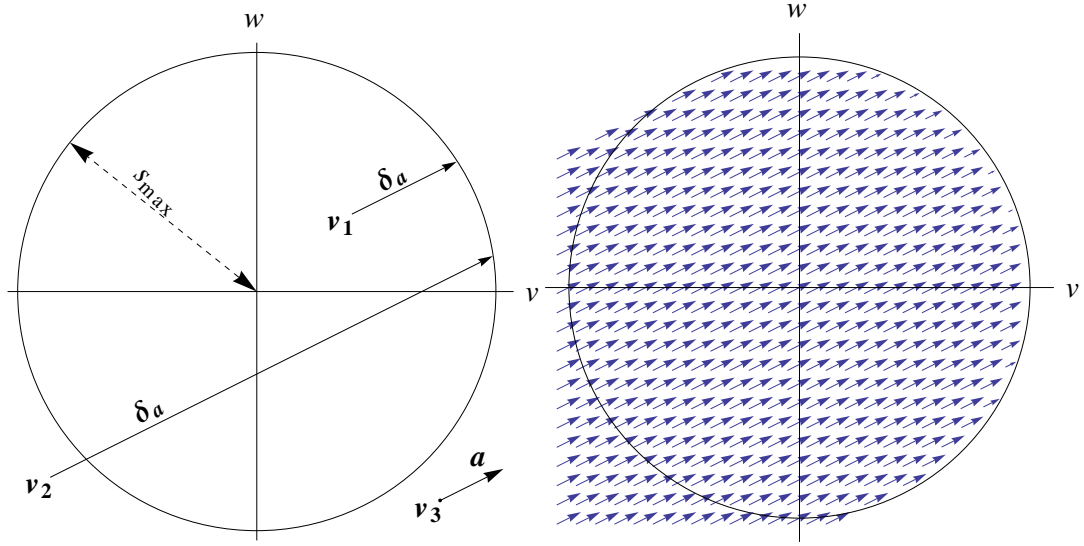


Figure II.4: Left: δ_a shown at three velocities in the direction of $\check{\mathbf{a}}$. Note that $\delta_a(\mathbf{v}_3, \check{\mathbf{a}}) = 0$ since the vector from it in the direction of $\check{\mathbf{a}}$ does not intersect the circle. Right: Modified advection field

In *CaseI* there are only two directions so $\check{a} \in \{-1, +1\}$ and we define

$$\delta_a(v, \check{a}) = \max(0, s_{\max} - \check{a}v). \quad (\text{II.32})$$

Observe that with these choices of φ , δ_d and δ_a and if we also assume that the initial value of \mathbf{v}_α satisfies $\|\mathbf{v}_\alpha(0)\| \leq s_{\max}$ then $\|\mathbf{v}_\alpha(t)\| \leq s_{\max}$ for all t .

One last detail is in *CaseI* where we model pedestrians moving on a circle. In this case, care must be taken in calculating the normal vector and the distance between two pedestrians. Specifically, such calculations must be done on a circle instead of on

a line. If the circle on which the pedestrians move has perimeter x_{\max} then

$$\nu(x_\alpha, x_\beta) = \text{mod}(x_\alpha - x_\beta + \frac{x_{\max}}{2}, x_{\max}) - \frac{x_{\max}}{2}, \quad (\text{II.33a})$$

$$n_{\alpha\beta} = \text{sgn}(\nu(x_\alpha, x_\beta)), \quad (\text{II.33b})$$

$$D_{\alpha\beta} = |\nu(x_\alpha, x_\beta)| - (r_\alpha + r_\beta), \quad (\text{II.33c})$$

where $\text{sgn}(\cdot)$ is the signum function and $\text{mod}(\cdot, \cdot)$ is the modulo function defined as

$$\text{mod}(a, b) = a - \left\lfloor \frac{a}{b} \right\rfloor b, \quad (\text{II.34})$$

here $\lfloor \cdot \rfloor$ is the floor function.

II.2.3 Mean Field Limit

The system (II.28) is of the same form as the system (II.13) with $\mathbf{F}^{\text{ind}} = \mathbf{F}$ and $\mathbf{F}^{\text{pair}}(x_\alpha, x_\beta) = \mathbf{H}(x_\alpha - x_\beta)$. Except for the fact that, in *Case II*, the coupling term is divided by P_J instead of the total number of particles P . Also, observe that here \mathbf{H} does not depend on the velocity difference. Recall that as the number of particles increases the underlying empirical measure ρ^P (cf. (II.14)) of the system (II.13) converges to a marginal distribution ρ describing the state of a generic particle. However, this convergence required certain assumptions on the functions \mathbf{F} and \mathbf{H} , namely (II.19). Looking at $\mathbf{H} = \mathbf{F}^{\text{psy}}$ we note that it satisfies the necessary conditions (II.19d) and (II.19e) since the distance between two pedestrians is bounded. However, $\mathbf{F} = \mathbf{F}^{\text{ind}}$ fails to satisfy the condition (II.19c) because of the discontinuous indicator function I_μ in the definition of \mathbf{F}^{ind} , cf. (II.23). With these differences in mind, we look at the limit marginal distribution ρ . Indeed, if it exists then it should satisfy the

nonlinear Fokker-Planck equation (analogous to (II.18))

$$\partial_t \rho = -\mathbf{v} \cdot \nabla_{\mathbf{x}} \rho - \nabla_{\mathbf{v}} \cdot (\vartheta \rho) + \frac{1}{2} \nabla_{\mathbf{v}}^2 (\varsigma^2 \rho), \quad (\text{II.35a})$$

$$t > 0, \quad \mathbf{x} \in \mathbb{R}^d, \quad \mathbf{v} \in \mathcal{D}_v = [-s_{\max}, s_{\max}]^d,$$

$$\vartheta \rho - \nabla_{\mathbf{v}_n} (\varsigma^2 \rho) \Big|_{\partial \mathcal{D}_v} = 0 \quad (\text{II.35b})$$

$$\rho(0, \mathbf{x}, \mathbf{v}) = \mathring{\rho}(\mathbf{x}, \mathbf{v}), \quad (\text{II.35c})$$

$$\vartheta(\mathbf{x}, \mathbf{v}, [\rho]) = \varphi(\delta_a(\mathbf{v}, \check{\mathbf{a}}(\mathbf{x}, \mathbf{v}, [\rho]))) \mathbf{a}(\mathbf{x}, \mathbf{v}, [\rho]), \quad (\text{II.35d})$$

$$\varsigma(\mathbf{v}) = \sigma \varphi(\delta_d(\mathbf{v})), \quad (\text{II.35e})$$

$$\mathbf{a}(\mathbf{x}, \mathbf{v}, [\rho]) = \mathbf{F}^{\text{ind}}(\mathbf{x}, \mathbf{v}) + \mathbf{F}^{\text{psy}} * \rho_{J, \mathbf{x}}(t, \mathbf{x}, [\rho]), \quad (\text{II.35f})$$

$$\rho_{J, \mathbf{x}}(t, \bar{\mathbf{x}}, [\rho]) = \frac{J(\bar{\mathbf{x}}) \int_{\mathcal{D}_v} \rho(t, \mathbf{v}, \bar{\mathbf{x}}) \, d\mathbf{v}}{\int_{\mathbb{R}^d} J(\mathbf{x}) \int_{\mathcal{D}_v} \rho(t, \mathbf{v}, \mathbf{x}) \, d\mathbf{v} \, d\mathbf{x}}. \quad (\text{II.35g})$$

Observe that \mathbf{v} is in $[-s_{\max}, s_{\max}]^d$ only since the original stochastic process satisfies $\|\mathbf{v}(t)\| \leq s_{\max}$ for all t . Here, $\delta \mathcal{D}_v$ is the boundary of the domain \mathcal{D}_v and $\nabla_{\mathbf{v}_n}$ is the normal derivative in the v -dimension. Finally, note that the argument $[p]$ was added to the definition of \mathbf{a} and ϑ to indicate their dependence on the underlying marginal pdf.

Chapter III

Numerical Simulations and Solutions

In this chapter we describe the numerical methods that are used to solve the particle-based model (II.28) and the continuous model (II.35) in both *CaseI* and *CaseII*.

III.1 Computational Methods

In this section we discuss the basic concepts of the used computational methods in a relatively general setting. The discussion in this chapter regarding PDEs and their numerical methods is taken from [19–21], where more details can be found. Also, the discussion about numerical methods of SDEs and Monte Carlo methods can be found in [15, 16].

III.1.1 Numerical Schemes for Solving PDEs

First, we focus on the simpler one-dimensional PDEs and later discuss how we can extend these methods to multidimensional PDEs. Here, one-dimensional PDEs refer to time-dependent PDEs in one space dimension. In this thesis, we mainly look at

two basic PDEs

$$\partial_t q + \partial_x(aq) = 0, \quad x \in [x_{\text{start}}, x_{\text{end}}] \subseteq \mathbb{R}, \quad (\text{III.1})$$

$$\partial_t q - \partial_x^2(dq) = 0, \quad x \in [x_{\text{start}}, x_{\text{end}}] \subseteq \mathbb{R}, \quad (\text{III.2})$$

with zero influx boundary conditions and an initial profile of q . The first PDE (III.1) is an advection equation with variable coefficient $a(x)$ written in conservative form. The second equation (III.2) is a diffusion equation with variable coefficient $d(x)$.

We discretize the domain $[x_{\text{start}}, x_{\text{end}}] \subseteq \mathbb{R}$ of the space variable x into N_x cells with cell centers x_i and cell edges $x_{i-1/2}$ and $x_{i+1/2}$ for the i 'th cell, where $i = 0, 1, 2, \dots, N_x - 1$. The discretization is assumed to be uniform with mesh size $\Delta x = (x_{\text{end}} - x_{\text{start}})/N_x$. Therefore

$$x_{i-1/2} = x_{\text{start}} + i\Delta x, \quad (\text{III.3})$$

$$x_{i+1/2} = x_{\text{start}} + (i+1)\Delta x, \quad (\text{III.4})$$

$$x_i = x_{i-1/2} + \frac{\Delta x}{2}. \quad (\text{III.5})$$

We also assume a time discretization $0 = t_0 < t_1 < \dots < t_K = T$ for some final time T and denote the step size $\Delta t_k = t_k - t_{k-1}$ and the maximum step size $\Delta t = \max_k \Delta t_k$.

We use two different methods to obtain a numerical approximation to the solution of the PDE (III.1). The first is a first order, implicit, method of lines approach obtained by approximating the spatial derivatives in the upwind direction and then using backward Euler to integrate in time. Therefore, we approximate

$$\partial_x(a(x)q(t, x))|_{x=x_i} = \begin{cases} \frac{a(x_i)q(t, x_i) - a(x_{i-1})q(t, x_{i-1})}{\Delta x} + \mathcal{O}(\Delta x), & a(x_i) > 0, \\ \frac{a(x_{i+1})q(t, x_{i+1}) - a(x_i)q(t, x_i)}{\Delta x} + \mathcal{O}(\Delta x), & a(x_i) < 0, \end{cases} \quad (\text{III.6})$$

where a is assumed to preserve sign throughout the whole domain. If we also assume

zero flux at x_{start} or x_{end} , depending on the sign of a , then we have

$$\begin{aligned} \partial_x(a(x)q(t, x))|_{x=x_0} &= \frac{2a(x_0)q(t, x_0)}{\Delta x} + \mathcal{O}(\Delta x), & \text{if } a(x_0) > 0, \\ \partial_x(a(x)q(t, x))|_{x=x_{N_x-1}} &= \frac{2a(x_{N_x-1})q(t, x_{N_x-1})}{\Delta x} + \mathcal{O}(\Delta x), & \text{if } a(x_{N_x-1}) < 0. \end{aligned} \quad (\text{III.7})$$

$$(\text{III.8})$$

Then we set up the resulting system as

$$\mathbf{q}(t) = \begin{bmatrix} q(t, x_0) & q(t, x_1) & \cdots & q(t, x_{N_x-1}) \end{bmatrix}^T, \quad (\text{III.9})$$

$$A = \text{diag} \left(\begin{bmatrix} a(x_0) & a(x_1) & \cdots & a(x_{N_x-1}) \end{bmatrix} \right), \quad (\text{III.10})$$

$$\mathbf{q}'(t) = -\mathcal{D}A\mathbf{q}(t) + \mathcal{O}(\Delta x), \quad (\text{III.11})$$

$$\mathcal{D} = \frac{1}{\Delta x} \underbrace{\begin{bmatrix} 2 & & & & \\ 1 & 1 & & & \\ & 1 & 1 & & \\ & & \ddots & \ddots & \\ & & & 1 & 1 \end{bmatrix}}_{a>0} \quad \text{or} \quad \frac{1}{\Delta x} \underbrace{\begin{bmatrix} 1 & 1 & & & \\ & 1 & 1 & & \\ & & \ddots & \ddots & \\ & & & 1 & 1 \\ & & & & 2 \end{bmatrix}}_{a<0}. \quad (\text{III.12})$$

Denoting $\mathbf{q}^k \approx \mathbf{q}(t_k)$ we find the numerical approximate solution to the system of ordinary differential equations (ODEs) (III.11) using backward Euler

$$\mathbf{q}^{k+1} = \mathbf{q}^k - \Delta t_{k+1} \mathcal{D}A\mathbf{q}^{k+1}. \quad (\text{III.13})$$

This method, which we will refer to as upwind Backward Euler (uBE), is consistent and stable for any $\Delta t > 0$, cf. [20]. However, it has high numerical diffusion and is only first order accurate in time and space. Moreover, in higher dimensions it is relatively expensive since it requires solving a linear system in each step. Of course, such operations can be optimized by factorizing the matrix once and then using that

factorization to solve the system efficiently in subsequent time steps, assuming that the advection speed a does not depend on time and that uniform time stepping is used, i.e. $\Delta t_k = \Delta t$ for all k .

Another method that we will also use to solve the advection equation (III.1) is the explicit Lax-Wendroff-Leveque method with the Riemann solver for the conservative, variable coefficient advection equation, both discussed in detail in [19]. This method is second order accurate in space and time, conservative, non-oscillatory and has minimal numerical diffusion. Moreover, it does not require a to preserve sign. However, to be stable it has to satisfy the CFL condition, namely

$$\Delta t \leq \left(\max_x |a(x)| \right) \Delta x. \quad (\text{III.14})$$

We use the `clawpack` implementation of this method and its python front-end, `pyclaw` [22].

To obtain a numerical approximation to the solution of the PDE (III.2), we use the implicit Crank-Nicolson (CN) method [20]. This method is based on the method of lines and it is obtained by approximating the spatial derivatives by centered differences and then using the trapezoidal method to integrate in time. Using similar notation as before we write the CN method as

$$D = \text{diag} \left(\begin{bmatrix} d(x_0) & d(x_1) & \cdots & d(x_{N_x-1}) \end{bmatrix} \right), \quad (\text{III.15})$$

$$(I - \frac{1}{2} \Delta t_k \mathcal{D}^2 D) \mathbf{q}^{k+1} = (I + \frac{1}{2} \Delta t_k \mathcal{D}^2 D) \mathbf{q}^k, \quad (\text{III.16})$$

$$\mathcal{D}^2 = \begin{bmatrix} 1 & 1 & & & \\ 1 & 2 & 1 & & \\ & & \ddots & \ddots & \ddots \\ & & & 1 & 2 & 1 \\ & & & & 1 & 1 \end{bmatrix}. \quad (\text{III.17})$$

The CN method is second order accurate in space and time, stable for any Δt and has no numerical second-order diffusion. However, it can be highly oscillatory if applied to an advection equation with high advection speeds and, like the uBE method, requires solving a linear system at each time step.

We now consider multidimensional PDEs. We discretize each dimension using similar ideas to those already discussed in one-dimensional PDEs, and tensorize all discretizations to obtain a multidimensional grid. Then, extending CN and uBE method on this multidimensional grid is relatively straightforward. Leveque also discusses the multidimensional extension of the Lax-Wendroff-Leveque method in [19].

Another method for solving these multidimensional PDEs using one-dimensional schemes is dimensional splitting [20]. Dimensional splitting is a form of fractional step methods to solve more complicated PDEs by combining solution of their simpler parts. To present the method, consider the PDE

$$\partial_t q + \mathcal{F}(q) + \mathcal{G}(q) = 0, \quad (\text{III.18})$$

for two operators \mathcal{F} and \mathcal{G} . For example, if $\mathcal{F}(q) = \partial_x(aq)$ and $\mathcal{G}(q) = \partial_x^2(dq)$ then (III.18) is an advection-diffusion equation. If q is a function of two space variables x and y and $\mathcal{F}(q) = \partial_x(a_x q)$ and $\mathcal{G}(q) = \partial_y(a_y q)$ then (III.18) is a two-dimensional advection equation. The fractional step method works by solving the following two PDEs

$$\partial_t q_1 + \mathcal{F}(q_1) = 0, \quad (\text{III.19a})$$

$$\partial_t q_2 + \mathcal{G}(q_2) = 0, \quad (\text{III.19b})$$

in succession from t_k to t_{k+1} , with appropriate boundary conditions that we discuss

later, and the following initial conditions

$$q_1(t_k, \cdot) = \hat{q}(t_k, \cdot), \quad (\text{III.20a})$$

$$q_2(t_k, \cdot) = q_1(t_{k+1}, \cdot), \quad (\text{III.20b})$$

$$\hat{q}(t_{k+1}, \cdot) = q_2(t_{k+1}, \cdot), \quad (\text{III.20c})$$

where $\hat{q}(t_k, \cdot)$ is an approximation of the solution $q(t_k, \cdot)$ to (III.18). The order of solving these two PDEs is arbitrary and can be reversed. In dimensional splitting each part is solved for each coordinate value of the other dimension. Returning to the two-dimensional advection equation example the first part

$$\partial_t q_1 + \partial_x(a(x, y_i)q_1(t, x, y_i)) = 0, \quad (\text{III.21a})$$

$$q_1(t_k, \cdot, y_i) = \hat{q}(t_k, \cdot, y_i), \quad (\text{III.21b})$$

is solved for each y_i in the multidimensional grid with appropriate boundary conditions. Moreover, the second part

$$\partial_t q_2 + \partial_y(a(x_i, y)q_2(t, x_i, y)) = 0, \quad (\text{III.22a})$$

$$q_2(t_k, x_i, \cdot) = q_1(t_{k+1}, x_i, \cdot), \quad (\text{III.22b})$$

is solved for each x_i with appropriate boundary conditions. Imposing appropriate boundary conditions on the separate parts that are consistent with the imposed boundary condition on the original PDE can be nontrivial. Sometimes the boundary conditions can be naturally split. For example, in multidimensional problems we might have different, already split boundary conditions imposed on the boundaries of each dimension. Also for relatively simple PDEs and associated simple boundary conditions, Leveque [23,24] discusses a general method to obtain boundary conditions on the split parts.

The fractional step method is first order accurate in time and is usually stable if the numerical method of each part is stable [20]. It is a versatile method that, in addition to having useful modularity properties, allows using appropriate methods to solve different parts of the PDE efficiently. For example, consider a one-dimensional advection-diffusion equation. Solving both parts together using the CN method produces oscillatory solutions and is inefficient since it requires solving a linear system at each time step. Moreover, solving this linear system cannot be optimized if the advection coefficient depends on time. On the other hand, solving both parts using the Lax-Wendroff-Leveque method places severe restrictions on the time step size ($\Delta t = \mathcal{O}(\Delta x^2)$) [20]. However, using fractional time stepping, the advection part can be solved efficiently using the explicit Lax-Wendroff-Leveque method and the diffusion part might be solved using the implicit CN method. The output of the two methods can be combined to produce a numerical approximation to the solution of the original problem. In fact, when solving individual parts, we might take more intermediate time steps to satisfy stability requirements. The accuracy will, however, still be determined by the splitting time step size.

Given $d(\mathbf{x})$, consider another example

$$\partial_t q + \nabla_{\mathbf{x}} \cdot (dq) = 0, \quad (\text{III.23})$$

with given boundary conditions. Here \mathbf{x} is a d -dimensional variable. Dimensional splitting can be used to efficiently calculate an approximate solution of this diffusion equation. This is done, for example, by solving a series of one-dimensional equations of the form

$$\partial_t q + \partial_{\mathbf{x}_i} (dq) = 0, \quad (\text{III.24})$$

with appropriate boundary conditions, for each $i = 1 \cdots d$. In other words, this

equation is solved in each dimension by fixing all other dimensions. The solution of each equation is used as the initial condition of the next and we obtain, at the end, an approximate solution of (III.23).

III.1.2 Numerical Schemes for Solving SDEs

Given a time discretization $0 = t_0 < t_1 < \dots < t_K = T$ and denoting $\Delta t_k = t_{k+1} - t_k$ and $\Delta t = \max_k \Delta t_k$ and finally $\Delta W_k = W(t_{k+1}) - W(t_k)$, we simulate an approximation of a single realization of the stochastic process $X(t_k) \approx X_k$ that satisfies the SDE (II.8) by using the following forward Euler discretization

$$X_{k+1} = X_k + a(X_k, t_k)\Delta t_k + b(X_k, t_k)\Delta W_k, \quad \text{for } k = 0, \dots, K-1 \quad (\text{III.25a})$$

$$X_0 = \mathring{X}. \quad (\text{III.25b})$$

Note that ΔW_k is sampled from a normal distribution with zero mean and variance Δt_k (cf. Section II.1.1). Similar ideas apply to the SDE system (II.10). Assuming sufficient regularity conditions on a and b , it is well known (cf. [15]) that X_K converges weakly to $X(T)$ with $\mathcal{O}(\Delta t)$ and strongly with $\mathcal{O}(\Delta t^{1/2})$. That is

$$|\mathbb{E}\{X_K\} - \mathbb{E}\{X(T)\}| = \mathcal{O}(\Delta t), \quad (\text{III.26})$$

$$(\mathbb{E}\{|X_K - X(T)|^2\})^{1/2} = \mathcal{O}(\Delta t^{1/2}). \quad (\text{III.27})$$

The forward Euler method is an explicit method that has certain stability requirements. Specifically, assuming a decaying process, the advection term requires $\Delta t_k < 2|\delta_x a(t_k, X_k)|^{-1}$. In some systems, the value $|\delta_x a|$ can be large, as is the case in (II.28b) and (II.28c) due to the φ factor. In such cases, Δt_k might have to be small for some k and an adaptive method might be more suitable. In other cases, Δt_k might have to be small for all k , in which case an explicit method is compu-

tationally inefficient. However, the noise term cannot be easily handled implicitly and using adaptive time stepping on the full Euler step (III.25) will complicate the implementation due to the use of Brownian bridges [25] and a variable number of random samples of ΔW . For these reasons we use a fractional step method similar to the one discussed in Section III.1.1 in which the advection term is solved separately from the diffusion term,

$$\tilde{X}^{k+1} = \mathcal{A}(X^k, \Delta t_k), \quad (\text{III.28a})$$

$$X^{k+1} = \tilde{X}^{k+1} + b(\tilde{X}^{k+1}, t_k) \Delta W_k, \quad (\text{III.28b})$$

$$X^0 = \overset{\circ}{X}, \quad (\text{III.28c})$$

where $\mathcal{A}(x, t)$ is any stable, consistent numerical method to solve the advection part of (II.8) using x as initial condition and up to time t . For example, \mathcal{A} can be an explicit method with adaptive time stepping or an implicit method.

III.1.3 Monte Carlo Methods

Assume we want to approximate the expectation $\mathbb{E}\{Y\}$ for some random variable Y . If we have an expression for the pdf of Y , $\rho(y)$, then we can compute

$$\mathbb{E}\{Y\} = \int_{\Omega} y \rho(y) \, dy, \quad (\text{III.29})$$

where Ω is the domain of Y . However, in some cases this integral is hard to evaluate, or the explicit expression for ρ might be unknown, for example if Y is a complicated function of a high dimensional random variable. In such cases, assuming we can still sample Y , we can use the Monte Carlo (MC) estimator

$$\hat{\mathbb{E}}_M\{Y\} = \frac{1}{M} \sum_{m=1}^M Y_m, \quad (\text{III.30})$$

for some M , where $\{Y_m\}_{m=0}^M$ are independent samples of Y . Assuming the variance of Y is finite, i.e. $\sigma^2 = \text{Var}\{Y\} < \infty$, the Central Limit Theorem states that

$$\sqrt{M} \left(\hat{E}_M\{Y\} - E\{Y\} \right) \xrightarrow[M \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2), \quad (\text{III.31})$$

Therefore, if we fix a finite M , large enough such that the MC estimator is “close” to a normal random variable, we can with probability $1 - \theta$ bound the error in terms of M

$$P \left\{ |E\{Y\} - \hat{E}_M\{Y\}| \leq \frac{\Phi^{-1}(\theta)\sigma}{\sqrt{M}} \right\} \geq 1 - \theta. \quad (\text{III.32})$$

Here $P\{\cdot\}$ denotes a probability and Φ is the cumulative density function (cdf) of the standard normal distribution. Moreover, the Berry-Esseen theorem [16] determines the speed of the weak convergence in (III.31)

$$|F_M(y) - \Phi(y)| \leq C_{\text{BE}} \frac{E\{|Y - E\{Y\}|^3\}}{\sqrt{M}(1 + |y|)^3\sigma^3}, \quad (\text{III.33})$$

for a fixed constant C_{BE} . Here, $F_M(y)$ is the cdf of $\hat{E}_M\{Y\}$. Observe that the bound in (III.32) depends on the unknown parameter σ , the standard deviation of the random variable Y . However, this parameter can be approximated using the sample standard deviation. Therefore, we start with some number of samples M and then estimate σ and test if the error bound in (III.32) is below a certain tolerance, if not we compute the optimal number of samples and repeat the process. This method can be unstable due to errors in the approximation of the standard deviation. To increase stability, one can limit the increase of the number of samples. For example, a maximum of double the current number of samples can be imposed.

One advantage of Monte Carlo methods is that they are “embarrassingly parallel”, since the computation of each sample Y_m is, and should be, completely independent of other samples. Moreover, the cost of computing each of these samples is usually

comparable to others. Therefore, a Monte Carlo method can be parallelized easily by computing equal batches of samples on different computational nodes and then combining their results on a central node. One technical, yet very important, detail is to ensure that the pseudo-random number generators in different computation nodes are not correlated. Random seeds in different computation nodes run into that risk. One way to resolve this issue is to use a parallel random number generator like SPRNG [26]. Another way is to have a central node for generating samples of random variable and send batches of random samples to computation nodes. The latter method is efficient and easy to implement when the exact number of needed random samples is fixed and can be computed beforehand, and the work for generating the random samples is small compared to evaluation of Y_m .

III.1.4 Multilevel Monte Carlo

Again, assume we want to calculate the expectation $E\{Y\}$, but this time in the specific case where $Y = \psi(X(T))$ for some function ψ and final time T , where $X(t)$ is a random process described by an SDE. We can construct a similar estimator to (III.30), however this assumes that we can sample the random variable $X(T)$. For most SDE systems this is usually not the case and all we can do is approximate a sample of $X(T)$ using for example forward Euler with uniform time step size Δt as discussed in Section III.1.2. We denote by $Y_{\Delta t;m}$ the m 'th sample of Y calculated by approximating $X(T)$ using a Δt time step size. Therefore, we have two sources of errors: discretization error from the forward Euler approximation with $\mathcal{O}(\Delta t)$ and a statistical error from the Monte Carlo estimator with $\mathcal{O}(\sqrt{M})$, as shown in (III.32). That is, with certain probability we have

$$|E\{Y\} - \frac{1}{M} \sum_{m=1}^M Y_{\Delta t;m}| = \mathcal{O}(\Delta t) + \mathcal{O}(\sqrt{M}). \quad (\text{III.34})$$

Thus, to achieve a certain accuracy ϵ we have to choose $\Delta t \propto \mathcal{O}(\epsilon)$ and $M \propto \mathcal{O}(\epsilon^{-2})$ and the total work would be then $\mathcal{O}(\epsilon^{-3})$.

In the context of Itô SDEs, Giles [27] proposed the following multilevel Monte Carlo (MLMC) estimator to reduce the cost above from $\mathcal{O}(\epsilon^{-3})$ to $\mathcal{O}(\epsilon^{-2}(\log(\epsilon))^2)$,

$$\hat{\mathbb{E}}_{ML}\{Y\} = \sum_{l=0}^L Z_l, \quad (\text{III.35a})$$

with

$$Z_0 = \frac{1}{M_0} \sum_{m=1}^{M_0} Y_{\Delta t_0;m}, \quad (\text{III.35b})$$

$$Z_l = \frac{1}{M_l} \sum_{m=1}^{M_l} Y_{\Delta t_{l-1};m} - Y_{\Delta t_l;m}, \quad 0 < l \leq L, \quad (\text{III.35c})$$

where in Z_l both $Y_{\Delta t_{l-1};m}$ and $Y_{\Delta t_l;m}$ use the same sample of initial conditions and noise path $W(t)$. Moreover, the noise used in computing Z_l is independent of the noise used in computing $\{Z_k\}_{k=0, k \neq l}^L$. The basic idea behind MLMC is to reduce the variance of the Monte Carlo estimators to subsequently reduce the number of needed samples to bound the statistical error, while still keeping the discretization error under control. This is achieved by using deeper levels with finer discretization and few samples to reduce the discretization error and using other levels with coarse discretization and more samples to reduce the variance. Giles also proposed the choice $\Delta t_l = \mathcal{O}(M_l^{-1})$ and discussed a method for choosing M_l and the number of levels L to achieve a desired accuracy ϵ with high probability.

III.2 Implementation

In this section we discuss how we use the computational methods reviewed in Section III.1 to obtain numerical approximations to the solutions of the models discussed in Section II.2.

III.2.1 Particle-Based Simulation

We begin by calculating approximate solutions to the modified particle-based model discussed in II.2.2.

Simulating a single realization

For simulating a single realization of (II.28) with a uniform time step size Δt , we use the fractional step method (cf. (III.28))

$$\mathbf{x}_\alpha^{k+1} = \mathbf{x}_\alpha^k + \mathbf{v}_\alpha^k \Delta t, \quad (\text{III.36a})$$

$$\tilde{\mathbf{v}}_\alpha^{k+1} = \mathcal{FE}(\mathbf{X}^{k+1}, \mathbf{V}^k, \mathbf{x}_\alpha^{k+1}, \mathbf{v}_\alpha^k, \Delta t), \quad (\text{III.36b})$$

$$\mathbf{v}_\alpha^{k+1} = \tilde{\mathbf{v}}_\alpha^{k+1} + \varsigma(\tilde{\mathbf{v}}_\alpha^{k+1}) \Delta \mathbf{W}_\alpha^k, \quad (\text{III.36c})$$

for $\alpha = 1 \cdots P$. Here $\mathbf{X}^k = \{\mathbf{x}_\alpha^k\}_{\alpha=1}^P$ and $\mathbf{V}^k = \{\mathbf{v}_\alpha^k\}_{\alpha=1}^P$. Moreover, $\mathcal{FE}(\mathbf{X}, \mathbf{V}, \mathbf{x}, \mathbf{v}, \Delta \hat{t})$ is an adaptive forward Euler method that numerically solves the advection part of (II.28b) for a single time step of size $\Delta \hat{t}$ starting with the initial value \mathbf{v} , such that the solution remains “stable”; possibly by using many intermediate time steps. Usually, adaptivity is used to achieve certain accuracy by measuring the error using a method such as Richardson extrapolation [20]. In this case we decrease the size of the time step if the solution is unstable. A good measure of stability of this system is if the velocity magnitude does not increase above s_{\max} . That is, either the velocity magnitude decreases or stays bounded by s_{\max} . This, as we know from Section II.2.2, is a qualitative property of the exact solution to (II.28) due to the multiplication by the cut-off factor φ . In summary, the adaptive algorithm of $\mathcal{FE}(\mathbf{X}, \mathbf{V}, \mathbf{x}, \mathbf{v}, \Delta \hat{t})$ is

1. Set $\Delta t_0 = \Delta \hat{t}$, $t = 0$, $k = 0$ and $\tilde{\mathbf{v}}_0 = \mathbf{v}$.
2. Calculate $\mathbf{z} = \tilde{\mathbf{v}}_k + \vartheta(\mathbf{X}, \mathbf{V}, \mathbf{x}, \tilde{\mathbf{v}}_k) \Delta t_k$
3. If $\|\mathbf{z}\| \leq \max(\|\tilde{\mathbf{v}}_k\|, s_{\max})$ then set $\tilde{\mathbf{v}}_{k+1} = \mathbf{z}$, $t = t + \Delta t_k$ and $\Delta t_{k+1} = \Delta \hat{t} - t$.

4. Else set $\Delta t_{k+1} = \Delta t_k/2$ and go to 2.
5. If $\Delta t_{k+1} = 0$ then return $\tilde{\mathbf{v}}_{k+1}$.
6. Otherwise set $k = k + 1$ and go to 2.

To prevent the algorithm from looping infinitely, we abort it when $\Delta t_k < \sqrt{\epsilon_M}$ where ϵ_M is machine epsilon. In such a case, we revert to a projected Euler (cf. [18]) method where we return the maximum velocity in the direction $\vartheta(\mathbf{X}, \mathbf{V}, \mathbf{x}, \tilde{\mathbf{v}}_k)$.

Computing Quantities of Interest

The main goal here is to calculate a vector of certain quantities of interest or observables that depend on the states of all P particles, such as the average position or average velocity. To this end, we use the multilevel Monte Carlo method with a certain specified tolerance to control the statistical and discretization errors as discussed in Section III.1.4 and [27]. We also use a parallel implementation of MLMC in which Monte Carlo samples are simulated on different computational nodes.

III.2.2 Numerical Solution of the Continuous Model

Next, we look at the numerical solution of the PDE (II.35). In *CaseI*, this is a two-dimensional PDE in space. In *CaseII*, the PDE is four-dimensional in space. General desirable properties for the numerical methods include positivity and mass-conservation. First, the numerical methods discussed in Section III.1.1 approximate solutions to PDEs over bounded domains. As such we must truncate the unbounded position domain of the PDE (II.35). Specifically we solve (II.35) over $\mathbf{x} \in [0, x_{\max}]^d$ and $\mathbf{v} \in [-s_{\max}, s_{\max}]^d$ where $d = 1, 2$ for the *CaseI* and *CaseII*, respectively. When dealing with bounded domains we have to introduce conditions at the domain boundaries, in this case these conditions are called artificial boundary conditions since they are imposed to solve the PDE numerically and are not part of the original PDE. These

artificial boundary conditions have to be consistent with the solution to the PDE with the unbounded domain when the initial conditions are bounded. In *CaseI* we impose periodic boundary conditions in the x -dimension. Moreover, in *CaseII* (see Figure II.1), we impose zero inflow on the top, left and bottom boundaries. These boundaries will not have an outflow since the walls will push the density away from boundaries. On the other hand, we do not impose any boundary conditions on the right boundary since this is an outflow boundary. We emphasize that the obstacles inside the domain are not considered “boundaries” in the sense of the numerical scheme, since they are part of the solution domain and hence the solution is approximated inside them. As such no conditions are imposed near the obstacles.

Dimensional splitting is used to update the position and velocity parts separately. Moreover, we use the fractional step method to solve the \mathbf{v} -advection and \mathbf{v} -diffusion separately. This is an efficient method in this case, since diffusion and advection have very different stability requirements, and time-splitting will allow us to use different methods well suited for these different operations.

In summary, in *CaseI* we solve the following PDEs that are each one-dimensional in space

$$\partial_t \rho_1 + v \cdot \nabla_x \rho_1 = 0, \quad (\text{III.37a})$$

$$\rho_1(t, 0, \cdot) = \rho_1(t, x_{\max}, \cdot),$$

$$\partial_t \rho_2 + \nabla_v \cdot (\vartheta \rho_2) = 0, \quad (\text{III.37b})$$

$$\rho_2(t, \cdot, -s_{\max}) = 0, \quad \text{or} \quad \rho_2(t, \cdot, s_{\max}) = 0,$$

$$\partial_t \rho_3 - \frac{1}{2} \nabla_v^2 (\varsigma^2 \rho_3) = 0, \quad (\text{III.37c})$$

$$\partial_v \rho_3(t, \cdot, -s_{\max}) = \partial_v \rho_3(t, \cdot, +s_{\max}) = 0,$$

In *CaseII* we solve the following PDEs that are each two-dimensional in space

$$\partial_t \rho_1 + \mathbf{v} \cdot \nabla_{\mathbf{x}} \rho_1 = 0, \quad (\text{III.38a})$$

$$\rho_1(t, [0, \cdot], \cdot) = \rho_1(t, [\cdot, 0], \cdot) = \rho_1(t, [\cdot, x_{\max}], \cdot) = 0,$$

$$\partial_t \rho_2 + \nabla_{\mathbf{v}} \cdot (\vartheta \rho_2) = 0, \quad (\text{III.38b})$$

$$\rho_2(t, \cdot, [-s_{\max}, \cdot]) = 0, \quad \text{or} \quad \rho_2(t, \cdot, [+s_{\max}, \cdot]) = 0,$$

$$\text{or} \quad \rho_2(t, \cdot, [\cdot, -s_{\max}]) = 0, \quad \text{or} \quad \rho_2(t, \cdot, [\cdot, +s_{\max}]) = 0,$$

$$\partial_t \rho_3 - \frac{1}{2} \nabla_{\mathbf{v}}^2 (\varsigma^2 \rho_3) = 0, \quad (\text{III.38c})$$

$$\rho_2(t, \cdot, [-s_{\max}, \cdot]) = \rho_2(t, \cdot, [+s_{\max}, \cdot]) = 0,$$

$$\rho_2(t, \cdot, [\cdot, -s_{\max}]) = \rho_2(t, \cdot, [\cdot, +s_{\max}]) = 0.$$

In both (III.37) and (III.38), the PDEs are solved in succession from t_k to t_{k+1} and using the following initial conditions

$$\rho_1(t_k, \cdot, \cdot) = \hat{\rho}(t_k, \cdot, \cdot), \quad (\text{III.39a})$$

$$\rho_2(t_k, \cdot, \cdot) = \rho_1(t_{k+1}, \cdot, \cdot), \quad (\text{III.39b})$$

$$\rho_3(t_k, \cdot, \cdot) = \rho_2(t_{k+1}, \cdot, \cdot), \quad (\text{III.39c})$$

$$\hat{\rho}(t_{k+1}, \cdot, \cdot) = \rho_3(t_{k+1}, \cdot, \cdot), \quad (\text{III.39d})$$

where $\hat{\rho}$ is the approximate solution of (II.35). The order of solving these parts is arbitrary and the same accuracy can be achieved by using a different ordering.

The selection of boundary conditions on ρ_2 in both (III.37b) and (III.38b) depend on the value of the flux at the boundaries. Specifically, these conditions are imposed only if the flux at the boundaries is inward. Observe that the same previously discussed artificial boundary conditions in \mathbf{x} -dimension are imposed unchanged on ρ_1 in (III.37a) and (III.38a). On the other hand, in general, because of the nonlinear

and nonlocal dependence on ρ , deriving boundary conditions on ρ_2 and ρ_3 that are equivalent to the zero-flux boundary condition on ρ in the \mathbf{v} -dimension is a hard problem [23,24]. However, this problem is simpler in our case since the advection and diffusion terms are multiplied by the cut-off function φ which equals zero at the outflow boundaries. In (III.37c) and (III.38c) the PDE trivially become $\partial_t \rho_3 = 0$ at the boundaries of the \mathbf{v} -dimension, therefore boundary conditions are in fact unnecessary for the well-posedness of the system. Similarly, at the outflow boundaries, (III.37b) and (III.38b) also trivially become $\partial_t \rho_2 = 0$, and therefore boundary conditions are not needed here either. At inflow boundaries, imposing zero influx is compatible with the suggested zero-flux boundary conditions in the \mathbf{v} -dimension.

We adopt the discretization discussed in Section III.1.1 with the same notation. For each \mathbf{v} on the grid, (III.37a) and (III.38a) are transport equations with constant coefficients. To obtain a numerical approximation of their solutions we use the Lax-Wendroff-Leveque scheme described in [19] and implemented in `Clawpack` and its python front-end, `pyclaw` [22]. Moreover, for each \mathbf{x} on the grid, we approximate the solutions of (III.37c) and (III.38c) by using the Crank-Nicholson method with zero flux boundary conditions as discussed in Section III.1.1. To numerically solve (III.37b) and (III.38b) we first linearize these equations by evaluating ϑ using $\rho_2(t_k, \cdot, \cdot)$. With this approximation, for every \mathbf{x} on the grid, (III.37b) and (III.38b) are advection equations with variable coefficients which can be solved numerically using the explicit Lax-Wendroff-Leveque method. However, for this explicit method, high advection speeds might place severe restrictions on the time step size, as discussed in Section III.1.1. For this reason we use the uBE implicit method to numerically solve this PDE when the advection speeds are big enough such that the Lax-Wendroff-Leveque method is inefficient. In fact, since we have to solve (III.37b) and (III.38b) for every position \mathbf{x} in the grid, we can select one of the two methods (Lax-Wendroff-Leveque or uBE) based on the maximum magnitude of ϑ for each \mathbf{x} .

We note that other higher order methods also exist to efficiently solve (III.37b) and (III.38b) with high advection speed, such as Crank-Nicolson, iWENO [28], and Lagrangian schemes [29]. Compared to uBE, Crank-Nicolson is highly oscillatory when the advection speed is high. These oscillations produce negative spikes in the probability density; which must be non-negative by definition. Similarly we tried the iWENO, or implicit WENO, method which is supposed to be stable for larger time step sizes, compared to the explicit WENO method. However, the method is again oscillatory unless the time step satisfies the stability requirements of the explicit WENO method. On the other hand, Lagrangian schemes are based on numerically solving the characteristic equations and interpolating the initial condition to approximate the solution at later time. These schemes seem the most promising in our case. However, they are only conservative if the used interpolation method is conservative. Constructing a conservative interpolation scheme that does not degrade the order of the method is not trivial, so we decided to leave it for future work. We refer to [29,30], where a conservative interpolation scheme is proposed, using the solution of separate PDEs for the spatial first-order derivatives of ρ .

One last detail is the evaluation of the advection speed ϑ which requires the evaluation of a double integral (refer to (II.35)). The inner integral in the velocity domain can be calculated once per time step and then used for all \mathbf{x} on the grid. On the other hand the second integral of the convolution must be re-evaluated for every \mathbf{x} on the grid. If naively implemented, this would introduce quadratic cost with respect to the number of cells in the \mathbf{x} -dimensions. We reduce this cost by using the Fast Fourier Transform (FFT) to evaluate the convolution efficiently for all \mathbf{x} -points on the grid. Observe that implementing FFT with non-uniform or non-rectangular mesh is nontrivial. This is the main reason for the choice of uniform rectangular grids in the discretization of the domain of the \mathbf{x} -dimensions.

Chapter IV

Numerical Results

This chapter presents numerical results obtained by approximating solutions to the modified particle-based model (II.28) and the continuous mean-field model (II.35) in both *CaseI* and *CaseII*. Recall that in *CaseI* pedestrians move clockwise on a circle with no obstacles. *CaseII*, on the other hand, models pedestrians exiting a room with obstacles. We start by describing the methods we use to test convergence of different models, the computed observables and the used parameters in Section IV.1. Section IV.2 presents results obtained using the modified particle-based model, Section IV.3 presents results obtained using the continuous model and Section IV.4 compares the results from the two models. Finally, IV.5 discusses the computational difficulties with *CaseII*.

IV.1 Observables and Parameters

To test convergence we compare a vector of observables computed using the different models or different discretizations of the same model. In *CaseI* these observables are

KDE-like (Kernel Density Estimation), in the sense that they are of the form

$$g_x(\tilde{x}_i) = \frac{1}{h_x} \int \Upsilon_1 \left(\frac{\nu(\tilde{x}_i, x) - 2r}{h_x} \right) \rho_x(x) \, dx, \quad (\text{IV.1a})$$

$$g_v(\tilde{v}_i) = \frac{1}{h_v} \int \Upsilon_1 \left(\frac{\tilde{v}_i - v}{h_v} \right) \rho_v(v) \, dv, \quad (\text{IV.1b})$$

for some $\{\tilde{x}_i\}_{i=1}^I$ and $\{\tilde{v}_i\}_{i=1}^I$ and positive constants h_x and h_v . Here, ρ_x and ρ_v are the marginal probability density functions of position and velocity at final time T , respectively. Also, ν is defined in (II.33a). The kernel Υ_1 is simply a Gaussian

$$\Upsilon_1(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2). \quad (\text{IV.2})$$

The vector of observables in *CaseI* is then

$$\mathbf{g} = (g_x(\tilde{x}_0), \dots, g_x(\tilde{x}_I), g_v(\tilde{v}_0), \dots, g_v(\tilde{v}_I)), \quad (\text{IV.3})$$

Similarly, in *CaseII* the observables are of the form

$$g_{\mathbf{x}}(\tilde{x}_i, \tilde{y}_j) = \frac{1}{h_x h_y} \int \int \Upsilon_2 \left(\frac{\tilde{x}_i - x}{h_x}, \frac{\tilde{y}_j - y}{h_y} \right) \rho_{J, \mathbf{x}}([x, y]) \, dx \, dy, \quad (\text{IV.4})$$

$$g_{\mathbf{v}}(\tilde{v}_i, \tilde{w}_j) = \frac{1}{h_v h_w} \int \int \Upsilon_2 \left(\frac{\tilde{v}_i - v}{h_v}, \frac{\tilde{w}_j - w}{h_w} \right) \rho_{J, \mathbf{v}}([v, w]) \, dv \, dw, \quad (\text{IV.5})$$

for some $\{\tilde{x}_i\}_{i=1}^I$, $\{\tilde{y}_i\}_{i=1}^I$, $\{\tilde{v}_i\}_{i=1}^I$ and $\{\tilde{w}_i\}_{i=1}^I$ and positive constants h_x , h_y , h_v and h_w .

The kernel Υ_2 is again a Gaussian

$$\Upsilon_2(x, y) = \frac{1}{2\pi} \exp(-(x^2 + y^2)/2). \quad (\text{IV.6})$$

Moreover

$$\rho_{J,\mathbf{x}}(\bar{\mathbf{x}}) = \frac{J(\bar{\mathbf{x}}) \int_{\mathcal{D}_v} \rho(T, \mathbf{v}, \bar{\mathbf{x}}) \, d\mathbf{v}}{\int_{\mathbb{R}^d} J(\mathbf{x}) \int_{\mathcal{D}_v} \rho(T, \mathbf{v}, \mathbf{x}) \, d\mathbf{v} \, d\mathbf{x}}, \quad (\text{IV.7})$$

$$\rho_{J,\mathbf{v}}(\bar{\mathbf{v}}) = \frac{\int_{\mathcal{D}_x} J(\mathbf{x}) \rho(T, \bar{\mathbf{v}}, \mathbf{x}) \, d\mathbf{x}}{\int_{\mathbb{R}^d} J(\mathbf{x}) \int_{\mathcal{D}_v} \rho(T, \mathbf{v}, \mathbf{x}) \, d\mathbf{v} \, d\mathbf{x}}. \quad (\text{IV.8})$$

The vector of observables in *CaseII* is then

$$\mathbf{g} = \left(g_{\mathbf{x}}(\tilde{x}_0, \tilde{y}_0), \dots, g_{\mathbf{x}}(\tilde{x}_0, \tilde{y}_I), g_{\mathbf{x}}(\tilde{x}_1, \tilde{y}_0), \dots, g_{\mathbf{x}}(\tilde{x}_i, \tilde{y}_j), \dots, g_{\mathbf{x}}(\tilde{x}_I, \tilde{y}_I), \right. \\ \left. g_{\mathbf{v}}(\tilde{v}_0, \tilde{w}_0), \dots, g_{\mathbf{v}}(\tilde{v}_0, \tilde{w}_I), g_{\mathbf{v}}(\tilde{v}_1, \tilde{w}_0), \dots, g_{\mathbf{v}}(\tilde{v}_i, \tilde{w}_j), \dots, g_{\mathbf{v}}(\tilde{v}_I, \tilde{w}_I), \right). \quad (\text{IV.9})$$

Given a reference vector of observables \mathbf{g}_{ref} and an approximated one $\hat{\mathbf{g}}$, errors are computed as

$$\text{Relative Error} = \frac{\|\hat{\mathbf{g}} - \mathbf{g}_{\text{ref}}\|_{\infty}}{\|\mathbf{g}_{\text{ref}}\|_{\infty}}. \quad (\text{IV.10})$$

Refer to Table IV.1 for the chosen parameters used in the simulations and computation of observables. Table IV.2 lists values of some observables in some cases for reference. Finally note that we only present convergence plots at a specific final time. However, similar plots were generated at different intermediate times showing similar results.

IV.2 Particle-Based Model

Figures IV.1 and IV.2 show sample realizations of the particle-based simulation in *CaseI* and *CaseII* using the parameters in Table IV.1. The marginal densities of pedestrians in the position and velocity plane were approximated using a KDE-like method and are shown in the same Figure. Note that this is different from the traditional KDE method since the samples of the particles' positions and velocities

Parameter Name	<i>CaseI</i>	<i>CaseII</i>
A for pedestrian-pedestrian interaction	10	40
A for pedestrian-obstacle interaction	-	2
B for pair interaction	0.016	0.0032
r Pedestrian radius	0.016	0.016
T	1.5	2.5
σ	0.0707	0.0707
x_{\max}	1	1
y_{\max}	-	1
s_{\max}	1	1
v_0	0.25	0.3
τ	1	1
ζ	10^{-3}	10^{-3}
$\{\tilde{x}_i\}_{i=1}^{10}$ for KDE-like observables	$(i-1)/10 + 0.05$	$(i-1)/10 + 0.05$
$\{\tilde{y}_i\}_{i=1}^{10}$ for KDE-like observables	-	$(i-1)/10 + 0.05$
$\{\tilde{v}_i\}_{i=1}^{10}$ for KDE-like observables	$2(i-1)/10 - 0.95$	$2(i-1)/10 - 0.95$
$\{\tilde{w}_i\}_{i=1}^{10}$ for KDE-like observables	-	$2(i-1)/10 - 0.95$
h_x, h_y	1/60	1/60
h_v, h_w	2/60	2/60

Table IV.1: Parameters used in the simulations for *CaseI* and *CaseII*. The parameters in *CaseII* are similar to those chosen by Helbing in his code.

are not independent as required by the traditional KDE algorithm [31].

The observables were approximated using a fixed number of pedestrians P using an approach like Monte Carlo. That is, in *CaseI*

$$g_x^P(\tilde{x}_i) = \frac{1}{h_x P} \sum_{\alpha=1}^P \Upsilon_1 \left(\frac{\nu(\tilde{x}_i, x_\alpha) - 2r}{h_x} \right), \quad (\text{IV.11a})$$

$$g_v^P(\tilde{v}_i) = \frac{1}{h_v P} \sum_{\alpha=1}^P \Upsilon_1 \left(\frac{\tilde{v}_i - v_\alpha}{h_v} \right), \quad (\text{IV.11b})$$

where x_α and v_α are the position and velocity of pedestrian α , respectively. Similarly in *CaseII*

$$g_{\mathbf{x}}^P(\tilde{x}_i, \tilde{y}_j) = \frac{1}{h_x h_y P_J} \sum_{\alpha=1}^{P_J} \Upsilon_2 \left(\frac{\tilde{x}_i - x_\alpha}{h_x}, \frac{\tilde{y}_j - y_\alpha}{h_y} \right), \quad (\text{IV.12a})$$

$$g_{\mathbf{v}}^P(\tilde{v}_i, \tilde{w}_j) = \frac{1}{h_v h_w P_J} \sum_{\alpha=1}^{P_J} \Upsilon_2 \left(\frac{\tilde{v}_i - v_\alpha}{h_v}, \frac{\tilde{w}_j - w_\alpha}{h_w} \right), \quad (\text{IV.12b})$$

Observable	Particle-based system, $P = 160$, relative tolerance $= 0.5\%$	Continuous mean-field limit, $N_x = 800$, $N_v = 400$, $K = 1500$, relative tolerance $= 0.5\%$
$g_x(\tilde{x}_3)$	1.05	1.04
$g_x(\tilde{x}_4)$	1.25	1.25
$g_x(\tilde{x}_5)$	1.35	1.35
$g_x(\tilde{x}_6)$	1.37	1.37
$g_x(\tilde{x}_7)$	1.29	1.29
$g_v(\tilde{v}_3)$	0	0
$g_v(\tilde{v}_4)$	0	0
$g_v(\tilde{v}_5)$	0.30	0.31
$g_v(\tilde{v}_6)$	1.72	1.69
$g_v(\tilde{v}_7)$	2.07	2.08

(a) *CaseI* at $T = 1.5$.

Observable	Particle-based system, $P = 160$, relative tolerance $= 10\%$	Continuous mean-field limit, $N_x = 40$, $N_v = 40$, $K = 800$, relative tolerance $= 100\%$
$g_x(\tilde{x}_3, \tilde{y}_4)$	4.9	4
$g_x(\tilde{x}_4, \tilde{y}_4)$	7.3	4
$g_x(\tilde{x}_4, \tilde{y}_3)$	6.0	4
$g_x(\tilde{x}_6, \tilde{y}_3)$	6.0	4
$g_x(\tilde{x}_7, \tilde{y}_3)$	4.5	4
$g_v(\tilde{v}_6, \tilde{w}_6)$	7.1	8
$g_v(\tilde{v}_6, \tilde{w}_5)$	2.5	2
$g_v(\tilde{v}_7, \tilde{w}_6)$	9.2	7
$g_v(\tilde{v}_7, \tilde{w}_5)$	3.2	2
$g_v(\tilde{v}_7, \tilde{w}_7)$	0.5	2

(b) *CaseII* at $T = 2.5$.

Table IV.2: Values of some observables at final time.

where $\mathbf{x}_\alpha = [x_\alpha, y_\alpha]$ and $\mathbf{v}_\alpha = [v_\alpha, w_\alpha]$ are the position and velocity of pedestrian α , respectively. Again, this is different from a traditional Monte Carlo method since $\{\mathbf{x}_\alpha\}_{\alpha=0}^P$ and $\{\mathbf{v}_\alpha\}_{\alpha=0}^P$ are not independent.

Figure IV.3 plots the relative error of the chosen vector of observables for different number of pedestrians P . The errors in this Figure were calculated by comparing each computed vector of observables to a reference vector of observables computed using a larger number of pedestrians P , indicated in the Figure's caption. We call this error the modelling error. In *CaseI* the relative error tolerance used to calculate the vector of observables is $\epsilon = 0.5\%$. In *CaseII* the relative error tolerance is $\epsilon = 10\%$. These figures suggest that the particle-based model reaches a limit as the number of pedestrians increases. Moreover, it seems that the observables converge as $\mathcal{O}(P^{-1})$ in *CaseI* and as $\mathcal{O}(P^{-0.5})$ in *CaseII*; where P is the number of particles.

IV.3 Continuous Model

Next we look at the numerical results obtained by approximating solutions to the continuous model (II.35) using the parameters in Table IV.1. Figures IV.4 and IV.5 show the L^1 convergence of approximate solutions to a reference solution with respect to discretization in position and velocity planes and time discretization, in *CaseI* and *CaseII*. In these figures, K refers to the number of time steps. Moreover, $N_x + 1$ refers to the number of discretization points in each dimension in the position plane. In other words, the number of cells in *CaseI* would be N_x and in *CaseII* the number of cells would be N_x^2 . Similarly, $N_v + 1$ refers to the number of discretization points in each dimension in the velocity plane. In each of these convergence plots, multiple approximations of the solution are compared to a reference approximation. Moreover, in each figure, the different approximations differ in only one of the discretization parameters (N_x, N_v and K) while keeping the others fixed. Similarly, the reference

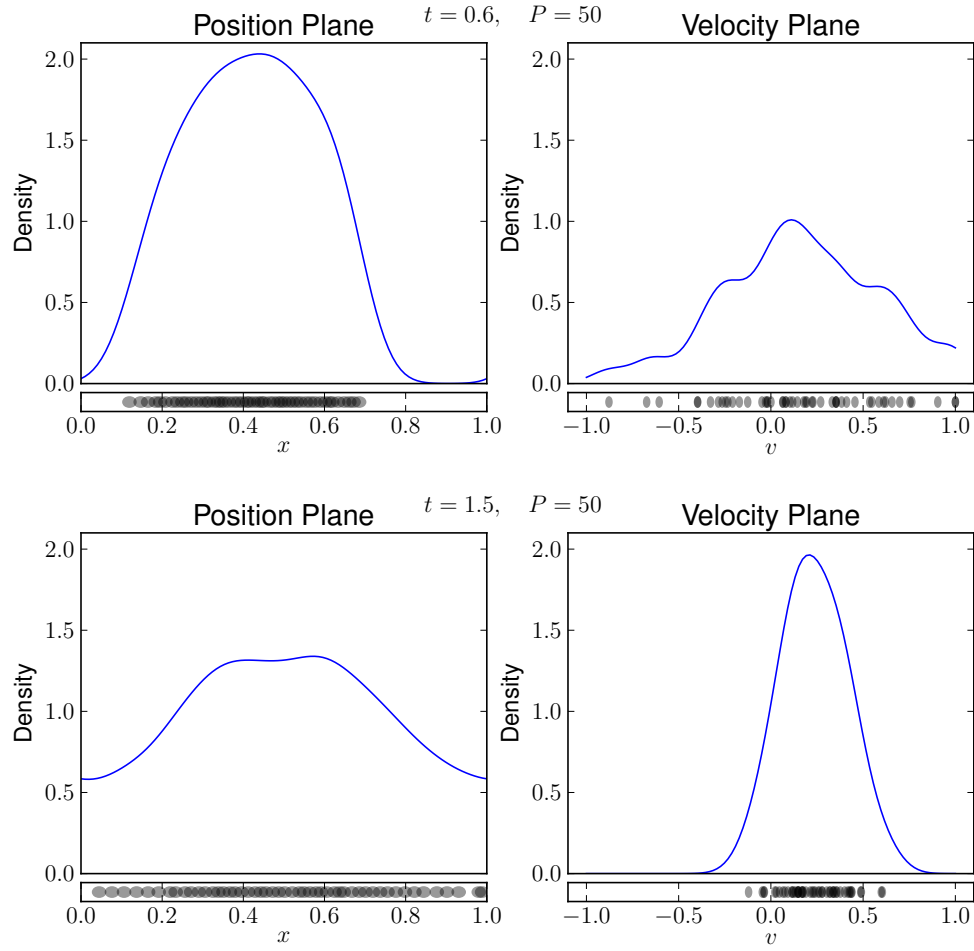


Figure IV.1: Sample realizations for *Case I* at two different times. The KDE of the marginal densities are superimposed on the positions and velocities of pedestrians.

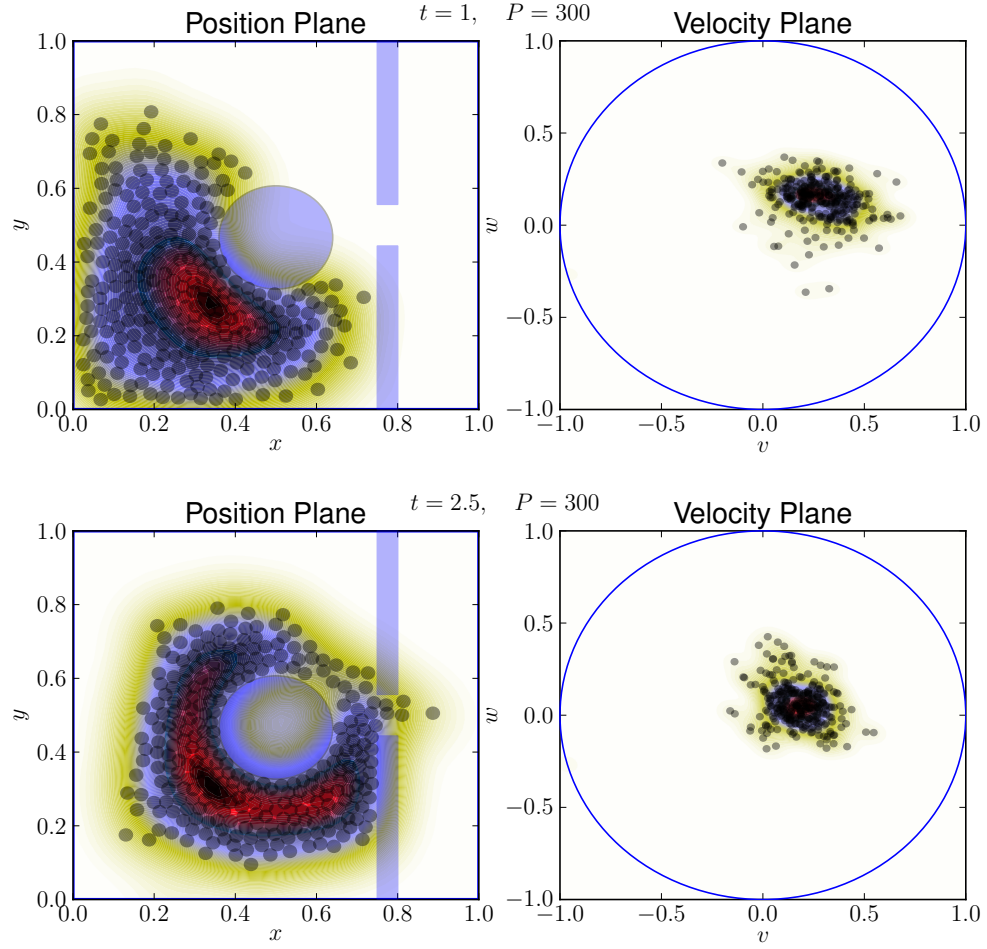
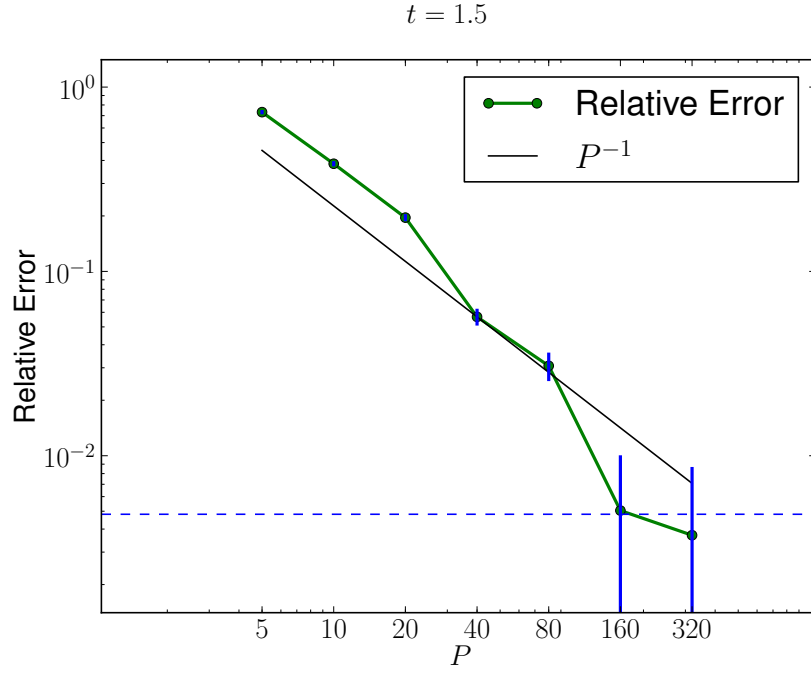
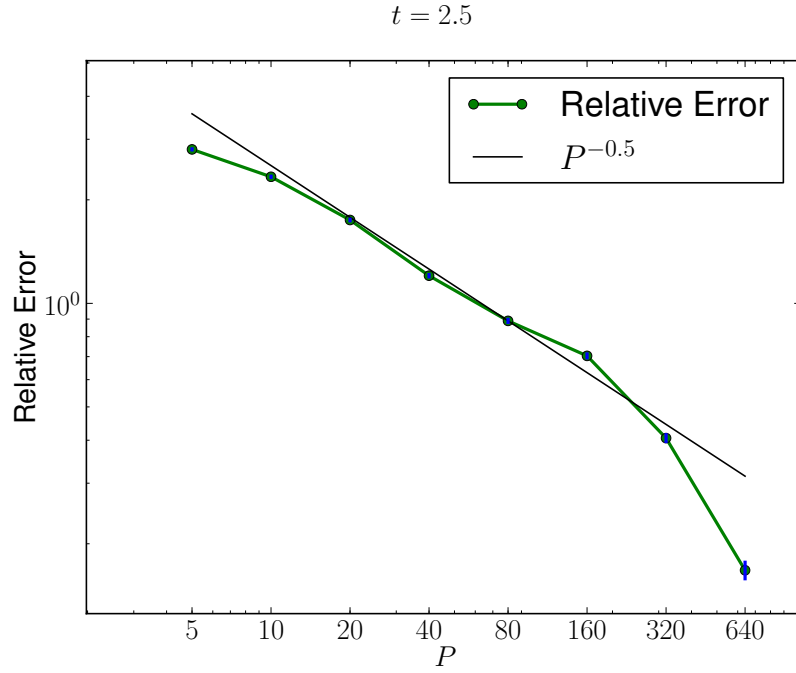


Figure IV.2: Sample realizations for *Case II* at two different times. The KDE of the marginal densities are superimposed on the positions and velocities of pedestrians.



(a) *Case I*. Reference solution was approximated using $P = 640$ pedestrians.



(b) *Case II*. Reference solution was approximated using $P = 1280$ pedestrians.

Figure IV.3: Convergence of observables computed using the particle-based model as the number of pedestrians increases. Blue dashed line is the relative error tolerance of the reference solution.

approximation has the same fixed discretization parameters and a fine discretization in the variable one. For example, in Figure IV.4a, N_x and N_v are the same for each approximation including the reference approximation and K is different for each approximation, here $K = 64000$ for the reference approximation. Finally, the error is the L^1 difference between each approximation and the reference solution.

In both *CaseI* and *CaseII*, we observe seemingly first order convergence with respect to the number of time steps. This is expected since the fractional time-stepping approach that we use is only first order accurate. We also observe at least first order convergence with respect to the number of cells in velocity plane. Similar convergence is observed in *CaseI* with respect to the number of cells in the position plane. The reason we do not see the expected second order convergence of the Lax-Wendroff-Leveque method is that we use the first order uBE method for some position slices depending on the advection speed. On the other hand, Figure IV.5b does not suggest convergence of the approximated solution with respect to the number of cells in the position plane in *CaseII*. We believe that this is because the PDE is under-resolved with such a coarse discretization of the position plane and convergence would be observed if finer discretization is used. However, due to the computational difficulties discussed in IV.5, we were not able to refine the discretization any further.

Figures IV.6 and IV.7 show the convergence of the previously described vector of observables when computed using approximate solutions to the continuous model. The error is the relative error between these vectors of observables to a reference vector of observables computed using a reference solution. Here, we observe similar convergence results to those previously observed in Figures IV.4 and IV.5. Moreover, we observe what seems to be more than first order convergence even with respect to the number of cells in the position plane in *CaseII*.

Finally, for *CaseI* we compute a reference solution with a tolerance of 0.5% to bound the relative error of the vector of observables by using Richardson extrapolation

to estimate the discretization error. We use this solution in the next section to compare the two models; continuous mean-field model and the particles-based model. For *CaseII*, obtaining such an approximation of the solution with small relative error tolerance proved to be difficult due to the computational complexity of the used numerical methods and the high dimensionality of the integral PDE, which in this case is four-dimensional and time-dependent. For this reason, the approximated solution that was obtained had an relative error tolerance of 100%.

IV.4 Comparison

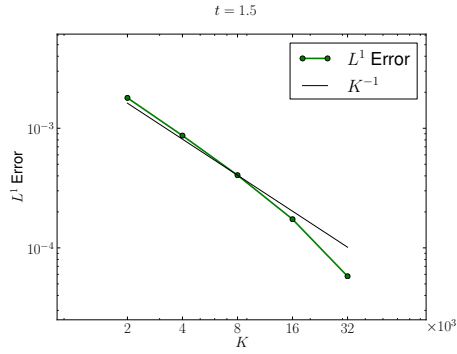
Finally we look at the convergence of the vector of observables that were calculated using the particle-based model to those calculated using the reference numerical solution of the continuous model. Figure IV.8 shows that this vector of observables seem to converge as P^{-1} in *CaseI* and as $P^{-0.5}$ in *CaseII*.

IV.5 Computational Challenges in *CaseII*

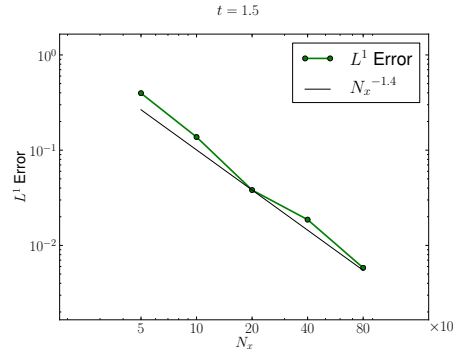
It can be argued that the observed errors in *CaseII* that are shown in Figures IV.7b and IV.8b are too large to make statements about convergence. Moreover, Figure IV.5b does not even suggest L^1 convergence of the approximated solution. However, even with large errors these figures show a possibility of convergence of the particle-based model to a mean-field model. These numerical observations justify further computations to show convergence with smaller errors and further investigation to rigorously prove such a convergence.

To generate figures similar to Figures IV.7b, IV.8b and IV.5b with smaller errors, finer discretization must be used to approximate the solution to the continuous model of the mean-field limit of *CaseII*. Recall that this continuous model is an integral, four-dimensional, time dependent PDE. The numerical methods discussed and

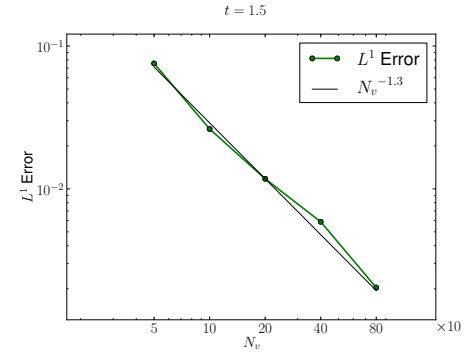
implemented in this work are non-adaptive and use uniform discretizations in all dimensions. Moreover, these methods are observed to be first-order as shown in IV.8b. As such the cost of these methods grow rapidly with finer discretization to the point where they become computationally infeasible. Better and higher order methods that use the regularity and locality of the solution (as seen in Figure IV.2) can be used to efficiently compute approximate solutions to the continuous model. To this end, we suggest using adaptive Finite Element methods and Sparse Grid methods.



(a) $N_x = 200, N_v = 200$ in all approximations and $K = 64000$ in the reference approximation.

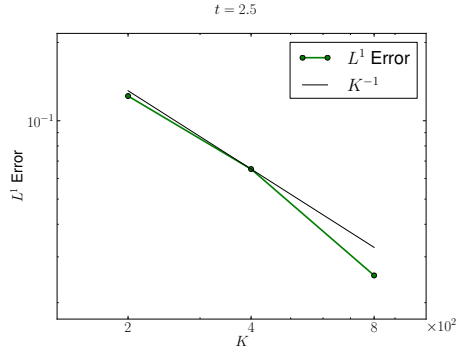


(b) $N_v = 140, K = 1500$ in all approximations and $N_x = 1600$ in the reference approximation.

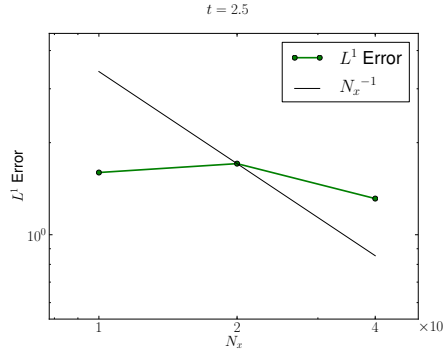


(c) $N_x = 200, K = 12000$ in all approximations and $N_v = 1600$ in the reference approximation.

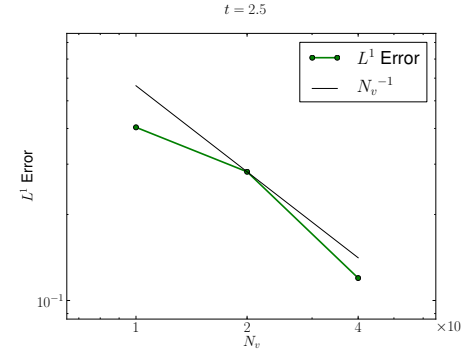
Figure IV.4: L^1 convergence of approximate solution to continuous model in *CaseI*.



(a) $N_x = 40, N_v = 40$ in all approximations and $K = 1600$ in the reference approximation.

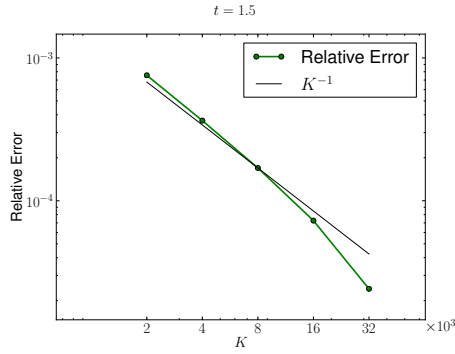


(b) $N_v = 40, K = 800$ in all approximations and $N_x = 80$ in the reference approximation.

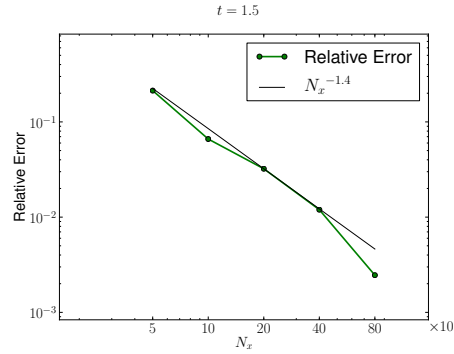


(c) $N_x = 40, K = 800$ in all approximations and $N_v = 80$ in the reference approximation.

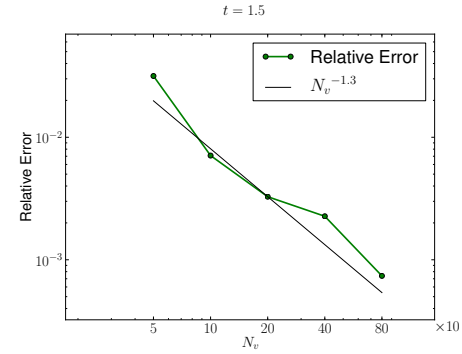
Figure IV.5: L^1 convergence of approximate solution to continuous model in *CaseII*.



(a) $N_x = 200, N_v = 200$ in all approximations and $K = 64000$ in the reference approximation.

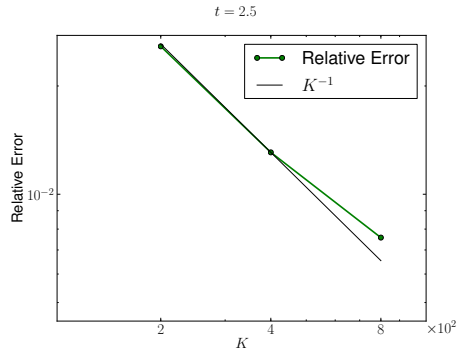


(b) $N_v = 140, K = 1500$ in all approximations and $N_x = 1600$ in the reference approximation.

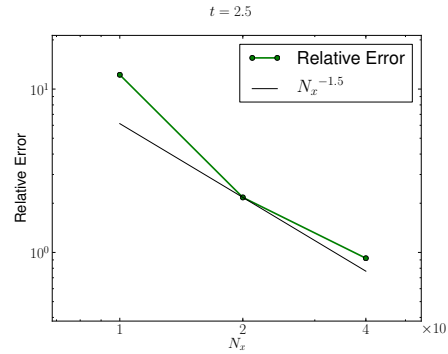


(c) $N_x = 200, K = 12000$ in all approximations and $N_v = 1600$ in the reference approximation.

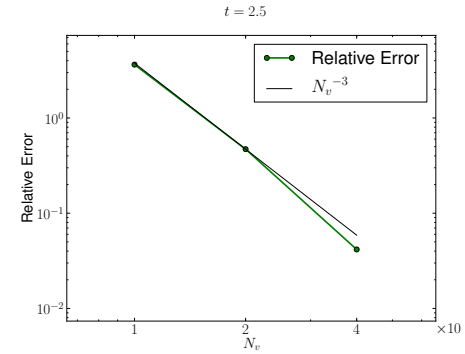
Figure IV.6: Convergence of vector of observables computed using the approximate solutions of the continuous model in *Case I*.



(a) $N_x = 40, N_v = 40$ in all approximations and $K = 1600$ in the reference approximation.

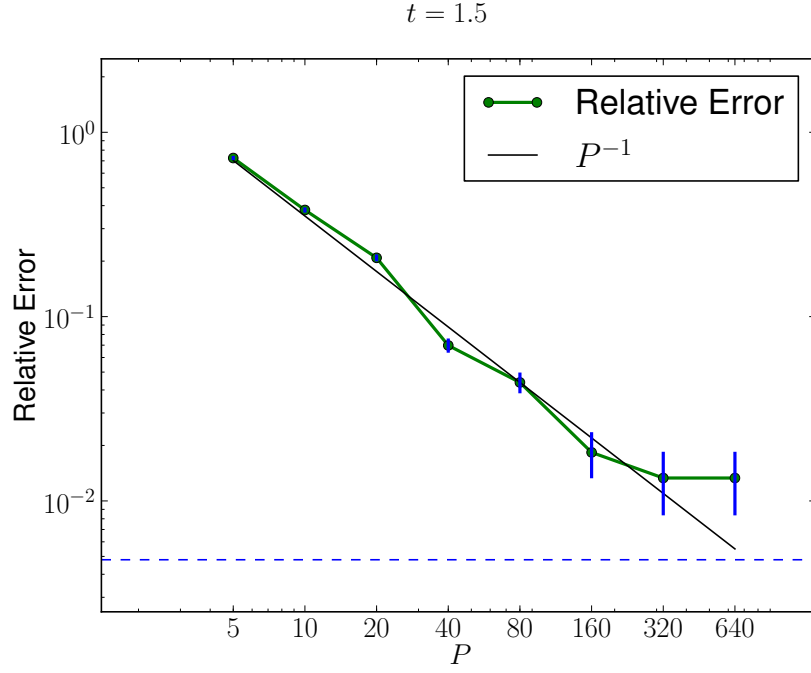


(b) $N_v = 40, K = 800$ in all approximations and $N_x = 80$ in the reference approximation.

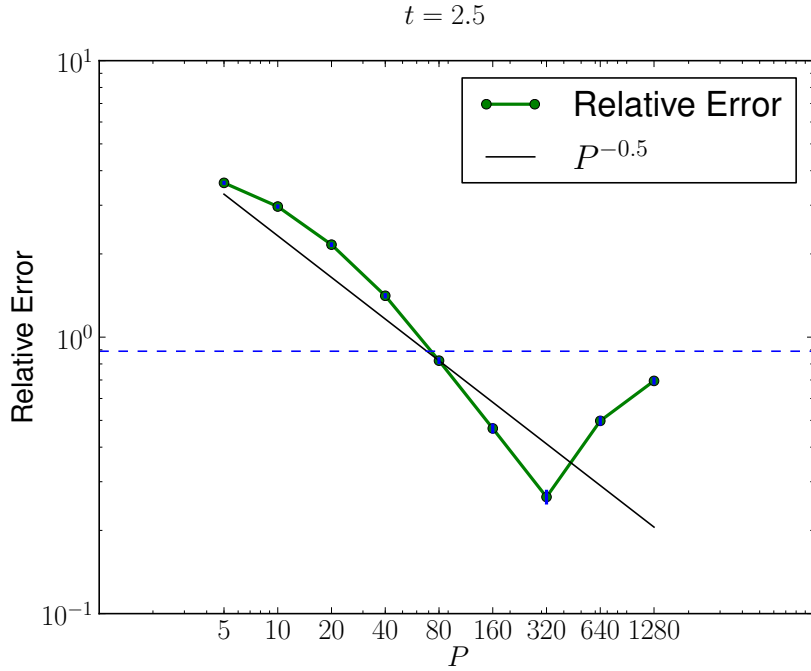


(c) $N_x = 40, K = 800$ in all approximations and $N_v = 80$ in the reference approximation.

Figure IV.7: Convergence of vector of observables computed using the approximate solutions of the continuous model in *Case II*.



(a) *Case I*. Reference approximation to the solution of the continuous mean-field was calculated using $N_x = 800, N_v = 400, K = 1500$.



(b) *Case II*. Reference approximation to the solution of the continuous mean-field was calculated using $N_x = 40, N_v = 40, K = 800$.

Figure IV.8: Convergence plot of observables computed using particle-based model to those computed using the continuous mean-field model. Blue dashed line is the relative error tolerance of the reference solution.

Chapter V

Particle Multilevel Monte Carlo

In this chapter we discuss a possible extension to MLMC that utilizes the strong convergence of the particle-based system to compute an observable in the mean-field. Note that this is the complete opposite of the objective discussed previously; namely, approximating observables of particle-based systems using the mean-field. Nonetheless, we list this application to illustrate the concept of Particle Multilevel Monte Carlo. This can also be an important method in case solving for the pdf of the mean-field limit is expensive or difficult, for example if the state in the nonlinear Fokker-Planck equation of the pdf is high dimensional.

Assume we are interested in computing $E\{Y\}$ where

$$Y = \psi(\bar{\mathbf{X}}(T), \bar{\mathbf{V}}(T)), \quad (\text{V.1})$$

for some Lipschitz function ψ . Here $\bar{\mathbf{X}}$ and $\bar{\mathbf{V}}$ satisfy (II.17). One way to compute such a quantity would be to use a Monte Carlo estimator. However, this assumes that we can sample $\bar{\mathbf{X}}$ and $\bar{\mathbf{V}}$, which is generally not the case since the marginal mean-field pdf ρ in (II.17) is generally unknown, unless we solve the corresponding Fokker-Planck equation (II.18). Nevertheless, we can sample an approximation of $\mathbf{X}(T)$ and $\mathbf{V}(T)$ by using P particles in the particle system and K time steps in the

Forward Euler method. Denote by $Y_{P,K;m}$ the m 'th sample of such an approximation, then the Monte Carlo estimator is

$$\hat{E}_{\text{MC}}\{Y\} = \sum_{m=1}^M Y_{P,K;m} \quad (\text{V.2})$$

We know from previous discussions in Chapters II and IV and [27] that the mean square error (MSE) is of the form

$$MSE = \mathbb{E} \left\{ |\mathbb{E}\{Y\} - \hat{E}_{\text{MC}}\{Y\}|^2 \right\} = \underbrace{\mathcal{O}(P^{-2w})}_{\text{Modelling error}} + \underbrace{\mathcal{O}(M^{-1})}_{\text{Statistical error}} + \underbrace{\mathcal{O}(K^{-2})}_{\text{Time discretization error}}, \quad (\text{V.3})$$

where w is the order of weak convergence of the particle-system to the mean-field limit as the number of pedestrians increases. For example, $w = 1$ for *Case I* pedestrian flow model, and $w = 0.5$ for *Case II* pedestrian flow model, both discussed in Chapter II. The total work to compute the estimator $\hat{E}_{\text{MC}}\{Y\}$ is $\mathcal{O}(MKP^r)$. For example, in the particle-based models discussed in this work $r = 2$. The cost is quadratic with respect to the number of particles due to the brute force calculation of the psychological force at each time step (cf. (II.28)). Hence, to bound the MSE by ϵ^2 the total needed work is $\mathcal{O}(\epsilon^{-3-r/w})$. Giles [27] proposed multilevel Monte Carlo in the context of SDEs to reduce the computational work required to bound the statistical and time discretization error from $\mathcal{O}(\epsilon^{-3})$ to $\mathcal{O}(\epsilon^{-2} \log^2(\epsilon))$. Using this method with a fixed number of particles reduces the total work to $\mathcal{O}(\epsilon^{-2-r/w} \log^2(\epsilon))$.

Inspired by Giles' work, we propose the following estimator

$$\hat{E}_{\text{ML}}\{Y\} = \sum_{l=0}^L Z_l, \quad (\text{V.4a})$$

$$Z_0 = \frac{1}{M_0} \sum_{m=1}^{M_0} Y_{P_0, K_0; m}, \quad (\text{V.4b})$$

$$Z_l = \frac{1}{M_l} \sum_{m=1}^{M_l} (Y_{P_l, K_l; m} - Y_{P_{l-1}, K_{l-1}; m}), \quad 1 \leq l \leq L. \quad (\text{V.4c})$$

The key idea here is that when computing Z_l both terms $Y_{P_l, K_l; m}$ and $Y_{P_{l-1}, K_{l-1}; m}$ use the same driving noise to reduce the total variance of the estimator $\hat{E}_{\text{ML}}\{Y\}$. Moreover, the samples of $\{Z_l\}_{l=0}^L$ are assumed to be independent. For example, if $P_l = 2P_{l-1}$ this can be done by sampling the initial condition of the P_l particles and the noise path $W(t)$ for each of these particle. With this, the sample $Y_{P_l, K_l; m}$ is computed. Then the P_l particles are split into two equal groups and $Y_{P_{l-1}, K_{l-1}; m}$ is computed for each of these groups where each particle uses the same initial condition and the same noise path $W(t)$. Finally the two values from the two groups are averaged.

To find the optimal number of samples M_l for each level, we minimize the work while constraining the variance to be $\epsilon^2/2$, that is

$$\begin{cases} \min_{M_l} \sum_{l=0}^L M_l K_l P_l^r \\ \text{subject to } \sum_{l=0}^L M_l^{-1} V_l \leq \frac{\epsilon^2}{2} \end{cases}, \quad (\text{V.5})$$

where $V_l = \text{Var}\{Z_l\}$. This optimization problem can be easily solved using Lagrange multipliers to find the optimal M_l , that is

$$M_l = 2\epsilon^{-2} \sqrt{\frac{V_l}{K_l P_l^r}} \left(\sum_{l=0}^L \sqrt{P_l^r V_l K_l} \right), \quad 0 \leq l \leq L. \quad (\text{V.6})$$

Also, for the value of L , observe that if the statistical error of Z_l for $l \geq 1$ is not

dominant then $|Z_l|$ is an approximation of the weak error of $|\mathbb{E}\{Y\} - \mathbb{E}\{Y_{P_{l-1}, K_{l-1}}\}|$, where $Y_{P_{l-1}, K_{l-1}}$ is the approximation of Y using P_{l-1} particles and K_{l-1} time steps. Based on this, we increase the number of levels L if $|Z_L|$ is below a certain tolerance, provided that the statistical error of Z_L is below the same tolerance. Recall from Section III.1.3 that the Central Limit Theorem provides an estimate of the statistical error based on the standard deviation of Z_L .

Next to calculate the total work, we assume that the variance of the differences satisfies $V_l = \mathcal{O}(K_l^{-1} + P_l^{-s})$ for some $s > 0$. We make the choice optimal $K_l = \mathcal{O}(P_l^s)$. Then substituting $V_l = \mathcal{O}(K_l^{-1})$ and (V.6) in (V.5), we can compute the order of the total work

$$\text{Total Work} = \sum_{l=0}^L M_l K_l P_l^r \quad (\text{V.7a})$$

$$= 2\epsilon^{-2} \left(\sum_{l=0}^L \sqrt{P_l^r V_l K_l} \right) \sum_{l=0}^L \left(\sqrt{\frac{V_l}{K_l P_l^r}} \right) K_l P_l^r \quad (\text{V.7b})$$

$$= \mathcal{O} \left(2\epsilon^{-2} \left(\sum_{l=0}^L P_l^{r/2} \right)^2 \right). \quad (\text{V.7c})$$

Provided that the statistical error of Z_L is bounded by $\mathcal{O}(\epsilon)$ we bound Z_L by $\mathcal{O}(\epsilon)$ to get

$$\mathcal{O}(P_L^{-w}) + \mathcal{O}(K_L^{-1}) = \mathcal{O}(P_L^{-w}) + \mathcal{O}(P_L^{-s}) = \mathcal{O}(P_L^{-f}) = \mathcal{O}(\epsilon). \quad (\text{V.8})$$

Here, we again used the assumption $K_L = \mathcal{O}(P_L^s)$ and denoted $f = \min(w, s)$. Also, assuming that we choose $P_l = \mathcal{O}(2^l)$ gives $L = \mathcal{O}(-\log(\epsilon)/f)$ and substituting in (V.7) gives a total work of $\mathcal{O}(\epsilon^{-2-r/f})$. Comparing this to the previous total work estimate of Giles' MLMC with fixed number of particles $\mathcal{O}(\epsilon^{-2-r/w} \log^2(\epsilon))$ we observe that, for the choices we made, if $w \leq s$ then the total work of the proposed MLMC method is $\mathcal{O}(\epsilon^{-2-r/w})$. In other words, the log factor is eliminated.

Figure V.1 shows the results of running standard Giles' MLMC with a fixed num-

ber of particles (denoted mlmc1), Particle MLMC with fixed number of time steps (denoted mlcm2) and finally the full MLMC with varying number of particles and varying number of time steps (denoted mlmc3) on *CaseI* particle-based model of pedestrian flow discussed in this work. In this system, $w = 1$ as observed numerically in Chapter IV. Also $r = 2$ due to the brute force calculation of the psychological force at each time step (cf. (II.28)). Moreover, the choice $K_l = \mathcal{O}(2^l)$, $P_l = \mathcal{O}(2^l)$ was made in all MLMC variations. In Figure V.1a, we observe the expected first order rates of weak convergence with respect to number time steps and number of particles, $Z_l = \mathcal{O}(K_l^{-1} + P_l^{-1})$. On the other hand, Figure V.1b suggests that $s \geq 1$ in *CaseI* particle-based model of pedestrian flow. Figure V.1c shows how the number of samples decrease in deeper levels. This agrees with the main idea of MLMC that was mentioned in Section III.1.4 where fewer samples with higher discretization requirements are used to reduce the discretization errors while more samples with less discretization requirements are used to reduce the statistical error. Moreover, recalling that the work per sample is $\mathcal{O}(KP^2)$ and that K is fixed in mlmc1 and P is fixed in mlmc2, Figure V.1d shows the expected computational time per sample in each case. Finally, Figure V.1f shows the expected work order $\mathcal{O}(\epsilon^{-4})$, which agrees with our prediction.

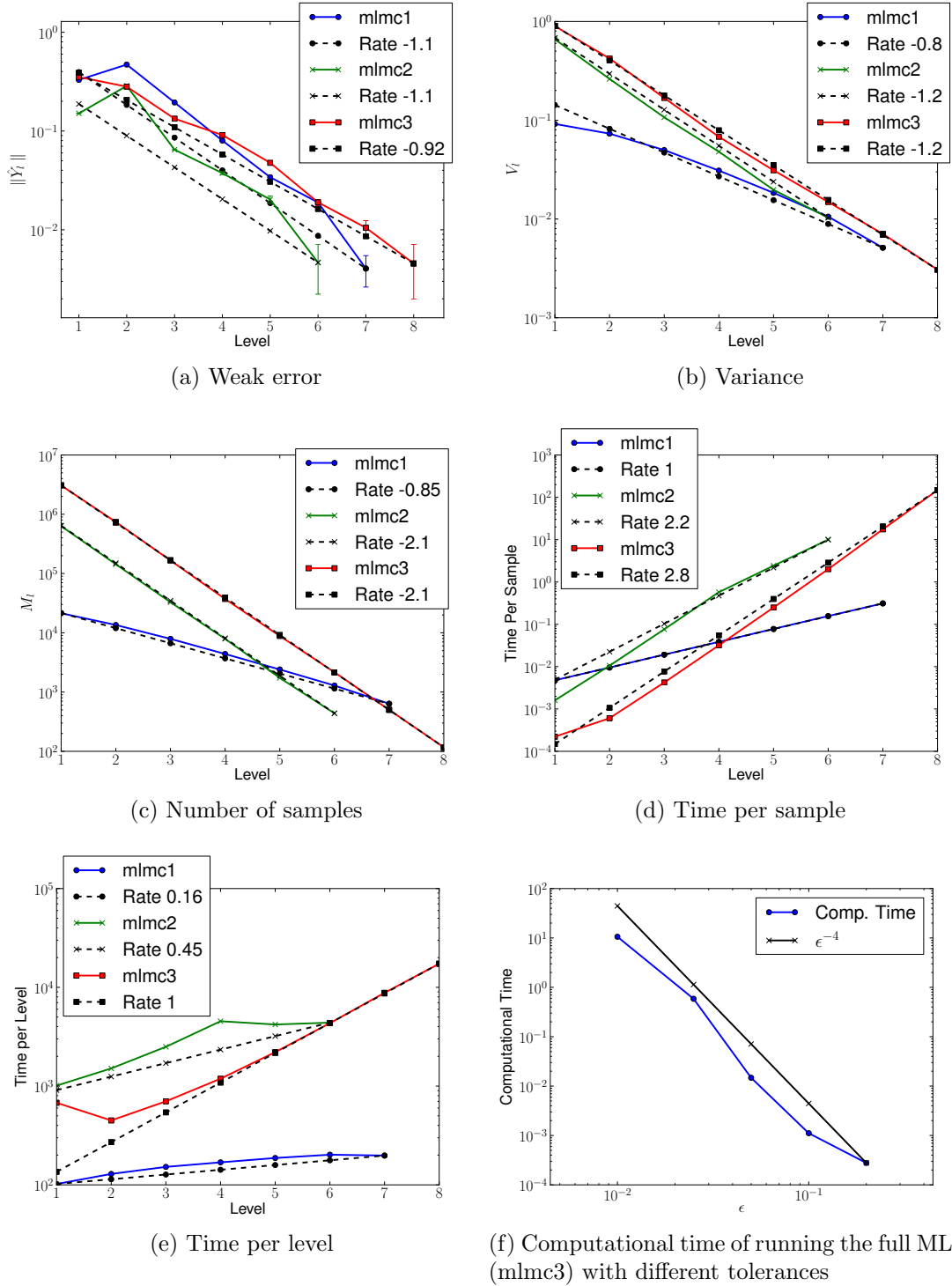


Figure V.1: Statistics about different MLMC methods when applied to the *Case I* particle-based model of pedestrian flow. Here, mlmc1 denotes standard MLMC with respect to time discretization and mlmc2 denotes particle MLMC with respect to number of particles. Finally, mlmc3 denotes full MLMC with respect to both.

Chapter VI

Conclusions

In this work we proposed modifications to Helbing's social force model for pedestrian flow [1] and proposed a mean-field approximation for it as the number of pedestrians increase. The existence of a mean-field limit of a specific SDE system was proved by Bolley et al. [12] under some restrictive conditions. The system we consider here is slightly different and does not satisfy all the required conditions. Yet we still conjecture a mean-field limit that we observe numerically. The evolution of mean-field limit is described by a marginal probability density function for each pedestrian that satisfies an integro-differential nonlinear Fokker-Planck equation. We discussed numerical methods to approximate a vector of observables using both the particle-based model and the continuous model and we presented numerical results that show the convergence of the particle-based model to a certain limit as the number of pedestrians increases. We also presented numerical results that indicate that such a limit is the solution of the aforementioned nonlinear Fokker-Planck equation.

As future work, and as discussed in IV.5, we plan to implement sparse approximation methods (cf. [32]) to efficiently solve the Fokker-Planck equation, especially in *Case II* where it is a integral, four dimensional, time-dependent PDE. Other methods can also be implemented to reduce the dimensionality of the problem, such as ap-

proximating the solution by a Gaussian and solving for its time-dependent mean and covariance matrix. Moreover, we plan to implement a multi-pole algorithm to use in the particle-based simulations to compute the psychological force in linear complexity with respect to the number of particles. Also, further parallelization of the particle-based simulations can be implemented that possibly uses the computational power of Graphical Processing Units (GPUs). Regarding Particle MLMC, we plan to look at possible usage of antithetic estimators [15] that could improve the rate of strong convergence and thus reduce the total work. Moreover, the current formulation assigns the same tolerance to both the modelling and time discretization errors. We plan to look at a different choice of tolerances that might increase the efficiency of the full MLMC method.

Further theoretical work might include extending the work of Bolley et al. [12] to prove the existence of a mean-field limit in more general settings that include the system we consider here. Also, optimization problems can be solved to determine the optimal environment parameters by using the deterministic solution of the continuous mean-field model instead of running expensive Monte Carlo simulations to the particle-based model. Finally, the current model can be easily extended to model other situations by changing the environment (including obstacles and exits) and choosing different target forces (cf. (II.24)). Moreover, more state variables might be added to the SDE system to track different objectives of the pedestrians. For example, we intend to model pilgrims circling the Kabaa in Mecca a specific number of times.

REFERENCES

- [1] D. Helbing, P. Molnár, I. Farkas, and K. Bolay, “Self-organizing pedestrian movement,” *Environment and Planning B: Planning and Design*, vol. 28, no. 3, pp. 361–383, 2001.
- [2] R. Leggett, “Real Time Crowd Simulation: A Review,” 2004. [Online]. Available: <http://www.leggett.net.org.uk/docs/crowdsimulation.pdf>
- [3] D. Helbing and P. Molnar, “Social force model for pedestrian dynamics,” *Physical Review E*, vol. 51, no. 5, pp. 4282–4286, 1995.
- [4] D. Helbing, I. Farkas, and T. Vicsek, “Simulating dynamical features of escape panic,” *Nature*, vol. 407, no. 6803, pp. 487–490, Sep. 2000.
- [5] E. Bouvier and E. Cohen, “Simulation of human flow with particle systems,” in *Simulators International XII. Proceedings of the 1995 Simulation MultiConference*. SCS, 1995, pp. 349–54.
- [6] C. W. Reynolds, “Flocks, herds and schools: A distributed behavioral model,” *ACM SIGGRAPH Computer Graphics*, vol. 21, no. 4, pp. 25–34, Aug. 1987.
- [7] V. Blue and J. Adler, “Cellular automata microsimulation for modeling bi-directional pedestrian walkways,” *Transportation Research Part B: Methodological*, vol. 35, no. 3, pp. 293–312, 2001.

- [8] J. Dijkstra, J. Jessurun, and H. Timmermans, “A multi-agent cellular automata model of pedestrian movement,” *Pedestrian and evacuation dynamics*, pp. 173–181, 2002.
- [9] R. Hughes, “The flow of human crowds,” *Annual review of fluid mechanics*, vol. 35, pp. 169–182, 2003.
- [10] L. Henderson, “The statistics of crowd fluids,” *Nature*, vol. 229, no. 5284, pp. 381–383, 1971.
- [11] A. Lachapelle and M. Wolfram, “On a mean field game approach modeling congestion and aversion in pedestrian crowds,” *Transportation Research Part B: Methodological*, vol. 45, no. 10, pp. 1572–1589, 2011.
- [12] F. Bolley, J. Canizo, and J. Carrillo, “Stochastic Mean-Field Limit: Non-Lipschitz Forces and Swarming,” *Mathematical Models and Methods in Applied Sciences*, vol. 21, no. 11, pp. 2179–2210, 2010.
- [13] F. Cucker and S. Smale, “Emergent Behavior in Flocks,” *IEEE Transactions on Automatic Control*, vol. 52, no. 5, pp. 852–862, May 2007.
- [14] A. Johansson, D. Helbing, H. Z Al-Abideen, and S. Al-Bosta, “From Crowd Dynamics to Crowd Safety: A Video-Based Analysis,” *Advances in Complex Systems*, vol. 11, no. 04, pp. 497–527, 2008.
- [15] J. Carlsson, K. Moon, A. Szepessy, R. Tempone, and G. Zouraris, *Stochastic Differential Equations: Models and Numerics*, 2010. [Online]. Available: www.math.kth.se/~szepessy/sdepde.pdf
- [16] I. Karatzas and S. Shreve, *Brownian motion and stochastic calculus*. Springer, 1991.

- [17] P. Cardaliaguet, “Notes on Mean Field Games,” 2010. [Online]. Available: <http://www.ceremade.dauphine.fr/~cardalia/MFG100629.pdf>
- [18] B. Pacchiarotti, C. Costantini, and F. Sartoretto, “Numerical approximation for functionals of reflecting diffusion processes,” *SIAM Journal on Applied Mathematics*, vol. 58, no. 1, pp. 73–102, 1998.
- [19] R. J. LeVeque, *Finite volume methods for hyperbolic problems*. Cambridge University Press, 2002, vol. 31.
- [20] —, *Finite difference methods for ordinary and partial differential equations*. Society for Industrial and Applied Mathematics, 2007.
- [21] S. Salsa, *Partial differential equations in action: from modelling to theory*. Springer, 2010.
- [22] D. I. Ketcheson, K. T. Mandli, A. Ahmadi, A. Alghamdi, M. Quezada, M. Parsani, M. G. Knepley, and M. Emmett, “PyClaw: Accessible, extensible, scalable tools for wave propagation problems,” *SIAM Journal on Scientific Computing*, vol. 34, no. 4, pp. C210–C231, 2012.
- [23] R. J. LeVeque, “Time-Split Methods for Partial Differential Equations.” PhD Thesis, Stanford University, 1982.
- [24] —, “Intermediate Boundary Conditions for Time-Split Methods Applied to Hyperbolic Partial Differential Equations,” *Mathematics of Computation*, vol. 47, no. 175, p. 37, Jul. 1986.
- [25] A. Szepessy, R. Tempone, and G. Zouraris, “Adaptive weak approximation of stochastic differential equations,” *Communications on Pure and Applied Mathematics*, vol. 54, no. 10, pp. 1169–1214, 2001.

- [26] M. Mascagni, D. Ceperley, and A. Srinivasan, “SPRNG: A scalable library for pseudorandom number generation,” in *Proceedings of the Ninth SIAM conference on parallel processing for scientific computing*, 1999.
- [27] M. B. Giles, “Multilevel Monte Carlo Path Simulation,” *Operations Research*, vol. 56, no. 3, pp. 607–617, May 2008.
- [28] S. Gottlieb, J. Mullen, and S. Ruuth, “A Fifth Order Flux Implicit WENO Method,” *Journal of Scientific Computing*, vol. 27, no. 1-3, pp. 271–287, Mar. 2006.
- [29] T. Yabe, R. Tanaka, T. Nakamura, and F. Xiao, “An Exactly Conservative Semi-Lagrangian Scheme (CIPCSL) in One Dimension,” *Monthly Weather Review*, vol. 129, no. 2, pp. 332–344, Feb. 2001.
- [30] T. Nakamura, R. Tanaka, T. Yabe, and K. Takizawa, “Exactly conservative semi-Lagrangian scheme for multi-dimensional hyperbolic equations with directional splitting technique,” *Journal of Computational Physics*, vol. 174, no. 1, pp. 171–207, 2001.
- [31] B. Silverman, “Density estimation for statistics and data analysis,” 1986.
- [32] H. Bungartz and M. Griebel, “Sparse grids,” *Acta Numerica*, vol. 13, pp. 147–269, Jun. 2004.