EFFICIENT ALGORITHMS FOR ASYMMETRIC FLOW FIELD FLOW FRACTIONATION

TIGRAN NAGAPETYAN

TECHNISCHE UNIVERSITÄT
KAISERSLAUTERN

Datum der Disputation: February 11, 2014

Vom Fachbereich Mathematik der Technischen Universität Kaiserslautern zur
Verleihung des akademischen Grades Doktor der Naturwissenschaften
(Doctor rerum naturalium, Dr. rer. nat.) genehmigte Dissertation

1. Gutachter:	Prof. Oleg Iliev
2. Gutachter:	PD. Dr. John G. M. Schoenmakers

D 386

Tigran Nagapetyan:
Efficient algorithms for Asymmetric Flow Field Flow Fractionation, © February 11, 2014

SUPERVISORS:
Prof. Oleg Iliev

REFEREES:
Prof. Oleg Iliev
PD. Dr. John G. M. Schoenmakers

LOCATION:
Kaiserslautern

I shall be telling this with a sigh
Somewhere ages and ages hence:
Two roads diverged in a wood, and I -
I took the one less traveled by,
And that has made all the difference.

— Robert Frost.

ACKNOWLEDGMENTS

First and foremost, I would like to express my immense gratitude to Prof. Oleg Iliev for direction of my research. Productive discussions with him and his thorough guidance and priceless support have most directly impacted my research. I am particularly thankful for coaching and mentoring me throughout the whole process and maintaining firendly relationships, as well as for giving me sufficient freedom to pursue my various projects in the scope of my research.

The project has directly benefited from the very productive collaboration with Prof. Klaus Ritter, who has kindly provided assistance in my work and whose scientific advice, knowledge and comments have had a profound role in the development of my research.

In addition, I have been very privileged to collaborate with the Wyatt Gmbh, and I am especially grateful to Dr. Christoph Johann for sharing useful information, data and tips, as well as for productive discussions regarding the experiments, as we strived to determine the nature of Asymmetric Flow Field Flow Fractionation (AF4) process.

I wholeheartedly acknowledge ITWM, in particular the Head of the Flow and Material Simulations Department Dr. Konrad Steiner, for the ITWM Fellowship; I have appreciated the local expertise and the excellently equipped facilities which allowed me to carry out my experiments. The work on this thesis was supported by Dr. Ralf Kirsch, Dr. Shiquan Zhang and many other great people from ITWM, which has become a source of scientific insight as well as good advice and cooperation.

Furthermore, I have greatly enjoyed the opportunity to work with Prof. Denis Belomestny, Prof. Mike Giles, Prof. Yvon Maday, Prof. Rene Pinnau. Our collaboration provided me with a unique opportunity to learn and work on various scientific subjects.

I have special people to thank, going back to my pre-ITWM days: I praise the enormous amount of help and teaching by professors of the Moscow State University, Faculty of Computational Mathematics and Cybernetics, and in particular Prof. Aram Arutyunov, Prof. Alexander Shananin and Dr. Alexander Gasnikov, whose scientific work taught me technical depth and attention to detail.

The path that has eventually laid my way to here is owed to my parents, Artur Nagapetyan and Astghik Petrosyan, and sister Lilit who have had key roles in my life and have been a source of energy and inspiration ever since. I am immensely grateful to them for every care and moral support provided to me.

Last but not least, I am grateful to my friends Nadir Bayramov, Alexander Belyaev, Dmitry Chistikov, Yegor Derevenets, Tatiana Gornak, Anastasia Migunova, Ivan Serna, Vladimir Shiryaev, Eduard Toroschin for interesting and challenging discussions.

CONTENTS

i ASYMMETRIC FLOW FIELD FLOW FRACTIONATION 1

1 INTRODUCTION 3

1.1 Focusing–injection stage. 3

1.2 Elution stage. 4

1.3 Overloading 5

1.4 Problems 5

ii OPTIMIZATION OF THE FOCUSING STAGE 7

2 PDE - BASED MATHEMATICAL MODEL 9

2.1 Coupled model 9

2.1.1 Motivation and Simplifications 9

2.2 The State System 10

2.2.1 Properties of Stokes equation 10

2.2.2 Existence, uniqueness and approximation of the convection-diffusion problem 12

3 THE OPTIMAL CONTROL PROBLEM 13

3.1 Functional 13

3.2 Problem statement 13

3.3 Selection of control space 14

3.4 Projection gradient method 15

4 NUMERICAL EXPERIMENTS 17

4.1 Time-space discretization of convection-diffusion equation 17

4.2 Scaling of coordinates in the numerical simulations 18

4.3 Numerical experiments 20

4.3.1 Discussion of the results. 25

iii MONTE CARLO METHODS FOR APPROXIMATION OF DISTRIBUTION FUNCTIONS 31

5 MONTE CARLO AND MULTILEVEL MONTE CARLO METHODS 33

5.1 Introduction 33

5.2 Single-level Monte Carlo estimator 33

5.3 Multilevel Monte Carlo estimator 34

5.4 Complexity estimators 34

6 MULTILEVEL MONTE CARLO METHOD FOR CDF AND PDF APPROXIMATION ON A COMPACT INTERVAL 37

6.1 Approximation of Distribution Functions on Compact Intervals 39

6.1.1 Smoothing 39

6.1.2 Assumptions on Weak and Strong Convergence 41

6.1.3 The Multi-level Algorithm 44

6.2 Approximation of Densities on Compact Intervals 50

6.2.1 Smoothing 50

6.2.2 Assumptions on Weak and Strong Convergence 51

6.2.3 The Multi-level Algorithm 51

6.3 Approximation of Distribution Functions at a Single Point 53

6.3.1 Smoothing 53

6.3.2 Assumptions on Weak and Strong Convergence 53

6.3.3 The Multi-level Algorithm 56

7 APPLICATIONS AND NUMERICAL EXPERIMENTS 59

7.1 Applications 59
7.1.1 Smooth Path-independent Functionals for SDEs 60
7.1.2 Smooth Path-dependent Functionals for SDEs 62
7.1.3 Stopped Exit Times for SDEs 63
7.2 Numerical Experiments 64
7.2.1 Smooth Path-independent Functionals for SDEs 65
7.2.2 Smooth Path-dependent Functionals for SDEs 66
7.2.3 Stopped Exit Times for SDEs 68
iv MODELLING ASYMMETRIC FLOW FIELD FLOW FRACTIONATION WITH STOCHASTIC DIFFERENTIAL EQUATIONS WITH REFLECTION 71
8 STOCHASTIC DIFFERENTIAL EQUATIONS WITH REFLECTION 73
8.0.4 Boundary 73
8.0.5 Existence and Uniqueness 75
8.1 PDE representation 76
8.2 Strong Approximation of Stopped Hitting Times 78
8.3 Mathematical model 80
9 NUMERICAL SIMULATIONS 81
9.1 Rectangular geometry 81
9.2 Hollow Fiber device 81
9.3 Eclipse device 85
v REDUCED BASIS APPROACH FOR RETENTION TIME ESTIMATION 89
10 PDE APPROACH FOR RETENTION TIME ESTIMATION 91
10.1 Mean exit time PDE 91
10.2 Reduced basis method 92
10.2.1 Two-grid Reduced basis 92
10.3 Numerical results 93
10.3.1 Numerical results for the offline stage 94
10.3.2 Numerical results for the online stage 95
10.3.3 Retention time 95
10.3.4 Future work on the optimization of the elution stage 98
BIBLIOGRAPHY 101

LIST OF FIGURES

Figure 1 Sketch of 3D channel of the Eclipse fractionation device. 3
Figure 2 Sketch of the focusing–injection stage of AF4. 4
Figure 3 Sketch of the Elution Stage of AF4. 5
Figure 4 Domain Ω of the system, its subdomains and boundary conditions 11
Figure 5 Concentration in Ω_1, integrated in y-direction $(B(c(x,T)))$. $K = 1$, set U_1 21
Figure 6 Concentration in Ω_1, integrated in y-direction $(B(c(x,T)))$. $K = 1$, set U_2 21
Figure 7 Concentration in Ω_1, integrated in y-direction $(B(c(x,T)))$. $K = 1.5$, set U_1 22
Figure 8 Concentration in Ω_1, integrated in y-direction $(B(c(x,T)))$. $K = 1.5$, set U_2 22
Figure 9 Values of $\log_{10} J$ with $N = 10$, $K = 1$ for two different grids. 23
Figure 10 Values of $\log_{10} J$ with $N = 10$, $K = 1.5$ for two different grids. 24
Figure 11 Values of $\log_{10} J$ with $N = 20$ and $N = 40$, $K = 1$. 24
Figure 12 Values of $\log_{10} J$ with $N = 20$ and $N = 40$, $K = 1.5$. 25
Figure 13 Terminal values of u for $N = 10$ for two different grids 26
Figure 14 Initial and terminal values of u for $N = 10$ 27
Figure 15 Initial and terminal values of u for $N = 20$ 28
Figure 16 Initial and terminal values of u for $N = 40$. 29
Figure 17 Smoothing polynomials g. 65
Figure 18 Replication numbers per level (left) and computational gain (right). 66
Figure 19 Error vs. accuracy demand ε. 67
Figure 20 Replication numbers per level (left) and computational gain (right). 67
Figure 21 Error vs. accuracy demand ε. 68
Figure 22 Replication numbers per level (left) and computational gain (right). 69
Figure 23 Error vs. accuracy demand ε. 69
Figure 24 Example of the domain, satisfying Conditions 1 and 2. 74
Figure 25 Domain D of the system and the streamlines of the velocity. 82
Figure 26 CDF for exit times 82
Figure 27 Hollow Fiber device during the focusing-injection stage for AFFFF. 82
Figure 28 Hollow Fiber device during the Elution Stage for AFFFF. 83
Figure 29 Particle injection, size $r = 3.2$ nm. 83
Figure 30 Particle injection, size $r = 32$ nm. 83
Figure 31 Fractogram for $r = 3.2$ nm and $r = 4.05$ nm, HF device 84
Figure 32 Sketch of the Eclipse device 85
Figure 33 Sketch of the Eclipse device 86
Figure 34 90 degree geometry. 4 seconds injection. 86
Figure 35 90 degree geometry. 120 seconds injection. 87
Figure 36 30 degree geometry. 4 seconds injection. 87
Figure 37 30 degree geometry. 120 seconds injection. 88
Figure 38 Domain and boundary types 91
Figure 39 Initial and coarsened grids. 94
Figure 40 Kolmogorov n-width decay. 95
Figure 41 Max L^2 error on B_{trial} in dependence of number of basis elements. 96
Figure 42 Parameters, chosen for the basis. 96

Figure 43 Mean Exit Time obtained via Reduced basis method and Retention time obtained via (117). 99

LIST OF TABLES

Table 1 Orders of convergence of the multi-level algorithm 39
Table 2 Multilevel Monte Carlo method for restoring the coefficient $D = 76.2$ micron^2/sec. Cross flow equal to 30 micron/sec. 84
Table 3 Parameters of simulation (absolute values). Retention times (sec) of particles with radius 3.2 and 4.05 nanometers after 180 seconds of focusing. 85
Table 4 Online stage test for coarse grid with 625×450 points 97
Table 5 Online stage test for coarse grid with 625×150 points 97
Table 6 Online stage test for coarse grid with 375×450 points 98
Table 7 Online stage test for coarse grid with 375×150 points 98
Table 8 Online stage test for coarse grid with 375×90 points 99

ACRONYMS

MLMC	Multi-level Monte Carlo
SMC	Single-level Monte Carlo
CDF	Cumulative Distribution Function
PDF	Probability Density Function
MSE	Mean Squared Error
SDE	Stochastic Differential Equation
FFF	Field Flow Fractionation
AF4	Asymmetric Flow Field Flow Fractionation
PDE	Partial Differential Equation
MET	Mean Exit Time

Part I

ASYMMETRIC FLOW FIELD FLOW FRACTIONATION

We introduce Asymmetric Flow Field Flow Fractionation technique (denoted further as AF4), which is a technique from Field Flow Fractionation (Field Flow Fractionation (FFF)) family. We state problems to be solved in this thesis and give an overview of the results.

1

INTRODUCTION

Asymmetric Flow Field Flow Fractionation is a special case of the Field Flow Fractionation technique. During the AF4 the particles get separated according to their size, more exactly to their hydrodynamic radius ([12]). This technique has already proved its efficiency in many applications in pharmaceutics, biology, chemistry (see e. g. [24], [18], [57], [20]).

The fractionation relies on the interplay between laminar flow and Brownian diffusion and occurs in a thin channel equipped with inlet and outlet, and a special membrane as a bottom wall (see e. g. sketch of Eclipse device on Figure1). A horizontal flow of a solvent along the membrane is combined with a strong cross flow across the membrane. The membrane has pores about 1nm in diameter, so it is impermeable for the particles while the solvent can flow throughout it. The frit is a ceramic or metallic composition situated under the membrane. It has a porous structure with pores about a micron in diameter so the solvent can easily go through it.

AF4 consists of two main stages: a focusing–injection stage and an elution stage.

1.1 FOCUSING–INJECTION STAGE.

A sketch of flow directions during the focusing–injection stage can be seen in Figure 27. During the focusing–injection stage the solvent is entering the channel from both sides and leaves through the membrane (see Figure 2).

The ratio between the left and the right volumetric fluxes determines the position of the focusing line. In the real devices this ratio is usually considered as piecewise-constant function of time. Moreover, the usual practice is to have number of switches as small as possible. The particles are injected from the left side during a certain time interval (shorter than the total duration of the focusing–injection stage), and they are transported towards the membrane due to the strong cross flow. The injection can be done either through a special injection port, or through the inlet. The Brownian diffusion, which acts isotropically, prevents particles from resting at the membrane surface. The interplay between the force

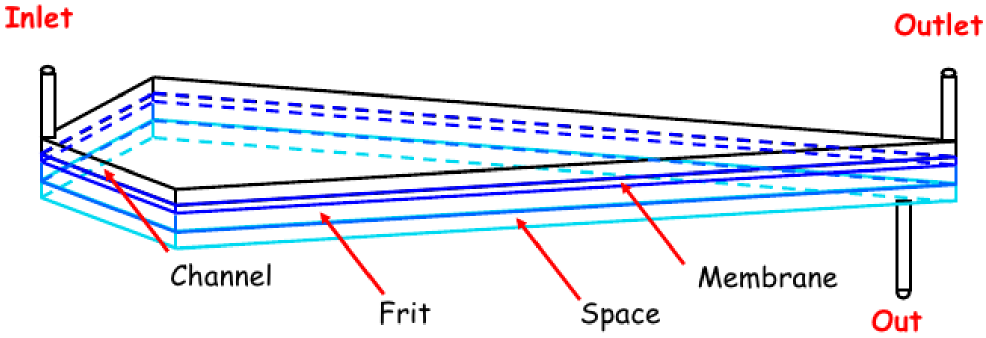


Figure 1: Sketch of 3D channel of the Eclipse fractionation device.

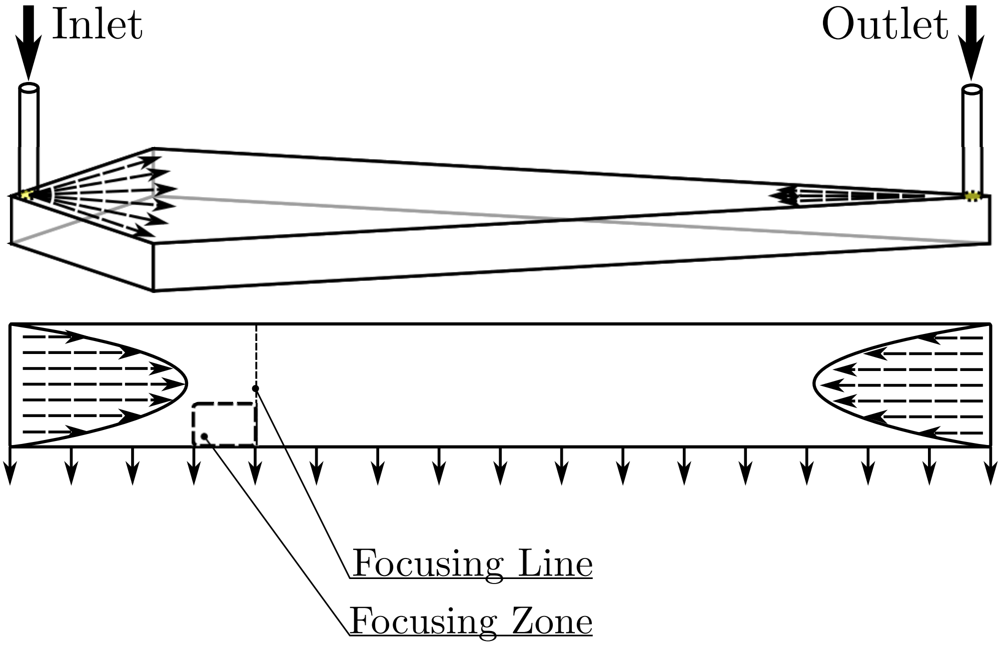


Figure 2: Sketch of the focusing–injection stage of AF4.

induced by the cross flow and the Brownian diffusion results in forming a boundary layer with average distance from the membrane depending on the particle size (i. e. on the diffusion coefficient) and on the intensity of the cross flow (see e. g. [12], [58]). Smaller particles with their larger diffusion coefficient form a layer which has a larger average distance to the membrane.

At the end of the focusing–injection stage the injected particles are located in a thin layer on the membrane, within a focusing zone around the focusing line, and this marks the starting point of the consecutive horizontal transport of the particles.

The main goals of the focusing-injection stage are:

A1. to concentrate almost all particles in a focusing zone by the end of the focusing stage;

A2. to avoid very high concentration of particles in a focusing zone; otherwise the particles will start to interact with each other, which can lead to a poor separation or non-durable use of the membrane (see section 1.3).

1.2 ELUTION STAGE.

The actual separation of the particles by their size is done during the elution stage. The parabolic profile of the horizontal flow (see Figure 3) is exploited to govern the separation: smaller particles are transported faster than the bigger particles towards the channel outlet because they experience a higher tangential flow velocity. An essential characteristic for description of the transport of particles with certain size and under given flow conditions is the `retention time`. The retention time is a characteristic time at which particles located at a prescribed point on the axis along the channel are expected to leave the channel. From mathematical perspective, the retention time is the median of exit times distribution and is one of the most important measures in the practical use of the fractionation (separation) devices along with the actual distribution of the exit times.

Distribution of exit times (sometimes also being called `chromatogram` or `fractogram`) and consequently the retention time, are produced as the results of the elution stage. In cases, when there is only one peak (or strictly speaking, when the distribution of exit times is unimodal), the retention

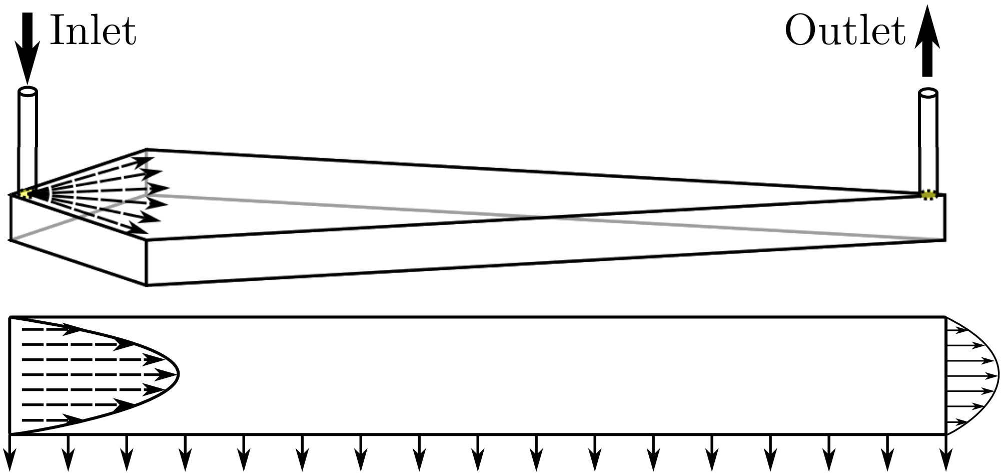


Figure 3: Sketch of the Elution Stage of AF4.

time can be seen as the mean exit time for particles to leave the channel, with respect to their initial position.

The main goal of the elution is to achieve the best possible separation of the particles, having in mind, that the whole elution stage should be as fast as possible. In order to achieve these goals, one needs to develop fast and reliable algorithms aimed to the problem of estimating the distribution function of exit times.

For a more detailed description of AF4 process see, e. g. [24],[58].

1.3 OVERLOADING

In many cases of FFF it is implicitly assumed that the considered suspensions are perfectly dilute. Hence one can assume that there is only one way coupling i. e. no back influence of particles on the flow, but the flow velocities enter in the equations governing particles transport. It is also assumed, that the particles do not interact with each other and do not change their hydrodynamic radius. In some cases, this assumption is not fulfilled which can result in a shift of the retention time and perturbation of chromatogram.

The phenomena of overloading was theoretically described and experimentally studied in [16]. Basically overloading means that there is a very high concentration of particles somewhere in the channel and consequently particles start to interact with each other (e. g. aggregation, additional repulsion forces). Assume, that the geometry of the separation device is fixed, then one can avoid overloading via controlling the crossflow and the injected amount of the sample in the separation channel. This is a challenging problem, and up to our knowledge, so far this problem has been approached only via experiments rather than via the numerical approach.

The problem of overloading is especially crucial for the focusing stage, as one need to transport all the particles into the focusing zone, so they experience the strong crossflow with almost no horizontal flow near the focusing line. The problem of overloading can be avoided, if one does not inject solvent at high speed, but on the other hand this can lead to the poor-focused concentration, which consequently will impact the fractograms.

1.4 PROBLEMS

The description of the AF4 brings up three distinct problems with both numerical and analytical challenges. We will state these problems, briefly discuss challenges and provide relevant solutions:

A. Optimal flow control for the focusing to avoid overloading (see Part ii)

  - one of the biggest challenges in this problem is the fact, that the governing convection-diffusion equation for this problem has a hugely (up to 10^6 times) dominating convection part;
  - the geometry of the device is highly stretched (up to 10^3 times);
  - the optimization algorithm should be fast and do not require high computational resources;
  - the hard bounds on the flow inflow can be considered, but then the focusing zone can be terribly large

B. Approximation of distribution function of exit times (see Part iii)

  - the shape of the distribution function of the exit times is of great importance in this application, so one needs algorithm, which can provide approximations for distribution function with error estimate in L^∞ norm
  - complexity estimates are important for determining the efficiency of the Monte Carlo algorithm

C. Stochastic Differential Equation (SDE)-based model for the particles behavior in the channel (see Part iv)

  - most of the SDE-based are developed for financial applications
  - model should take in the account, that the particles stay in the domain, with both reflecting and absorbing boundary
  - Multi-level Monte Carlo (MLMC) approach requires results on the strong convergence for exit times

D. Estimation of the retention time (see Part v)

  - huge dominance (up to 10^6 times) of the flow over the diffusion part;
  - stretched geometries;
  - the approximation algorithm should be fast and should not require high computational resources;

Part II

OPTIMIZATION OF THE FOCUSING STAGE

We introduce a Partial Differential Equation (PDE)-based mathematical model for the focusing stage modeling. We present a goal-oriented functional and algorithm for its minimization. Numerical experiments are described in the end of this Part with a further discussion of their results.

2

PDE - BASED MATHEMATICAL MODEL

The present chapter discusses an optimal control problem for the focusing-injection stage of the AF4 process, in which the distribution of particles suspended in a liquid needs to be controlled. In this problem, as in many of similar applications, the evolution of the particle concentration can be influenced only by the inflow/outflow of the liquid. From the mathematical perspective this leads to a constrained optimal boundary control problem for a convection-diffusion equation coupled with the Stokes equation. The theory to handle PDE constrained optimal problems is meanwhile well developed (see e. g. [36]). In particular, optimal control problems for convection-diffusion problems were considered in [13], for the Stokes and Navier-Stokes equations in [22, 60, 37, 33] and for the coupled problems in [3, 4, 45]. In these works the necessary derivative information was mainly provided by the solution of the adjoint equations.

Here, in contrast, we use the sensitivities approach due to the special structure of the objective functional for this specific application. This allows for the construction of a fast optimization algorithm, which is easy to parallelize. Parallelization is essential here, since the computing times for the forward problem are large already for the two-dimensional problem. Despite the fact, that we use sensitivities approach, we are able to provide explicit PDEs for each component of the gradient (see below).

2.1 COUPLED MODEL

2.1.1 *Motivation and Simplifications*

The solvent is usually an incompressible liquid, the flow is slow and therefore the Stokes equation can be used to describe the flow in the channel (see the sketch of the fractionation device geometry in Figure 1), as well as the flow in the free space under the membrane/frit region. The flow through the porous membrane and the underlying frit can either be described by the Darcy or the Brinkman model. Based on our previous experience, we consider only the Stokes-Brinkman model (for details see [38]). Further, the flow simulations with the 3D Stokes-Brinkman problem demonstrate that for the regimes that are relevant for the fractionation process, the velocity on the surface of the membrane is almost constant, with zero tangential component and non-zero vertical component. Based on this, even the Stokes-Brinkman equations can be replaced by the Stokes problem only in the channel. Numerical studies have also shown that even when we consider time-dependent Stokes equation with fixed boundary conditions, then its solution stabilizes in few seconds ([38]).

Remark 2.1.1 *Note, that this simplification is used also by pioneers of AF4 [58]. Furthermore, because the main objective of the research is the flow control for optimization of the fractionation procedure, another simplification from [58] is considered: We reduce the 3D problem to an effective 2D one. Clearly, that this simplifies our simulations, but it does not change the core nature of the optimization algorithm derived below.*

2.2 THE STATE SYSTEM

The state system is given by a convection-diffusion equation for the particle density, where the convective flow velocity is given by the solution of the quasi-stationary Stokes system:

$$\begin{cases} \Delta\vec{V} - \nabla p = 0, & x \in \Omega, \\ \nabla \cdot \vec{V} = 0, & x \in \Omega, \\ \vec{V} = \vec{g}(\cdot, u(t)), & x \in \partial\Omega. \end{cases} \tag{1}$$

Here V is the velocity, p the pressure and $\vec{g}(x, u(t))$ some prescribed spatial velocity distribution along the boundary (see Figure 4), which can be adjusted with respect to time via the control variable $u(t)$.

Then, the convection-diffusion equation for the concentration of particles $c(x, t)$ reads as:

$$\begin{cases} \partial_t c + \vec{V} \cdot \nabla c = D\Delta c, & (x,t) \in \Omega \times (0,T), \\ c(x,0) = c_0(x), & x \in \Omega, \\ \left(\vec{V}c - D\nabla c\right) \cdot \vec{n} = 0, & (x,t) \in \partial\Omega \times (0,T), \end{cases} \tag{2}$$

where the diffusion coefficient D is constant and the velocity $\vec{V}$ is the solution of (1).

Remark 2.2.1 *We assume, that the concentration is already injected of the domain. This assumption is restrictive, but can be met in the application (if one introduce additional injection tube into the geometry). Also, it allows us to consider simpler models for the modeling of AF4, which eventually leads to the faster algorithms.*

To state problem (2) in the weak form we introduce the spaces

$$W = L^2\left(0, T; H^1(\Omega)\right), \quad X = \{\phi \in W : \phi_t \in W^*\}$$

and the space-time cylinder $Q = \Omega \times (0, T)$, as well as $\Sigma = \partial\Omega \times (0, T)$.

Using the divergence theorem and the incompressibility of the vector flow $\vec{V}$ (i.e., $\nabla \cdot \vec{V} = 0$), for any $\varphi \in W$ we have

$$\begin{aligned} 0 &= \langle \tfrac{\partial}{\partial t} c, \varphi \rangle_{L^2(Q)} + \langle \nabla \cdot (\vec{V}c - D\nabla c), \varphi \rangle_{L^2(Q)} \\ &= \langle \tfrac{\partial}{\partial t} c, \varphi \rangle_{L^2(Q)} + \langle \varphi, (\vec{V}c - D\nabla c) \cdot \vec{n} \rangle_{L^2(\partial\Omega \times (0,T))} - \langle \vec{V}c - D\nabla c, \nabla\varphi \rangle_{L^2(Q)}. \end{aligned}$$

Applying the boundary conditions from (2), we obtain the following weak formulation:

Find $c \in X$ such that $c(x, 0) = c_0$ in $L^2(\Omega)$ and

$$\langle \partial_t c, \varphi \rangle_{W^*, W} - \left\langle \vec{V} \cdot c - D \cdot \nabla c, \nabla\varphi \right\rangle_{L^2(Q)} = 0 \tag{3}$$

for all $\varphi \in W$.

2.2.1 *Properties of Stokes equation*

First, we state a proposition for the existence and uniqueness of the Stokes system (1) (see [52]).

We define

$$L_0^2(\partial\Omega) = \left\{ \vec{f} \in L^2(\partial\Omega) : \int_{\partial\Omega} \vec{f} \cdot \vec{n} = 0 \right\}$$

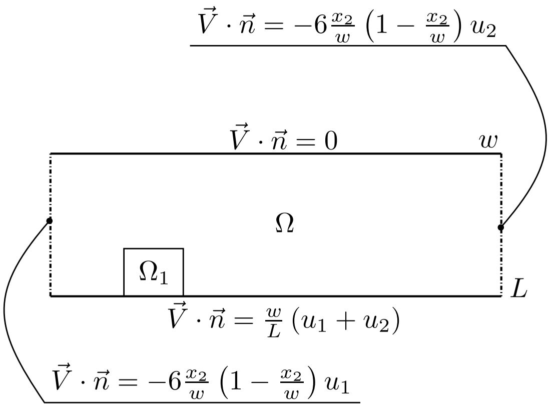

$\vec{V}\cdot\vec{n} = -6\frac{x_2}{w}\left(1-\frac{x_2}{w}\right)u_2$
$\vec{V}\cdot\vec{n} = 0$
w
Ω
Ω_1
L
$\vec{V}\cdot\vec{n} = \frac{w}{L}\left(u_1+u_2\right)$
$\vec{V}\cdot\vec{n} = -6\frac{x_2}{w}\left(1-\frac{x_2}{w}\right)u_1$

Figure 4: Domain Ω of the system, its subdomains and boundary conditions

for the boundary conditions (which have discontinuities at the corners of Ω in our numerical tests), and

$$L_0^2(\Omega) = \left\{ p \in L^2(\Omega) \colon \int\limits_{\Omega} p = 0 \right\}$$

for pressure functions.

By $H_\sigma^{1/2}(\Omega)$ we denote the subspace of $H^{1/2}(\Omega)$ of all weakly divergence free vector fields, that is, $\vec{f} \in H_\sigma^{1/2}(\Omega)$ if

$$\vec{f} \in H^{1/2}(\Omega) \quad \text{and} \quad \int_\Omega \vec{f}\cdot\nabla\phi = 0 \quad \text{for any } \phi \in C_0^\infty(\Omega).$$

Proposition 2.2.2 *Let Ω be a bounded Lipschitz domain in $\mathbb{R}^d$ and let $\vec{g} \in L_0^2(\partial\Omega)$. Then the system (1) admits a solution*

$$(\vec{V}, p) \in \left[H_\sigma^{1/2}(\Omega) \cap C^\infty(\Omega)\right] \times \left[L_0^2(\Omega) \cap C^\infty(\Omega)\right]$$

and there exists a constant $C_\Omega > 0$ such that

$$\left\|\vec{V}\right\|_{H^{1/2}(\Omega)} \leqslant C_\Omega \cdot \|\vec{g}\|_{L^2(\partial\Omega)}. \tag{4}$$

Further we have $\vec{g}$ as a linear function of $\binom{u_1}{u_2} \in U \subset \mathbb{R}^2$, where u_1 and u_2 are control parameters (see Figure 4).

In this case the Stokes equation admits a linear with respect to $\binom{u_1}{u_2}$ solution $(\vec{V}, p)$, that is,

$$\vec{V} = u_1 \cdot \vec{V}^{(1,0)} + u_2 \cdot \vec{V}^{(0,1)}, \quad p = u_1 \cdot p^{(1,0)} + u_2 \cdot p^{(0,1)}.$$

where $\vec{V}^{(1,0)}$ corresponds to boundary condition $\binom{u_1}{u_2} = \binom{1}{0}$ and $\vec{V}^{(0,1)}$ corresponds to boundary condition $\binom{u_1}{u_2} = \binom{0}{1}$. The uniqueness provided by the formulated above proposition yields the following result, which will be essential for the setup of our fast optimization algorithm.

Lemma 2.2.3 *Consider the Stokes system (1) with a boundary function $\vec{g} = u_1 \cdot \vec{g}^{(1,0)} + u_2 \cdot \vec{g}^{(0,1)}$, where $\vec{g}^{(1,0)}, \vec{g}^{(0,1)} \in L_0^2(\partial\Omega)$ and $u_1, u_2 \in \mathbb{R}$. Then the solution $(\vec{V}, p)$ depends linearly on $(u_1, u_2)^T$, i.e.*

$$\vec{V} = u_1 \cdot \vec{V}^{(1,0)} + u_2 \cdot \vec{V}^{(0,1)}, \quad p = u_1 \cdot p^{(1,0)} + u_2 \cdot p^{(0,1)},$$

where $(\vec{V}^{(1,0)}, p^{(1,0)})$ and $(\vec{V}^{(0,1)}, p^{(0,1)})$ correspond to the boundary conditions $\vec{g}^{(1,0)}$ and $\vec{g}^{(0,1)}$ respectively.

The lemma allows to compute the finite element approximations $(\mathbf{V}, \mathbf{p})$ of exact solution $(\vec{V}, p)$ as a linear combination of basis finite element approximations

$$(\mathbf{V}^{(1,0)}, \mathbf{p}^{(1,0)}) \text{ and } (\mathbf{V}^{(0,1)}, \mathbf{p}^{(0,1)}).$$

2.2.2 *Existence, uniqueness and approximation of the convection-diffusion problem*

The following theorem is a consequence of Theorem 11.1.1 in [50].

Theorem 2.2.4 *Let Ω be a bounded Lipschitz domain in $\mathbb{R}^n$. Further, let $\vec{V} \in L^\infty(Q)$ and $c_0 \in L^2(\Omega)$ be given. Then, there exists a unique solution $c \in X$ of problem (3).*

Remark 2.2.5 *Note that the bilinear form*

$$a(u, w) = \left\langle D \cdot \nabla u - \vec{V} \cdot u, \nabla w \right\rangle_{L^2(Q)}$$

is weakly coercive, i.e. there exist constants $\gamma_1, \gamma_2 > 0$ such that for all $u \in W$ it holds

$$a(u, u) \geqslant \gamma_1 \|u\|_W^2 - \gamma_2 \|u\|_{L^2(Q)}^2 .$$

For our numerical experiments we consider controls, which are piecewise-constant in time controls. Hence, we use the following backward Euler time discretization of the convection-diffusion equation (3):

$$\begin{gathered} \langle \hat{c}(t_k) - \hat{c}(t_{k-1}), \varphi \rangle_{L^2(\Omega)} - \\ \Delta t \cdot \left\langle \vec{V}(t_k) \cdot \hat{c}(t_k) - D \cdot \nabla \hat{c}(t_k), \nabla \varphi \right\rangle_{L^2(\Omega)} = 0 \\ \hat{c}(0) = c_0 \end{gathered} \tag{5}$$

for any $\varphi \in H^1(\Omega)$, where $\Delta t = \dfrac{T}{m}$, $t_k = k\Delta t$, $k = 1, \ldots, m$.

Remark 2.2.6 *Using the same argument as in Remark 2.2.5 one easily sees that the corresponding bilinear form is only weakly coercive in $H^1(\Omega)$ with respect to $L^2(\Omega)$. Coercivity can only be ensured for small time steps depending on $\left\|\vec{V}\right\|_{L^\infty(\Omega)}$. Hence, either we choose such a small time step or we assume that the Fredholm alternative holds.*

Further we will employ a stable edge-averaged finite element space discretization (see [9]) and obtain a system of differential equations for $c_h(t) \in V_h$:

$$\begin{gathered} \langle \tfrac{d}{dt} c_h(t), \varphi_h \rangle_{L^2(\Omega)} + b_h(c_h(t), \varphi_h) = 0, \; t \in (0, T) \\ \langle c_h(0), \varphi_h \rangle_{L^2(\Omega)} = \langle c_0, \varphi_h \rangle_{L^2(\Omega)} \end{gathered} \tag{6}$$

for any $\varphi_h \in V_h$. Here $V_h \subset H^1(\Omega)$ is a space of piecewise linear functions for a given mesh $\mathcal{T}_h$ with mesh size h.

Note that the backward Euler time discretization is unconditionally stable for this system of differential equations. Denote corresponding solution (after backward Euler time discretization) by $\hat{c}_h(t_k)$, $t_k = k\Delta t$. Then using the Rothe function $\hat{c}_{\Delta t}(x, t)$ defined by

$$\hat{c}_{h,\Delta t}(x, t)|_{[t_k, t_{k+1})} = \hat{c}_h(t_k),$$

where $k = 0, \ldots, m-1$, we can formulate the following convergence result:

Proposition 2.2.7 *For $\Delta t \to 0$, $h \to 0$ it holds*

$$\left\|\hat{c}_{h,\Delta t} - c\right\|_{L^2(Q)} \to 0.$$

3

THE OPTIMAL CONTROL PROBLEM

Now we state a specific optimal control problem on the domain $\Omega = (0, L) \times (0, w)$ shown in Figure 4.

3.1 FUNCTIONAL

We define the subdomain $\Omega_1 = (x_1^{(1)}, x_1^{(2)}) \times (0, \alpha w)$, where $0 < x_1^{(1)} < x_1^{(2)} < L$ and $0 < \alpha < 0.5$. The functional we are going to minimize is defined as

$$J(c,u) = \frac{1}{2}\int_{x_1^{(1)}}^{x_1^{(2)}} \left(\left[\int_0^{\alpha w} c(x_1,x_2,T)\, dx_2 - K \right]_+ \right)^2 dx_1 + \frac{\mu}{2}\left(\int_{\Omega\setminus\Omega_1} c(x,T) dx \right)^2, \tag{7}$$

where $c = c(u)$ is the concentration computed from (2) and $\vec{V} = \vec{V}(u)$.

It holds $(f)_+ = \max(f, 0)$ and $K > 0$ is a constant which gives the not to be exceeded average concentration on the sets $\{x_1\} \times [0, w]$.

The first term assures that the particles will end up in the focusing zone. Note that the focusing line will shift with every change of the control. The second term in (7) penalizes amount of concentration of the particles outside the focusing zone Ω_1, in which we aim to accumulate them. Both aims can be balanced by the positive constant μ.

3.2 PROBLEM STATEMENT

Consider the following linear operator

$$Bc = \int_0^{\alpha w} c(x_1,x_2)\, dx_2, \quad B\colon L^2(\Omega) \mapsto L^2(0,L) \tag{8}$$

Then, the functional (7) can be rewritten in a more convenient operator form

$$J(c,u) = \frac{1}{2}\langle (Bc-K)_+, (Bc-K)_+ \rangle_{L^2([x_1^{(1)},x_1^{(2)}])} + \frac{\mu}{2}\|c(T)\|^2_{L^1(\Omega\setminus\Omega_1)}. \tag{9}$$

Next, we formulate the constrained optimization problem. Define $Z = W \times L^2(\Omega)$ and the nonlinear operator $e = (e_1, e_2)\colon X \times \mathcal{U} \to Z^*$ by

$$\begin{cases} \langle e_1(c,u), \varphi \rangle_{W^*,W} = \langle \partial_t c, \varphi \rangle_{W^*,W} - \left\langle \vec{V}(u)c - D\nabla c, \nabla\varphi \right\rangle_{L^2(Q)} \\ \langle e_2(c,u), \varphi(x,0) \rangle_{L^2(\Omega)} = \langle c(x,0) - c_0(x), \varphi(x,0) \rangle_{L^2(\Omega)}, \end{cases} \tag{10}$$

for all $\varphi \in W$. Here $\vec{V}(u)$ is the (explicit) solution of the Stokes problem for a given boundary control $u = u(t)$.

Then the constrained minimization problem over the control space $\mathcal{U}$ can be shortly written as

$$\begin{aligned} &\min J(c,u) \text{ over } (c,u) \in X \times \mathcal{U}, \\ &\text{subject to } e(c,u) = 0. \end{aligned}$$

Let us introduce the solution operator $c(u)\colon \mathcal{U} \mapsto X$ (see [36]) and the reduced cost functional $\hat{J}(u) = J(c(u),u)$. Assume that the mapping $e_c(c,u)$ is a homeomorphism. Then, the implicit function theorem implies the existence of the derivative of the mapping $u \mapsto c(u)$ with respect to u in a direction δu which is given by

$$c'(u)[\delta u] = -e_c^{-1}(c(u),u)e_u(c(u),u)\delta u.$$

Noting that $J_u = 0$ one obtains by the chain rule

$$\langle \hat{J}'(u), \delta u\rangle_{\mathcal{U}^*,\mathcal{U}} = \langle J_c(c(u),u), -e_c^{-1}(c(u),u)e_u(c(u),u)\delta u\rangle_{X^*,X}.$$

Due to the special structure of our constraint and the upcoming choice of the control space, we are going to use this sensitivity formulation instead of the adjoint one. Note, that due to the linearity of the convection-diffusion equation the computation of $c'(u)[\delta u]$ requires the solution of the same equation just with a different right hand side and homogeneous initial condition.

Let $f_{u,\delta u} := c'(u)[\delta u]$. According to Lemma 2.2.3 the velocity $\vec{V} = \vec{V}(u)$ depends linearly on u, i.e. $\vec{V}_u(u)[\delta u] = \vec{V}(\delta u)$. Hence, $f_{u,\delta u} \in X$ can be computed from

$$\begin{cases} \langle \partial_t f_{u,\delta u}, \varphi\rangle_{W^*,W} - \left\langle \vec{V}(u) f_{u,\delta u} - D\nabla f_{u,\delta u}, \nabla\varphi \right\rangle_{L^2(Q)} = \langle \vec{V}(\delta u)c, \nabla\varphi\rangle_{L^2(Q)} \\ \langle f_{u,\delta u}(0), \varphi(\cdot,0)\rangle_{L^2(\Omega)} = 0, \quad \forall \varphi \in W \end{cases} . \tag{11}$$

Remark 3.2.1 *Due to the special structure of our constraints, the cost functional and the upcoming choice of the control space, we are going to use this sensitivity formulation instead of the adjoint one.*

Since the left-hand side of the convection-diffusion equation is linear with respect to the control function, the computation of $c'(u)[\delta u]$ *requires the solution of the state equation of the same form, but with different right-hand sides and with homogeneous initial condition. Therefore, one can solve each system for the component of the gradient independently and in parallel with the solution of the main state equation.*

Another advantage of this approach is that there is no need to save previous values of the state $c(x,t)$ *to compute the gradient, which already leads to great savings in memory in two-dimensional case.*

3.3 SELECTION OF CONTROL SPACE

By problem definition, functions $u_1(t)$, $u_2(t)$ that rule Stokes equation (1) are control parameters at any time t. We impose natural box constraints on the control:

$$u_1^a \leqslant u_1 \leqslant u_1^b, \quad u_2^a \leqslant u_2 \leqslant u_2^b.$$

Thus we obtain a parametric set U:

$$U = \left\{(u_1,u_2) : u_1 \in [u_1^a, u_1^b],\, u_2 \in [u_2^a, u_2^b]\right\},$$

which is a convex polygon.

Based on the Proposition 2.2.4, we will choose control functions $(u_1(t), u_2(t))$ to be piecewise continuous, i.e. they have a finite number of discontinuities of first kind on the interval $[0,T]$.

For the piecewise constant function we will need to evaluate derivatives with respect to an increment of arguments $u_1^i = u_1(\tau_{i-1})$, as well as of $u_2^i = u_2(\tau_{i-1})$ $(i = 1,\dots,N)$.

For example, designating $\dfrac{\partial}{\partial u_1^i} c(x,t) = f_1^i(x,t)$ and noting that

$$\frac{\partial}{\partial u_1^i} \vec{V}(\cdot, u(t)) = \chi_i(t)\vec{V}^{(1,0)},$$

where $\chi_i(t) = \mathcal{I}_{\{t\in[\tau_{i-1},\tau_i)\}}$, we have

$$\begin{cases} \langle \frac{\partial}{\partial t} f_1^i, \varphi \rangle_{W^*,W} - \left\langle \vec{V} f_1^i - D\nabla f_1^i, \nabla\varphi \right\rangle_{L^2(Q)} = \langle \vec{V}^{(1,0)} c, \nabla\varphi \rangle_{L^2(\Omega\times[\tau_{i-1},\tau_i])} \\ \langle f_1^i(0), \varphi(\cdot,0) \rangle_{L^2(\Omega)} = 0 \end{cases} .$$

In the numerical implementation we need to solve system (11) for various control variations δu to find the corresponding differential functional variation $\langle \hat{J}'(u), \delta u \rangle_{\mathcal{U}}$.

3.4 PROJECTION GRADIENT METHOD

As a result, after solving the systems for

$$c(x,t),\ f_1^i(x,t),\ f_2^i(x,t), \quad i = 1,\dots,N$$

we obtain a 2N-dimensional gradient vector of $c(x,t)$:

$$\begin{aligned} &\big(f_1^1(x,t),\dots,f_1^N(x,t), f_2^1(x,t),\dots,f_2^N(x,t)\big) \\ &\qquad =: d_u(x,t) \in [L^2(\Omega\times(0,T))]^{2N}. \end{aligned}$$

Then we derive an expression for the gradient of the reduced cost functional $\hat{J}(u)$: the i-th component of 2N-dimensional vector $\hat{J}'(u)$ can be written as

$$\begin{aligned} \big(\hat{J}'(u)\big)^i = &\quad \langle (Bc-K)_+, Bd_u^i \rangle_{L^2([x_1^{(1)},x_1^{(2)}])} \\ + &\quad \mu\,\|c(T)\|_{L^1(\Omega\setminus\Omega_1)} \int\limits_{\Omega\setminus\Omega_1} d_u^i(x,T)\,dx, \end{aligned}$$

where d_u^i is i-th component of the gradient vector $d_u(x,t)$.

Remark 3.4.1 *Similarly to the case with derivative of the solution* $c'(u)$, *gradient* $\hat{J}'(u)$ *is an approximation of Frechet derivative of functional* $\hat{J}(u)$ *with respect to control parameter* $u = u(t)$.

Using these results, we can formulate projection gradient (see [36]) method for minimization problem (9) with control set

$$\mathcal{U} = \left\{ (u_1^1,\dots,u_1^N,u_2^1,\dots,u_2^N) \colon\ (u_1^i,u_2^i) \in U \right\}.$$

This control set is convex and closed (which follows from the convexity of U), therefore, projection $\mathcal{P}_{\mathcal{U}}$ is well-defined and unique.

Denote initial guess $\mathbf{u}_0$ be in the interior of $\mathcal{U}$. Calculate the corresponding gradient $\mathbf{d}_0$. The algorithm has the following form.

```
1 : Set $k := 0,\ \sigma \in (0,1),\ \gamma_\varepsilon,\ \gamma$
2 : do
3 :     Evaluate $d_k := -\hat{J}'(u_k)$
4 :     Set $\mathbf{u}_{k+1} := \mathcal{P}_{\mathcal{U}}(\mathbf{u}_k + \gamma \mathbf{d}_k)$
5 :     Compute $J(\mathbf{u}_{k+1})$
6 :     If $\mathbf{u}_{k+1}$ satisfies Armijo's rule,
          set $\mathbf{k} := k+1$ and Go to 3
7 :     Else $\gamma := \gamma \cdot \sigma$ and Go to 4
8 : while $\|\mathbf{u}_{k+1} - \mathbf{u}_k\| / \|\mathbf{u}_1 - \mathbf{u}_0\| \geqslant \gamma_\varepsilon$
```

4

NUMERICAL EXPERIMENTS

4.1 TIME-SPACE DISCRETIZATION OF CONVECTION-DIFFUSION EQUATION

Recall that we consider a backward Euler time discretization of the convection-diffusion equation (3). For the resulting sequence of stationary convection-diffusion-reaction equations we apply an edge-averaged finite element (EAFE) scheme, following the articles [59], [9]. Note that this discretization we use for evaluating both the concentration $c(x,t)$ and derivatives $f_1^i(x,t)$, $f_2^i(x,t) \quad (i = 1, \dots, N)$.

We introduce the following notation thereby assuming that the triangulation $\mathcal{T}_h$ is shape regular: for a given $T_e \in \mathcal{T}_h$ denote by

- q_j $(1 \leqslant j \leqslant 3)$ the vertices of T_e;
- E_{ij} (or simply E) the edge connecting two vertices q_i and q_j;
- $\delta_E \phi = \phi(q_i) - \phi(q_j)$ for any continuous function ϕ on $E = E_{ij}$;
- $\tau_E = \delta_E x = q_i - q_j$ the directional vector of E.

We set $V_h \subset H^1(\Omega)$ to be the usual piecewise linear finite element space.

Define stiffness coefficients $(a_{ij}^{T_e})$ for any given $T_e \in \mathcal{T}_h$:

$$\forall c_h, v_h \in V_h : \int\limits_{T_e} D\nabla c_h \cdot \nabla v_h \, dx = \sum_{i,j} a_{ij}^{T_e} c_h(q_i) v_h(q_j)$$

and the coefficients $\omega_E^{T_e} = -a_{ij}^{T_e}$ with E connecting the vertices q_i and q_j.

Besides that, given $T_e \in \mathcal{T}_h$ and an edge $E \in T_e$, we define a function $\psi_E^{T_e}$ by

$$\frac{\partial \psi_E^{T_e}}{\partial \tau_E} = -\frac{1}{|\tau_E|}(D^{-1}\vec{V} \cdot \tau_E).$$

Let $\mathcal{H}_E^{T_e}(\vec{V})$ be the corresponding harmonic average of $e^{-\psi_E^{T_e}}$:

$$\mathcal{H}_E^{T_e}(\vec{V}) = \left[\frac{1}{|\tau_E|} \int\limits_E e^{\psi_E^{T_e}} \, ds \right]^{-1}.$$

Then the approximating bilinear form can be defined as

$$a_h(c_h, v_h) = \Delta t \sum_{E \in \mathcal{T}_h} \omega_E \mathcal{H}_E(\vec{V}) \delta_E(e^{\psi_E} c_h) \delta_E v_h + (c_h, v_h)_\Omega.$$

Here $(\cdot,\cdot)_\Omega = \sum\limits_{T_e \in \mathcal{T}_h} (\cdot,\cdot)_{L^2(T_e)}$. Due to the continuity of $\vec{V}(x)$ in Ω we have $\mathcal{H}_E^{T_e} \equiv \mathcal{H}_E$.

For any time step t_k $(k = 1, \dots, m)$ we consider the following finite element scheme: Find $c_h(t_k)$ such that

$$a_h[t_k](c_h(t_k), v_h) = (c_h(t_{k-1}), v_h)_\Omega, \tag{12}$$

where $a_h[t_k](c_h(t_k), v_h)$ corresponds to the velocity flow field $\vec{V}(\cdot) = \vec{V}(\cdot, t_k)$.

Approximation error estimates are also discussed in [59], [9].

At last we discuss some details of numerical implementation of the described scheme. Denote $\bar{\psi}_E(s) = \psi(|\tau_E|s)$ for $s \in [0, 1]$. One can see that

$$\frac{\partial \bar{\psi}_E}{\partial s} = -D^{-1}(\vec{V} \cdot \tau_E).$$

Consider the part $\mathbf{B}$ of matrix $\mathbf{A}$ of the bilinear form $a_h(\cdot, \cdot)$ such that $\mathbf{A} = \Delta t \cdot \mathbf{B} + \mathbf{A}_0$ which corresponds to the representation

$$a_h(c_h, v_h) = \Delta t \cdot b_h(c_h, v_h) + (c_h, v_h)_\Omega.$$

Here $\mathbf{A}_0$ denotes a mass matrix.

Then the element of matrix $\mathbf{B}$ of the bilinear form $b_h(\cdot, \cdot)$ corresponding the the edge $E = (q_i, q_j)$ can be computed as follows

$$b_{ij} = b_h(\varphi_j, \varphi_i) = \omega_E \mathcal{H}_E(\vec{V}) \delta_E(e^{\psi_E} \varphi_j) \delta_E \varphi_i,$$

where $\varphi_i(x)$ is a nodal basis function for a vertice q_i. Denoting $\alpha_E = (D^{-1}\vec{V} \cdot \tau_E)$ and assuming that in our discretization it is constant on any edge E, we have $\bar{\psi}_E(s) = \alpha_E(1 - s)$ and

$$b_{ij} = -\omega_E \left[\int_0^1 e^{\bar{\psi}_E(s)} ds \right]^{-1} e^{\bar{\psi}_E(1)} = -\omega_E \frac{\alpha_E}{e^{\alpha_E} - 1}. \tag{13}$$

Similarly one obtains

$$b_{ji} = b_h(\varphi_i, \varphi_j) = -\omega_E \left[\int_0^1 e^{\bar{\psi}_E(s)} ds \right]^{-1} e^{\bar{\psi}_E(0)} = -\omega_E \frac{\alpha_E}{1 - e^{-\alpha_E}}. \tag{14}$$

Also we have

$$b_{ii} = - \sum_{E = (q_i, q_k) \in \mathcal{T}_h} b_{ki}, \tag{15}$$

since

$$\omega_E \mathcal{H}_E(\vec{V}) \delta_E(e^{\psi_E} \varphi_i) \delta_E \varphi_i = -b_{ki}, \quad \forall E = (q_i, q_k) \in \mathcal{T}_h.$$

One can see that the local parameter α_E is a linear function of (u_1, u_2).

Note that the expression for b_{ji} has the same form as for b_{ij} since

$$b_{ji} = -\omega_E \frac{-\alpha_E}{e^{-\alpha_E} - 1}, \quad E = (q_i, q_j).$$

4.2 SCALING OF COORDINATES IN THE NUMERICAL SIMULATIONS

Preliminary tests showed that we need to apply the scaling of coordinates to improve the accuracy of the numerical results.

Recall the convection-diffusion problem (in differential form):

$$\frac{\partial c}{\partial t} + \vec{V} \cdot \nabla c = D \triangle c.$$

Consider the following scaling of coordinates:

$$\hat{t} = 0.1 \cdot t, \quad \hat{x} = 100 \cdot x.$$

These new coordinates are used in the numerical simulations.

We come to a new system:

$$\frac{\partial c}{\partial \hat{t}} + \hat{\vec{V}} \cdot \hat{\nabla} c = \hat{D} \hat{\triangle} c.$$

Returning to old coordinates (t, x), we obtain

$$10 \frac{\partial c}{\partial t} + 10^{-2} \hat{\vec{V}} \cdot \nabla c = 10^{-4} \hat{D} \triangle c$$

$$\Rightarrow \frac{\partial c}{\partial t} + 10^{-3} \hat{\vec{V}} \cdot \nabla c = 10^{-5} \hat{D} \triangle c.$$

Recall that $\vec{V} = u_1 \vec{V}^{(1,0)} + u_2 \vec{V}^{(0,1)}$ and $\hat{\vec{V}} = \hat{u}_1 \hat{\vec{V}}^{(1,0)} + \hat{u}_2 \hat{\vec{V}}^{(0,1)}$. The solutions of the Stokes system satisfy the identities

$$\vec{V}^{(1,0)} = \hat{\vec{V}}^{(1,0)} \quad \text{and} \quad \vec{V}^{(0,1)} = \hat{\vec{V}}^{(0,1)}$$

(pointwise in the corresponding points x and $\hat{x}$).

As a result, we can relate the scaled parameters via their corresponding real values:

$$\hat{D} = D \cdot 10^5, \quad \hat{u} = u \cdot 10^3.$$

Note that the boundary condition $\left(\vec{V} c - D \nabla c\right) \cdot \vec{n} = 0$ transforms to the same form $\left(\hat{\vec{V}} c - \hat{D} \hat{\nabla} c\right) \cdot \vec{n} = 0$.

Next we consider the functional

$$J = \frac{1}{2} \int\limits_{x_1^{(1)}}^{x_1^{(2)}} \left(\left[\int_0^{\alpha w} c(x_1, x_2, T)\, dx_2 - K \right]_+ \right)^2 dx_1 + \frac{\mu}{2} \Big(\int\limits_{\Omega \setminus \Omega_1} c(x, T) dx \Big)^2.$$

For convenience denote the integrals I_1, I_2 such that $J = I_1 + \mu I_2$. Since the spacial domain is two-dimensional, we have

$$\hat{I}_1 = I_1 \cdot 10^6, \quad \hat{I}_2 = I_2 \cdot 10^8,$$

with $\hat{K} = K \cdot 10^2$.

To achieve a relation $\hat{J} = J \cdot 10^6$, we switch to a new parameter $\hat{\mu} = \mu \cdot 10^{-2}$. Then due to an expression for J' we also have $\hat{J}' = J' \cdot 10^6$.

Recall that the step in the gradient algorithm has the form

$$\mathbf{u}_{k+1} := \mathcal{P}_{\mathcal{U}}(\mathbf{u}_k + \gamma \mathbf{d}_k)$$

with $\mathbf{d}_k := -J'(\mathbf{u}_k)$. To satisfy $\hat{u} = u \cdot 10^3$ for any step of this algorithm, we scale the length of the step $\hat{\gamma} = \gamma \cdot 10^{-3}$.

The Armijo's rule can be written as

$$J(\mathbf{u}_{k+1}) - J(\mathbf{u}_k) \leqslant \epsilon \langle J'(\mathbf{u}_k), \mathbf{u}_{k+1} - \mathbf{u}_k \rangle.$$

This inequality holds also for the scaled equivalents if $\epsilon u = \hat{\epsilon} \hat{u}$. As a result, $\hat{\epsilon} = \epsilon \cdot 10^{-3}$.

In particular, the real parameters are the following:

$$L = 0.1,\ w = 0.00029,\ T = 100,\ \Delta t = 0.05,\ D = 8 \cdot 10^{-11},$$
$$\mathbf{u}_0 = (1.5 \cdot 10^{-3},\ 8.5 \cdot 10^{-3}),\ \mu = 10^4,\ \epsilon = 0.005,\ \gamma = 10^4.$$

They correspond to the values of parameters used in the code:

$$L = 10.0,\ w = 0.029,\ T = 10,\ \Delta t = 0.005,\ D = 8 \cdot 10^{-6},$$
$$\mathbf{u}_0 = (1.5,\ 8.5),\ \mu = 100,\ \epsilon = 0.000005,\ \gamma = 10.$$

4.3 NUMERICAL EXPERIMENTS

We present here some numerical results of minimization problem with simplified (compared with practical tasks) parameters of PDE and control.

For evaluating $c(x,t)$, $f_1^i(x,t)$, $f_2^i(x,t) \quad (i = 1,\dots,N)$ we applied edge-averaged finite element (EAFE) scheme, following the articles [59], [9].

At first we consider a certain case of boundary conditions for velocity flow $\vec{V}$ shown on Figure 4. The solution for the Stokes equation (1) will read as

$$V_1 = -6(u_1+u_2)\frac{x_1}{L}\cdot\frac{x_2}{w}\left(1-\frac{x_2}{w}\right)+6u_1\frac{x_2}{w}\left(1-\frac{x_2}{w}\right),$$

$$V_2 = -(u_1+u_2)\left(1-3\left(\frac{x_2}{w}\right)^2+2\left(\frac{x_2}{w}\right)^3\right)\cdot\frac{w}{L}$$

Rectangular domain Ω has length $L = 0.1$ m and width $w = 290\cdot 10^{-6}$ m. Subdomain Ω_1 is described as

$$\Omega_1 = \{(x_1,x_2):\ x_1 \in [0.1\cdot L,\ 0.2\cdot L],\ x_2 \in [0,\ 0.1\cdot w]\}.$$

We consider a diffusion coefficient $D = 8\cdot 10^{-11}\ \mathrm{m}^2/\mathrm{sec}$, while the focusing time is $T = 100$ seconds.

Number of switching moments was selected $N = T/\Delta t$ with fixed switching time steps $\Delta t_1 = 5$ sec and $\Delta t_2 = 10$ sec. Moreover, as an initial guess for optimization problem with $\Delta t_2 = 5$ we use the answer of the optimization problem with $\Delta t_1 = 10$ sec.

Two cases of box constraints on u_1 and u_2 are considered. Corresponding constraining sets we denote by U_1 and U_2. Then

$$U_1 = \left\{(v_1,v_2)\colon v_2 \in [4.5\cdot 10^{-4},\ 10^{-2}],\ \frac{v_1}{v_2} \in \left[\frac{1}{9},\frac{1}{4}\right]\right\}, \tag{16}$$

$$U_2 = \left\{(v_1,v_2)\colon v_2 \geqslant 4.5\cdot 10^{-4},\ \frac{v_1}{v_2} \in \left[0,\ \frac{1}{3}\right]\right\}. \tag{17}$$

Triangular mesh $\mathcal{T}_h$ of the domain Ω was selected uniform, with the numbers of partitions equal to

$$n_x = 3960, \quad n_y = 350$$

along x-axis and y-axis respectively. Due to the box constraints on control u and initial condition $c_0(x)$ accumulated on $[0, 0.1\cdot L]\times[0,w]$, only the part of the mesh $\mathcal{T}_h$ on subdomain

$$\Omega_\theta = [0,\theta\cdot L]\times[0,w], \quad \theta = 0.3,$$

was used for computations of $c(x,t)$ (since its values on the rest of the mesh are physically negligible). Parameter μ in the functional (7) is set to be equal 10^4.

We have done tests for two different K:

- $K = \dfrac{1}{0.1\cdot L}\int\limits_{\Omega} c_0(x)\,dx$ — stands for ideally flat distribution. The value of (8) is presented on Figures 5-6.
- $K = \dfrac{1.5}{0.1\cdot L}\int\limits_{\Omega} c_0(x)\,dx$ — more weak constraint. The value of (8) is presented on Figures 7-8.

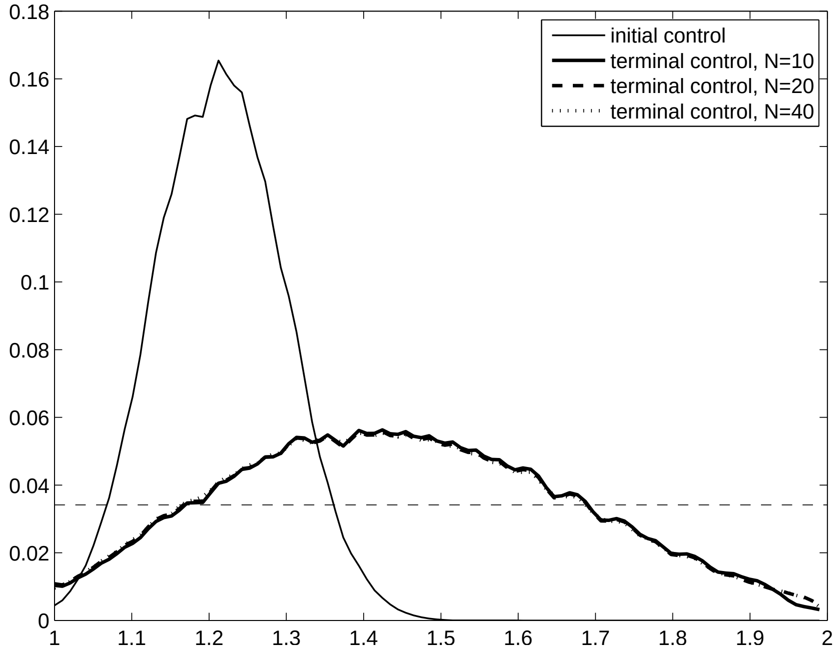


Figure 5: Concentration in Ω_1, integrated in y-direction ($B(c(x,T))$). $K = 1$, set U_1

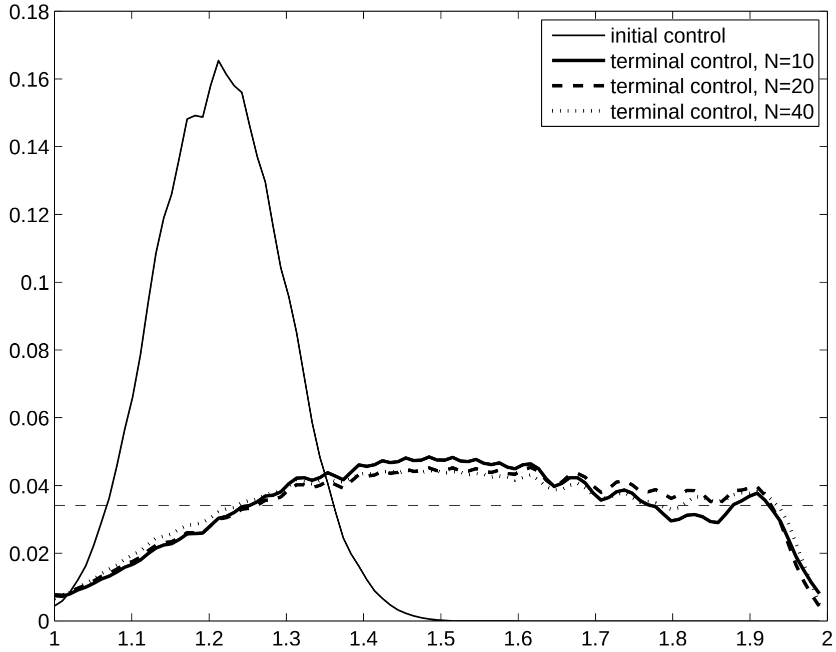


Figure 6: Concentration in Ω_1, integrated in y-direction ($B(c(x,T))$). $K = 1$, set U_2

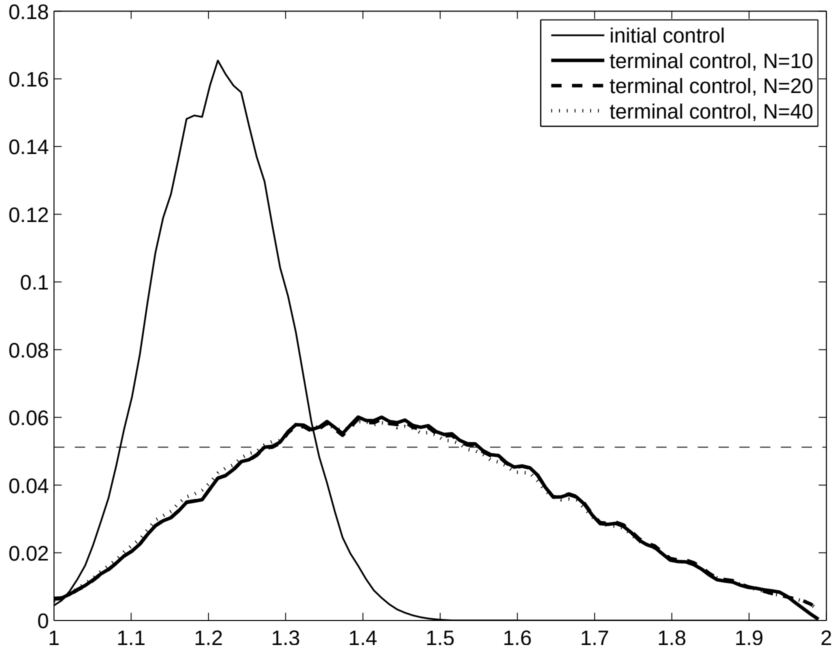


Figure 7: Concentration in Ω_1, integrated in y-direction ($B(c(x, T))$). $K = 1.5$, set U_1

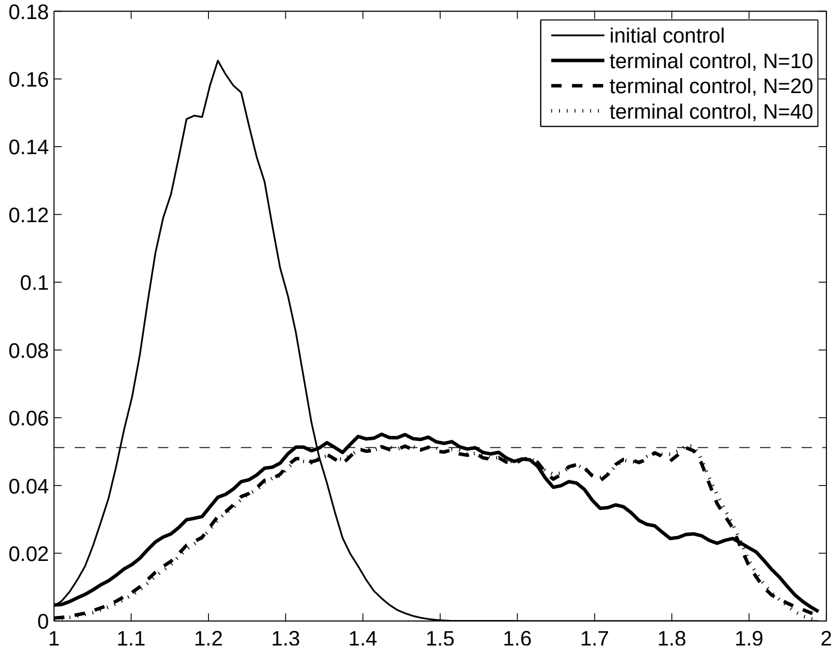


Figure 8: Concentration in Ω_1, integrated in y-direction ($B(c(x, T))$). $K = 1.5$, set U_2

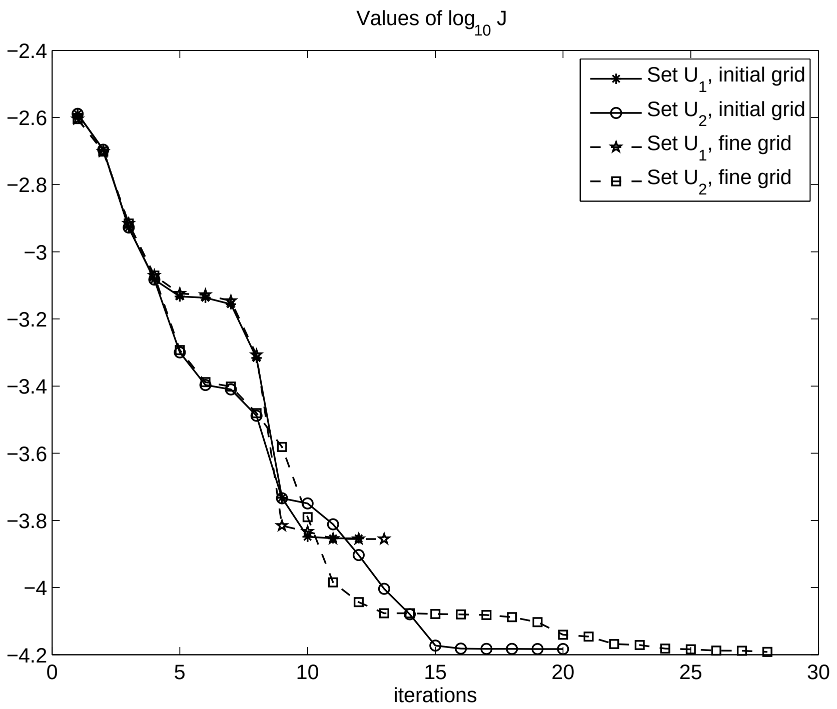


Figure 9: Values of $\log_{10} J$ with $N = 10$, $K = 1$ for two different grids.

The value of stopping criteria γ_ε in the following experiments is set to $\gamma_\varepsilon = 5 \cdot 10^{-3}$.

The first group of experiments were run with the lowest number $N = 10$ of switching points for two types of grids. They are in some sense close to each other for a given K and control set U. For all 4 considered options for K and U the decay of the functional (7) is presented on the Figures 9-10.

After reaching a stopping criteria for $N = 10$ we turned to the case $N = 20$ choosing the initial grid for the space discretization and taking the value of control u from the last meaningful iteration of the corresponding experiment for $N = 10$ as an initial guess of u. Again reaching a stopping criteria in this new case, we then turn again to a higher number of control switching points $N = 40$, in the same manner choosing an initial guess of u from the last iteration of the corresponding experiment for $N = 20$. For this group of numerical experiments with N taking values 10, 20 and 40, the decay of the functional (7) is shown on the Figures 11-12.

In the next Figure 13 we depict the values of control u for $N = 10$ being compared for two different grids for the space discretization. Only for the last iterations (when stopping criteria is attained) the value of u is drawn. Additionally for the coarser grid on the Figure 14 initial guess of u and its value on the last iteration are compared. For the experiments with $N = 20$ which are already described above, the values of control u are demonstrated on the Figure 15. These values of control are compared for the first and last meaningful iterations of the same experiment.

In the last group of plot shown on the Figure 16, the values of control are depicted for the case of $N = 40$. Here again the values of u are compared for the first and last meaningful iterations of the same experiment.

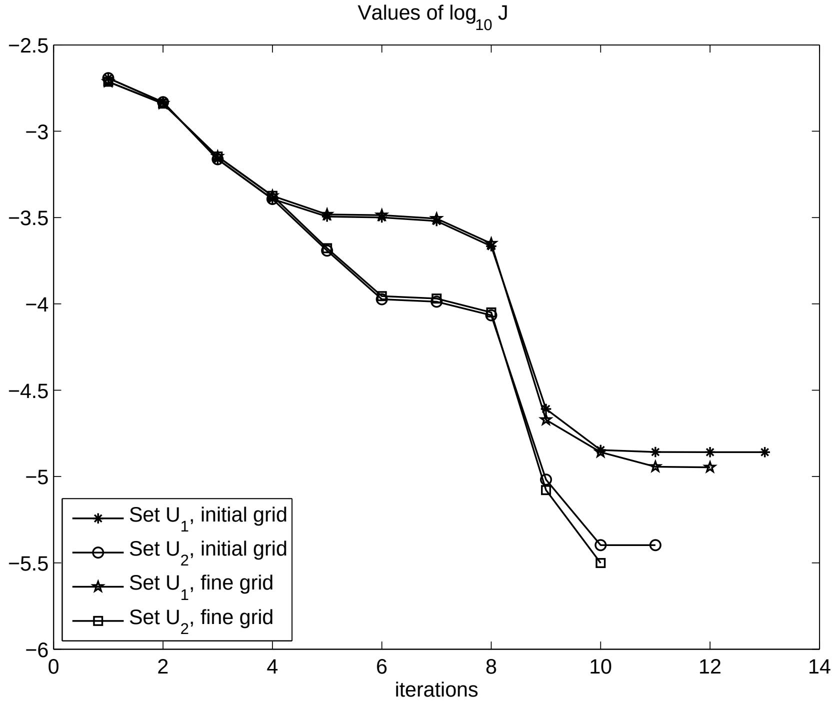


Figure 10: Values of $\log_{10} J$ with $N = 10$, $K = 1.5$ for two different grids.

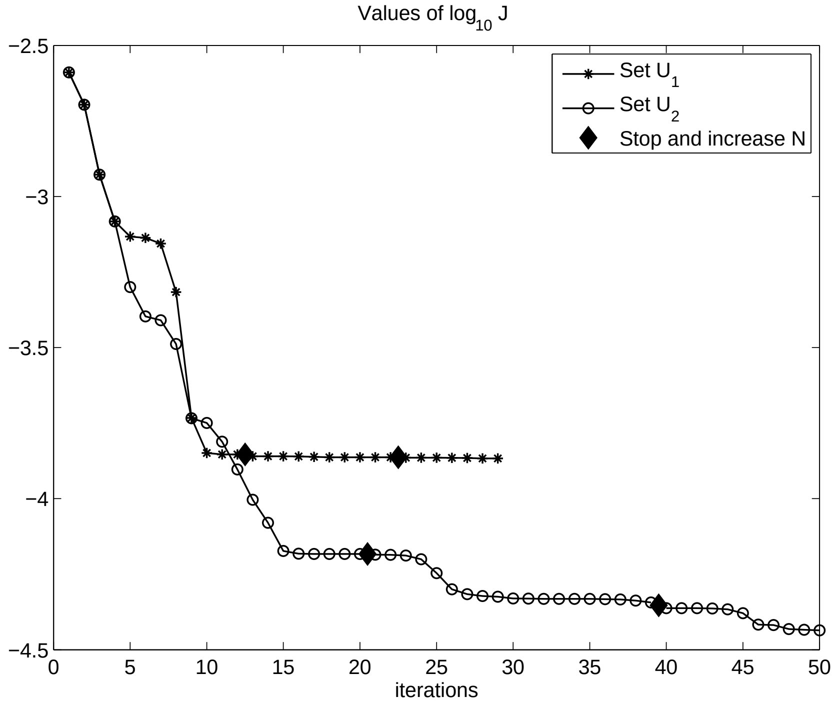


Figure 11: Values of $\log_{10} J$ with $N = 20$ and $N = 40$, $K = 1$.

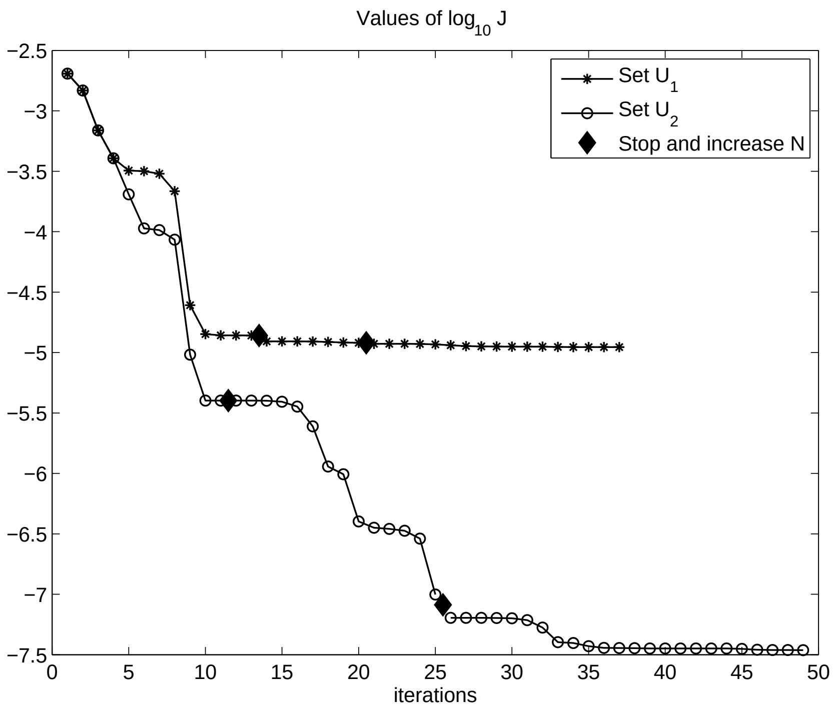
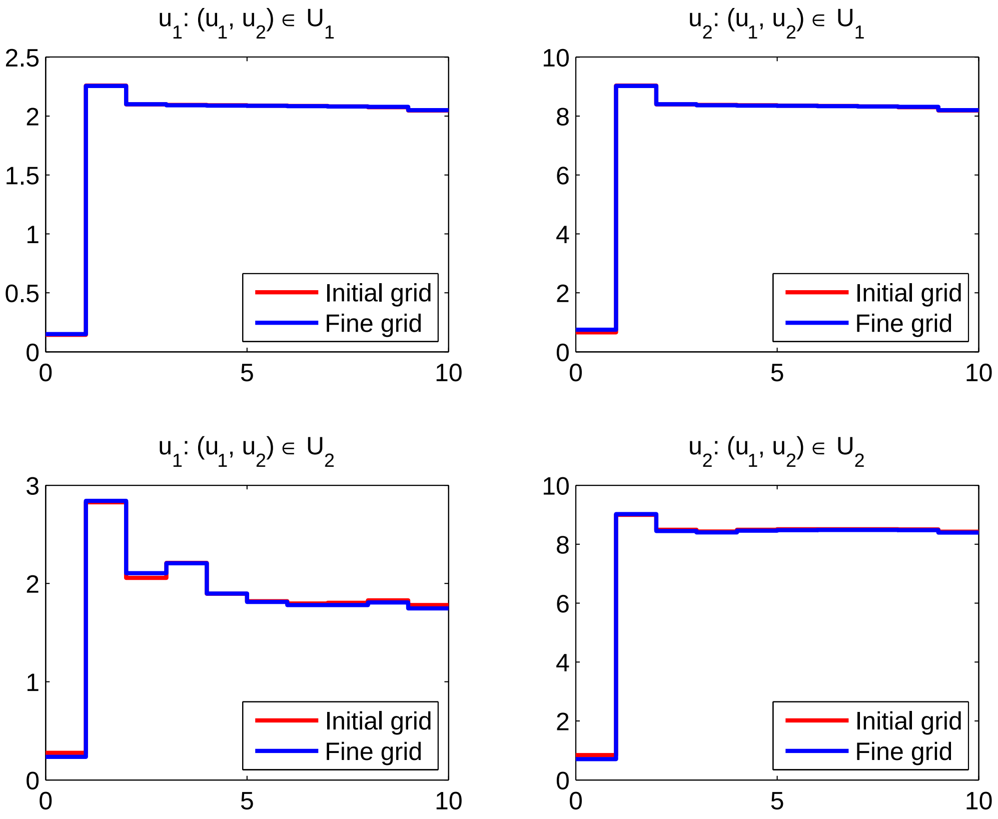


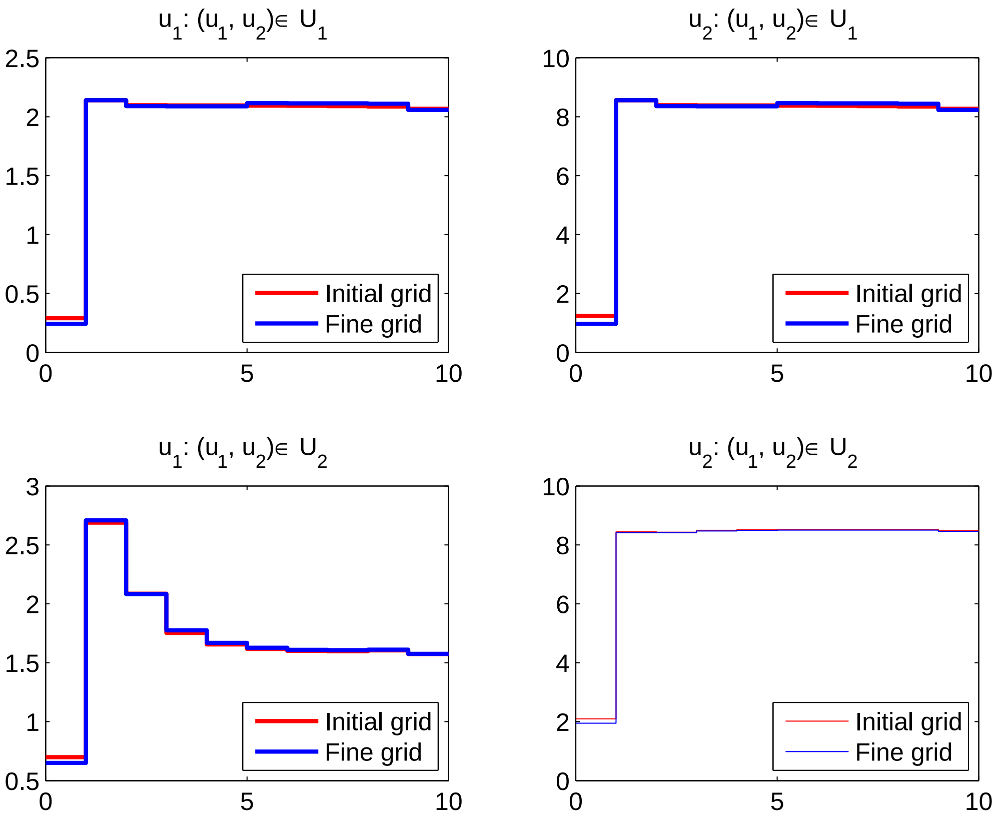
Figure 12: Values of $\log_{10} J$ with $N = 20$ and $N = 40$, $K = 1.5$.

4.3.1 *Discussion of the results.*

Our first numerical results are very promising. For the realistic setting with $K = \frac{1.5}{0.1 \cdot L} \int\limits_{\Omega} c_0(x)\, dx$ the functional value is sufficiently small for already 10 switches. Moreover, we have seen, that for some cases, we do not benefit from increasing number of switching points.

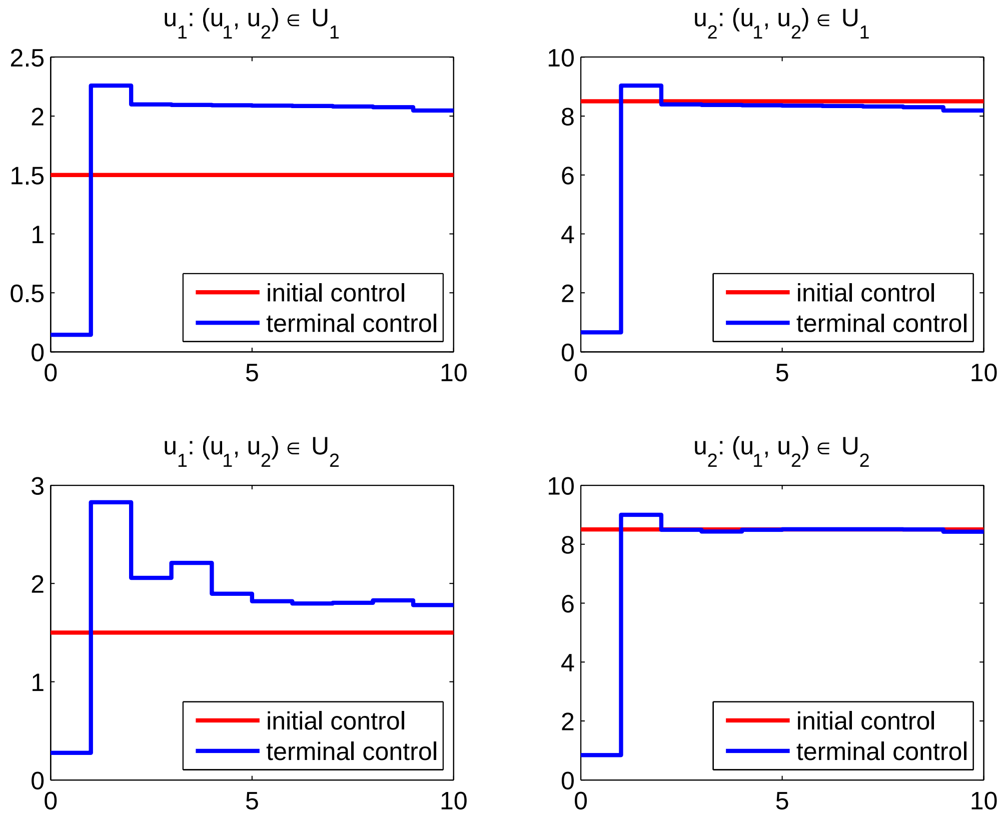


(a) $K = 1$

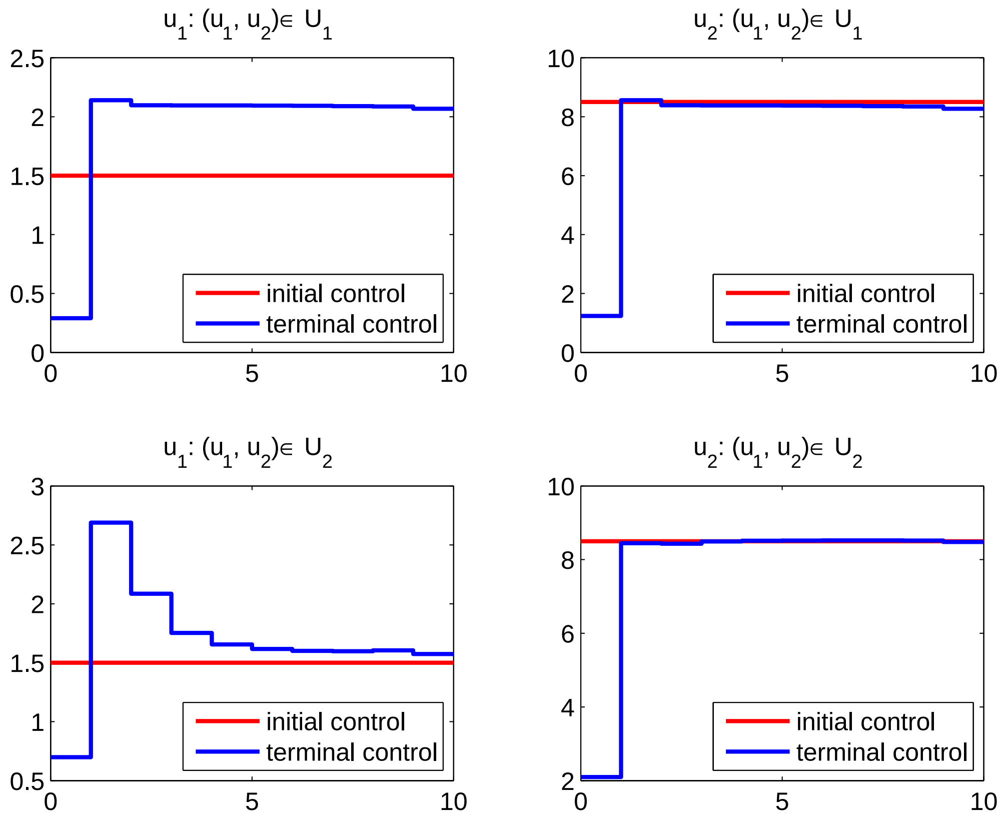


(b) $K = 1.5$

Figure 13: Terminal values of u for $N = 10$ for two different grids.

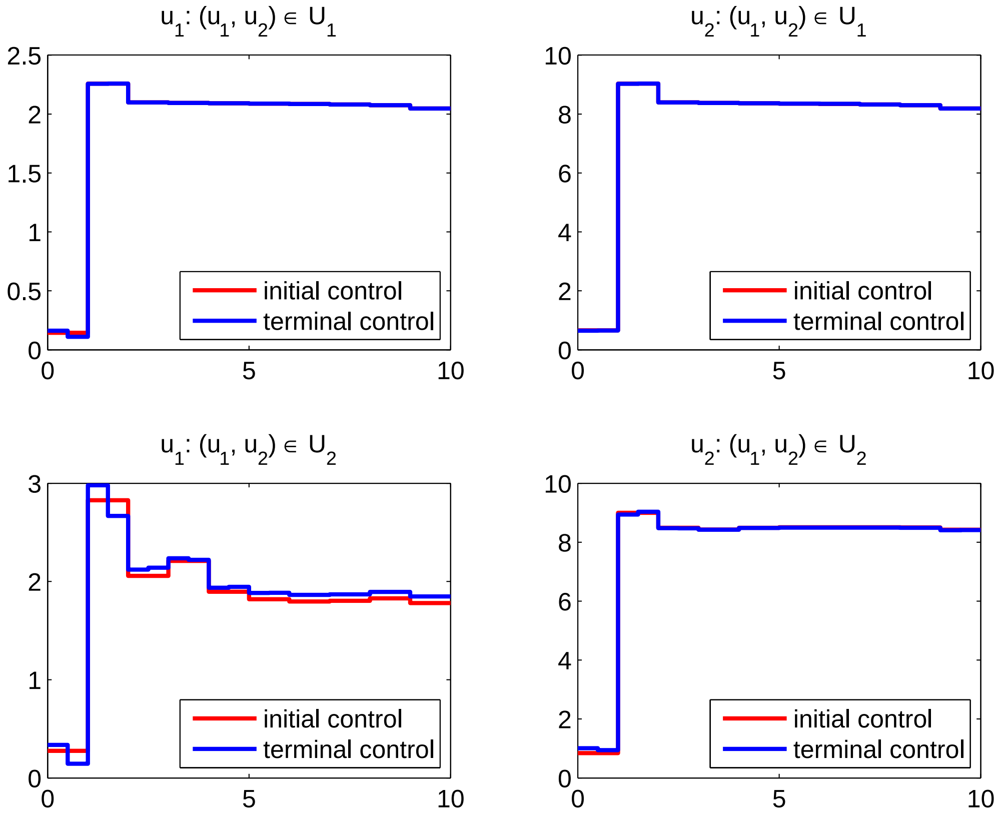


(a) $K = 1$

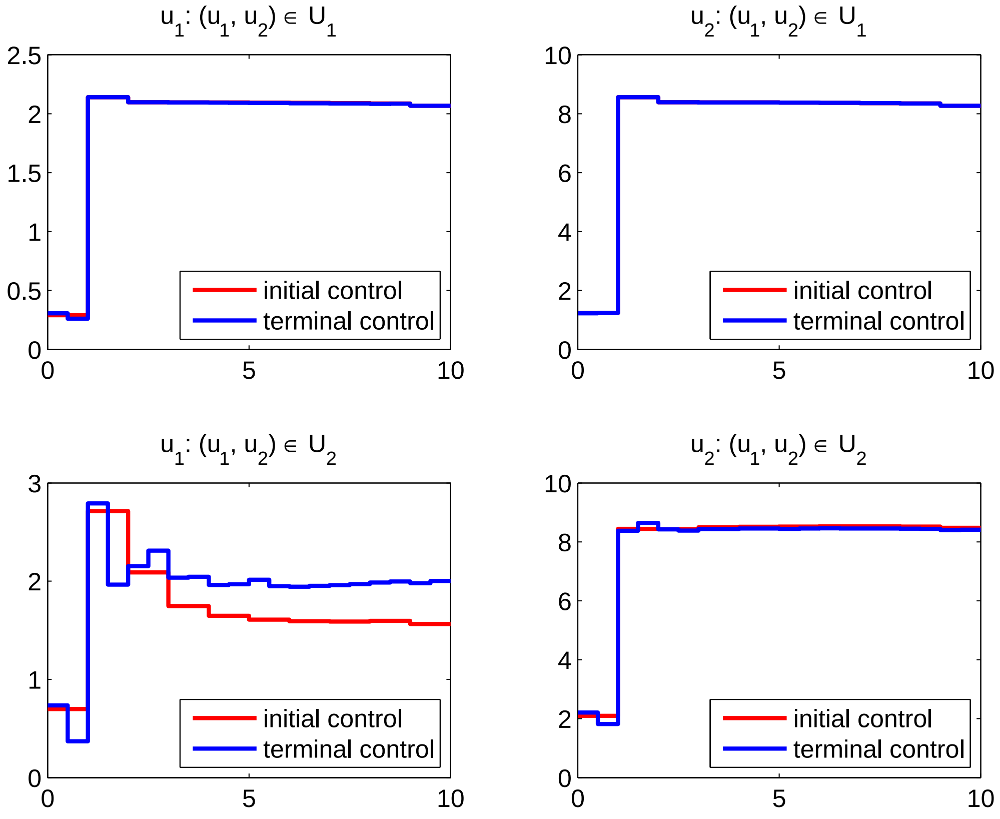


(b) $K = 1.5$

Figure 14: Initial and terminal values of u for $N = 10$.

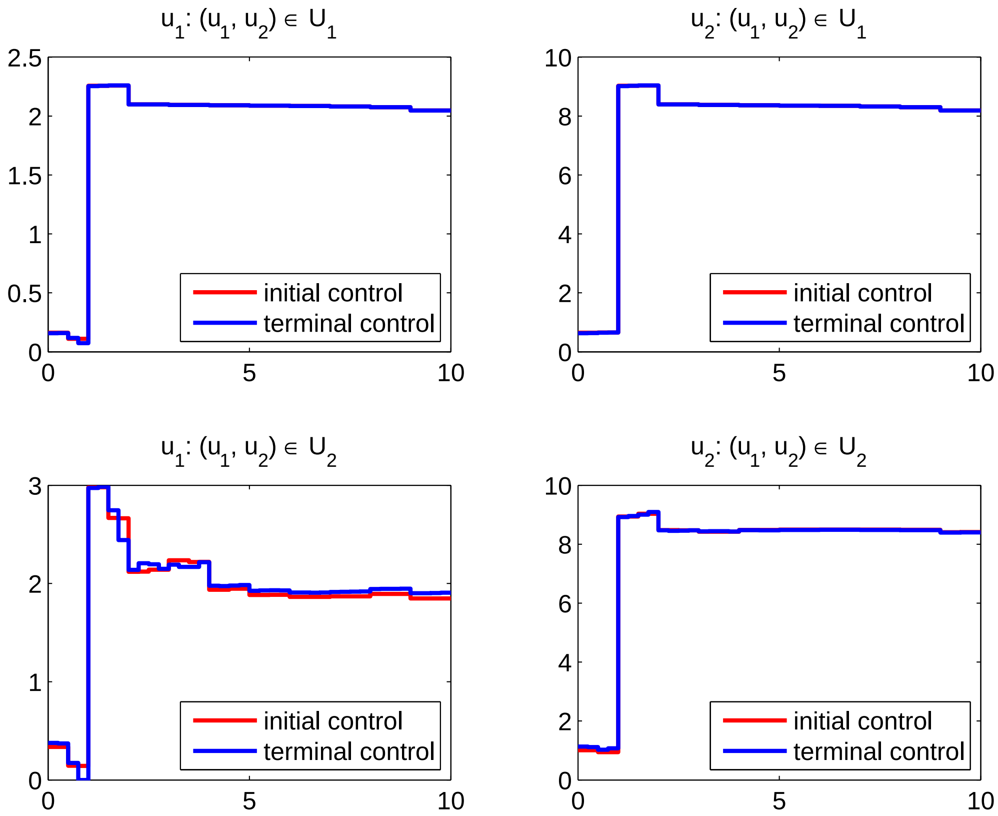


(a) $K = 1$

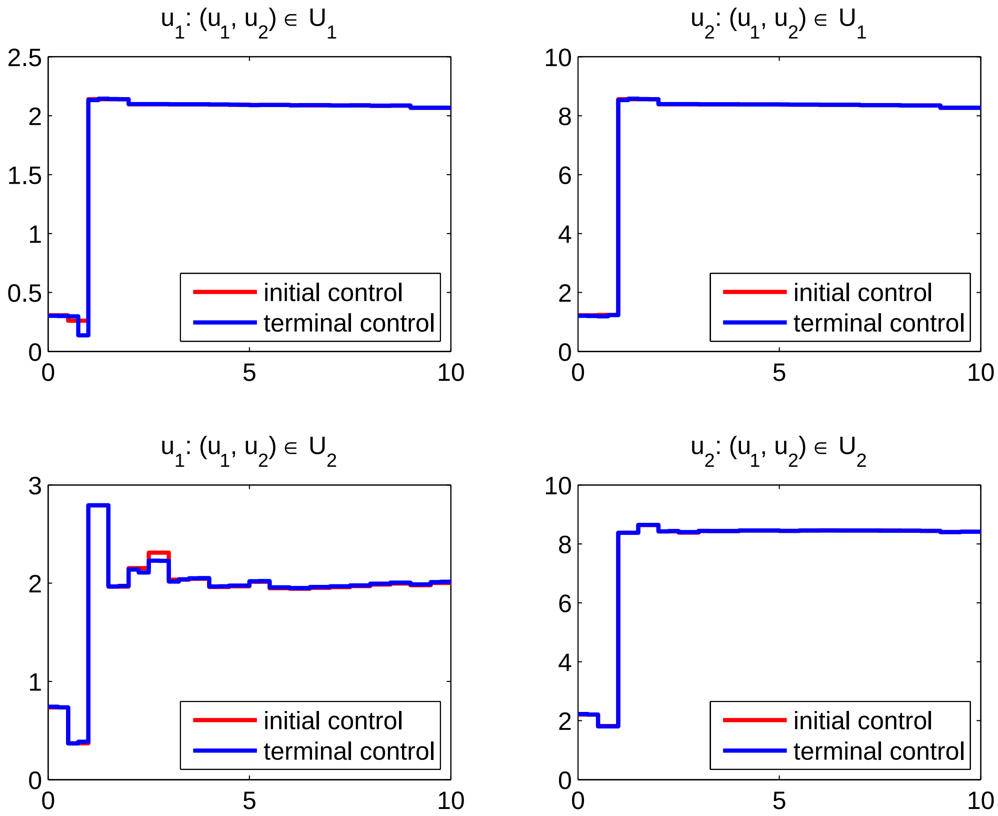


(b) $K = 1.5$

Figure 15: Initial and terminal values of u for $N = 20$.



(a) $K = 1$



(b) $K = 1.5$

Figure 16: Initial and terminal values of u for $N = 40$.

Part III

MONTE CARLO METHODS FOR APPROXIMATION OF DISTRIBUTION FUNCTIONS

In this part of the thesis we present a short overview of MLMC method with applications into finance. Then we present new MLMC algorithm for Cumulative Distribution Function (CDF) and Probability Density Function (PDF) approximations on a compact interval, and CDF approximation at a single point.

5

MONTE CARLO AND MULTILEVEL MONTE CARLO METHODS

5.1 INTRODUCTION

Monte Carlo methods are a large family of algorithms based on repeated random sampling in order to obtain numerical results. They are used in a variety of applications, where deterministic algorithms can be infeasible (e. g. high-dimensional integration, solving PDEs in high dimensions) or where the estimated quantity depends on a random variable with unknown distribution. Financial engineering is one of the most important applications for Monte Carlo algorithms (see [30]), due to the typical high-dimensionality of the problems arising in this field and their stochastic origin.

In this work we will focus on a concrete problem of estimating PDF and CDF, motivated initially by the AF4 application. We introduce Single-level Monte Carlo (SMC) (Section 5.2) and MLMC (Section 5.3) methods for estimating a functionals and compare those methods in terms of convergence rates, according to Definition 1 below.

5.2 SINGLE-LEVEL MONTE CARLO ESTIMATOR

We start from the abstract setting, where we need to estimate

$$\mathrm{E}\,P = \mathrm{E}\,f\,(Y)\,,$$

where Y is a random variable and $f\colon \mathbb{R} \to \mathbb{R}$ is a given function. We don't assume, that the distribution of Y can be simulated exactly. Instead of that we assume, that the simulation is feasible for random variables $Y^{(\ell)}$ that converge to Y in a suitable way, such that

$$\lim_{\ell\to\infty} \mathrm{E}\,f\left(Y^{(\ell)}\right) = \mathrm{E}\,f\,(Y)\,.$$

The standard Monte Carlo estimator based on N replications of $Y^{(\ell)}$ has the form

$$\hat{P} = \frac{1}{N}\sum_{i=1}^{N} f\left(Y_i^{(\ell)}\right). \tag{18}$$

We are interested in the Mean Squared Error (MSE):

$$\begin{aligned} \mathrm{E}\left[(\hat{P} - \mathrm{E}\,P)^2\right] &= \mathrm{E}\left[(\hat{P} + \mathrm{E}\,\hat{P} - \mathrm{E}\,\hat{P} - \mathrm{E}\,P)^2\right] \\ &= \mathrm{E}\left[(\hat{P} - \mathrm{E}\,\hat{P})^2\right] + \mathrm{E}\left[(\mathrm{E}\,\hat{P} - \mathrm{E}\,P)^2\right] \end{aligned} \tag{19}$$

As one can see the MSE has been decomposed into two errors:

- $\mathrm{E}\left[(\mathrm{E}\,\hat{P} - \mathrm{E}\,P)^2\right]$ – bias error;
- $\mathrm{E}\left[(\hat{P} - \mathrm{E}\,\hat{P})^2\right]$ – Monte Carlo variance.

The Monte Carlo variance is proportional to N^{-1} as

$$\mathrm{Var}\,\hat{P} = \mathrm{Var}\left(\frac{1}{N}\sum_{i=1}^{N} f(Y_i^{(\ell)})\right) = \frac{1}{N^2}\,\mathrm{Var}\left(\sum_{i=1}^{N} f(Y_i^{(\ell)})\right) = \frac{1}{N}\,\mathrm{Var}\left(f(Y^{(\ell)})\right).$$

5.3 MULTILEVEL MONTE CARLO ESTIMATOR

We will start from the simple observation:

$$\mathrm{E}\, f\left(Y^{(L)}\right) = \mathrm{E}\, f\left(Y_i^{(0)}\right) + \sum_{l=1}^{L} \mathrm{E}\left[f\left(Y^{(\ell)}\right) - f\left(Y^{(\ell-1)}\right)\right]. \tag{20}$$

The idea behind MLMC is to independently estimate each of the expectations on the right-hand side of (20) in a way, which minimises the overall variance for a given computational cost. The formal definition of the cost will be given later. The final estimator $\hat{P}$ can be seen as a sum of independent estimators

$$\hat{P} = \sum_{\ell=0}^{L} P_\ell, \tag{21}$$

where P_0 is an estimator for $\mathrm{E}\, f\left(Y_i^{(0)}\right)$ based on N_0 samples, and Y_ℓ are estimates for

$$\mathrm{E}\left[f\left(Y^{(\ell)}\right) - f\left(Y^{(\ell-1)}\right)\right]$$

based on N_ℓ samples. The simplest form for P_0 and P_ℓ is a mean value of f over all samples:

$$P_0 = \frac{1}{N_0} \sum_{i=1}^{N_0} f\left(Y_i^{(0)}\right),$$

$$P_\ell = \frac{1}{N_\ell} \sum_{i=1}^{N_\ell} \left[f\left(Y^{(\ell)}\right) - f\left(Y^{(\ell-1)}\right)\right], \ \ell = 1, \dots, L.$$

For the variance and bias for $\hat{P}$ we have

$$\operatorname{Var} \hat{P} = \operatorname{Var}\left[\sum_{\ell=0}^{L} P_\ell\right] = \sum_{\ell=0}^{L} \operatorname{Var} P_\ell,$$

$$\mathrm{E}\, \hat{P} = \mathrm{E}\left[\sum_{\ell=0}^{L} P_\ell\right] = \mathrm{E}\, f\left(Y^{(L)}\right).$$

5.4 COMPLEXITY ESTIMATORS

Definition 1 *We say that a sequence of randomized algorithms $\mathcal{A}_n$ converges with order $(\gamma, \eta) \in]0, \infty[\times \mathbb{R}$ if*

$$\lim_{n \to \infty} \operatorname{error}(\mathcal{A}_n) = 0$$

and if there exists a constant $c > 0$ such that

$$\operatorname{cost}(\mathcal{A}_n) \leqslant c \cdot (\operatorname{error}(\mathcal{A}_n))^{-\gamma} \cdot (-\log \operatorname{error}(\mathcal{A}_n))^{\eta}.$$

By cost we will understand mathematical expectation of random number calls made during the algorithm.

Remark 5.4.1 *In our notation the low convergence rate means lower computational cost in order to achieve given accuracy.*

Now for MLMC the overall cost is proportional to $\sum_{\ell=0}^{L} N_\ell / h^\ell$, h_ℓ is a time step discretization at the level ℓ.

The next two theorems will provide the complexity estimates for both methods.

Theorem 5.4.2 *Consider* $\mathrm{E}\,Y = \mathrm{E}(f(Y)) = \mathrm{E}(f(\varphi(W)))$ *and an approximation* $Y^\ell = \varphi^\ell(W)$*, where* φ^ℓ *is a discretization scheme with stepsize* $h = T/2^\ell$*. Consider the estimator* (18) *and assume that the following is true with some universal constant* c*:*

- $f\colon \mathbb{R} \to \mathbb{R}$ *and*
$$\operatorname{cost}(f(x)) \leqslant c, \quad \forall x \in \mathbb{R}$$
- *There exists* $M > 1$ *such that*
$$\operatorname{cost}(Y^{(\ell)}) \leqslant c \cdot M^\ell, \quad \forall \ell \in \mathbb{N}_0$$
- *There exists* $\alpha > 0$
$$\left|\mathrm{E}\left(f(Y) - f\left(Y^{(\ell)}\right)\right)\right| \leqslant c \cdot M^{-\alpha\ell}, \quad \forall \ell \in \mathbb{N}_0$$
- $\sup_{\ell \in \mathbb{N}_0} \operatorname{Var}(f(Y^{(\ell)})) < \infty$

Then the SMC *method achieves order of convergence*

$$\gamma = \frac{2 \cdot \alpha + 1}{\alpha}, \eta = 0.$$

Proof It is straightforward to verify, that $N \asymp \varepsilon^{-2}$ and $h \asymp \varepsilon^\alpha$, so

$$\operatorname{cost} = N/h = \varepsilon^{-2-\frac{1}{\alpha}}.$$

The next theorem was proved in [25].

Theorem 5.4.3 *Consider* $\mathrm{E}\,Y = \mathrm{E}(f(Y)) = \mathrm{E}(f(\varphi(W)))$ *and an approximation* $Y^\ell = \varphi^\ell(W)$*, where* φ *is a discretization scheme with stepsize* $h_\ell = T/2^\ell$*. Consider*

$$\hat{P} = \frac{1}{N_0} \sum_{i=1}^{N_0} f\left(Y_i^{(0)}\right) + \sum_{\ell=1}^{L_1} \frac{1}{N_\ell} \sum_{i=1}^{N_\ell} \left(f(Y_i^{(\ell)}) - f(Y_i^{(\ell-1)})\right) \tag{22}$$

and assume that

- $f \in \operatorname{Lip}(\mathbb{R})$ *and*
$$\operatorname{cost}(f(x)) \leqslant c, \quad \forall x \in \mathbb{R}$$
- *There exists* $M > 1$*, such that*
$$\operatorname{cost}(Y^{(\ell)}, Y^{(\ell-1)}) \leqslant c \cdot M^\ell, \quad \forall \ell \in \mathbb{N}_0$$
- *There exists* $\alpha > 0$*, such that*
$$|\mathrm{E}(f(Y) - f(Y^{(\ell)}))| \leqslant c \cdot M^{-\alpha\ell}, \quad \forall \ell \in \mathbb{N}_0$$
- *There exists* $\beta \in]0, \alpha]$ *such that*
$$(\mathrm{E}(Y - Y^{(\ell)})^2)^{1/2} \leqslant c \cdot M^{-\beta\ell}, \quad \forall \ell \in \mathbb{N}_0.$$

Then the convergence rate for MLMC *algorithm is*

$$\gamma = \begin{cases} 2, & \beta \geqslant \frac{1}{2} \\ \frac{2(\alpha-\beta)+1}{\alpha}, & \beta < \frac{1}{2} \end{cases}, \qquad \eta = \begin{cases} 2, & \beta = \frac{1}{2} \\ 0, & \beta \neq \frac{1}{2} \end{cases}.$$

Proof Due to the fact, that $f \in \mathrm{Lip}(\mathbb{R})$ we have

$$\begin{aligned} \mathrm{Var}\left(f(Y) - f(Y^{(\ell)})\right) &\leqslant \mathrm{E}\left(f(Y) - f(Y^{(\ell)})\right)^2 \\ &\leqslant c \cdot \mathrm{E}\left(Y - Y^{(\ell)}\right)^2 \leqslant M^{-2\cdot\beta\ell}. \end{aligned}$$

So now we can write an optimization problem.

Find N_ℓ, L such that:

$$\begin{cases} g_1(N_\ell, L) = \sum\limits_{\ell=0}^{L} \dfrac{N_\ell}{h_\ell} \to \min, \\ g_2(N_\ell, L) = \sum\limits_{\ell=0}^{L} \dfrac{h_\ell^{2\cdot\beta}}{N_\ell} \preceq \varepsilon^2, \\ g_3(L) = h^{2\cdot\alpha\cdot L} \preceq \varepsilon^2. \end{cases}$$

It is clear, that $L \geqslant \frac{1}{\alpha} \log_2 \varepsilon^{-1}$.

Now by differentiation we obtain:

$$\frac{\partial g_1}{\partial N_\ell} = \frac{1}{h_\ell}; \quad \frac{\partial g_2}{\partial N_\ell} = -\frac{h_\ell^{2\cdot\beta}}{N_\ell^2}$$

Hence by the Lagrange principle

$$\left.\frac{\partial g_1}{\partial N_\ell}\right|_{N^\star} = \lambda \left.\frac{\partial g_2}{\partial N_\ell}\right|_{N^\star} \Rightarrow \frac{1}{h_\ell} = -\lambda \frac{h_\ell^{2\cdot\beta}}{(N_\ell^\star)^2} \Leftrightarrow (N_\ell^\star)^2 = -\lambda h_\ell^{2\cdot\beta+1}$$

$$\Downarrow$$

$$N_\ell^\star = (-\lambda)^{\frac{1}{2}} h_\ell^{\frac{2\cdot\beta+1}{2}}.$$

Now we find λ from $g_2(N^\star) \asymp \varepsilon^2$:

$$\sum_{\ell=1}^{L} \frac{h_\ell^{2\cdot\beta}}{N_\ell^\star} = \sum_{\ell=1}^{L} \frac{h_\ell^{2\cdot\beta}}{(-\lambda)^{\frac{1}{2}} h_\ell^{\frac{2\cdot\beta+1}{2}}} = \varepsilon^2 \Longrightarrow \lambda = -\left(\frac{\sum\limits_{\ell=0}^{L} h_\ell^{\frac{2\cdot\beta-1}{2}}}{\varepsilon^2}\right)^2.$$

So, we have the number of experiments as given by

$$N_\ell^\star \asymp \varepsilon^{-2} \cdot h_\ell^{\frac{2\cdot\beta+1}{2}} \cdot \left(\sum_{\ell=0}^{L} h_\ell^{\frac{2\cdot\beta-1}{2}}\right)$$

and complexity

$$\mathrm{cost} = \sum_{\ell=1}^{L} \frac{N_\ell}{h_\ell} = 2 \cdot \varepsilon^{-2} \cdot \left(\sum_{\ell=0}^{L} h_\ell^{\frac{2\cdot\beta-1}{2}}\right)^2.$$

So now we can consider three main cases:

$$\left(\sum_{\ell=0}^{L} h_\ell^{\frac{2\cdot\beta-1}{2}}\right)^2 \asymp \begin{cases} 1, & \beta > \frac{1}{2} \\ (\log_2 \varepsilon)^2, & \beta = \frac{1}{2}, \\ \varepsilon^{\frac{2\cdot\beta-1}{\alpha}}, & \beta < \frac{1}{2} \end{cases}$$

which concludes the proof.

6

MULTILEVEL MONTE CARLO METHOD FOR CDF AND PDF APPROXIMATION ON A COMPACT INTERVAL

Let Y denote a real-valued random variable with distribution function F and density ρ. We study the approximation of F and ρ with respect to the supremum norm on a compact interval $[S_0, S_1]$, without assuming that the distribution of Y is explicitly known or that the simulation of Y is feasible. Instead, we suppose that a sequence of random variables $Y^{(\ell)}$ is at hand that converge to Y in a suitable way and that are suited to simulation.

We present a general approach, which is later on applied in the context of stochastic differential equations (SDEs). In this specific setting we aim at the distribution of Lipschitz continuous, path-independent or path-dependent functionals of the solution process, or the distribution of stopped exit times from bounded domains.

In the general setting a naive Monte Carlo algorithm for the approximation of ρ works as follows: Choose a level $\ell \in \mathbb{N}$ and a replication number $n \in \mathbb{N}$, generate n independent samples according to $Y^{(\ell)}$, and apply a kernel density estimator, say, to these samples. For the approximation of F one proceeds analogously, and here the empirical distribution function of the samples is the most elementary choice.

In this chapter we develop the multi-level Monte Carlo approach, which relies on the coupled simulation of $Y^{(\ell)}$ and $Y^{(\ell-1)}$ on a finite range of levels ℓ. For the multi-level approach to work well for the approximation of distribution functions or densities, a smoothing step is necessary on every level. The smoothing is based on rescaled translates of a suitable function g, which is meant to approximate either the indicator function of $]-\infty, 0]$ or the Dirac functional at zero. In a first stage the multi-level algorithm provides an approximation to F or ρ at discrete points, which is then extended to a function on $[S_0, S_1]$ in a standard and purely deterministic way.

For the approximation of F and ρ on $[S_0, S_1]$ our assumptions are as follows:

(i) The density ρ of Y is r-times continuously differentiable,

(ii) The simulation of the joint distribution of $Y^{(\ell)}$ and $Y^{(\ell-1)}$ is possible at cost $O(M^\ell)$ for every $\ell \in \mathbb{N}$, where $M > 1$.

(iii) A weak error estimate

$$\sup_{s\in[S_0,S_1]} \left| E\left(g((Y-s)/\delta) - g((Y^{(\ell)}-s)/\delta) \right) \right| \leqslant \\ \leqslant O\left(\min\left(\delta^{-\alpha_1} \cdot M^{-\ell\cdot\alpha_2}, M^{-\ell\cdot\alpha_3} \right) \right)$$

holds for all positive, sufficiently small δ and all $\ell \in \mathbb{N}_0$, where $\alpha_1 \geqslant 0$, $\alpha_2 > 0$, and $\alpha_2 \geqslant \alpha_3 \geqslant 0$.

(iv) A strong error estimate

$$E \min((Y - Y^{(\ell)})^2/\delta^2, 1) \leqslant O\left(\delta^{-\beta_4} \cdot M^{-\ell\cdot\beta_5} \right)$$

holds for all positive, sufficiently small δ and all $\ell \in \mathbb{N}_0$, where $\beta_4 \geqslant 0$ and $\beta_5 > 0$.

We also study the approximation of the distribution function F at a single point $s \in [S_0, S_1]$, and here (iv) is replaced by the following assumption:

(v) A strong error estimate

$$\sup_{s\in[S_0,S_1]} \mathrm{E}\left(g((Y-s)/\delta) - g((Y^{(\ell)}-s)/\delta)\right)^2 \leqslant$$
$$\leqslant O\left(\min\left(\delta^{-\beta_1}\cdot M^{-\ell\cdot\beta_2}, M^{-\ell\cdot\beta_3}\right)\right)$$

holds for all positive, sufficiently small δ and all $\ell \in \mathbb{N}_0$, where $\beta_1 \geqslant 0$, $\beta_2 > 0$, and $\beta_2 \geqslant \beta_3 \geqslant 0$.

The parameters of a multi-level algorithm $\mathcal{A}$ are the minimal and maximal level, the replication numbers per level, the smoothing parameter δ, and the number of discrete points to be used in the first stage. We derive upper bounds for $\mathrm{error}(\mathcal{A})$, the root mean square error, and $\mathrm{cost}(\mathcal{A})$, the computational cost, in terms of these parameters and the values of r, α_i, and β_i, and we present the asymptotically optimal choice of the parameters with respect to our upper bounds. This leads to a final estimate of the form

$$\mathrm{cost}(\mathcal{A}) \leqslant O\left(\mathrm{error}(\mathcal{A})^{-\theta+\varepsilon}\right)$$

for every $\varepsilon > 0$, where $\theta > 0$. Roughly speaking, θ is the order of convergence of the multi-level algorithm. See Theorems 6.1.5–6.3.3 for the precise statements involving also powers of $\log \mathrm{error}(\mathcal{A})$.

Here we only present a particular application of these theorems for functionals

$$\varphi : C([0,T], \mathbb{R}^d) \to \mathbb{R}$$

of the solution process X of a d-dimensional system of SDEs, i.e., $Y = \varphi(X)$. For simplicity we take the Euler scheme with equidistant time-steps for the approximation of X in the construction of the multilevel algorithm, and we assume that $r \geqslant 1$ for the rest of the introduction. Table 1 contains the values of θ for the approximation of F and ρ on $[S_0, S_1]$ as well as for the approximation of F at a single point $s \in [S_0, S_1]$. In the first row φ is assumed to be Lipschitz continuous, and based on a well known upper bound for the strong error of the Euler scheme we show that (iii)–(v) are satisfied with

$$\alpha_1 = 0, \qquad \alpha_2 = 1/2 - \varepsilon, \qquad \alpha_3 = 1/2 - \varepsilon$$

and

$$\beta_1 = 1 + \varepsilon, \qquad \beta_2 = 1 - \varepsilon, \qquad \beta_3 = 1/2 - \varepsilon, \qquad \beta_4 = 2, \qquad \beta_5 = 1 - \varepsilon$$

for every $\varepsilon > 0$. In the second row

$$\varphi(x) = \inf\{t \geqslant 0 : x(t) \in \partial D\} \wedge T$$

is a stopped exit time from a bounded domain $D \subset \mathbb{R}^d$, and based on a recent result from [14] we obtain

$$\alpha_1 = 1, \qquad \alpha_2 = 1/2, \qquad \alpha_3 = 1/4$$

and

$$\beta_1 = 1, \qquad \beta_2 = 1/2, \qquad \beta_3 = 1/4, \qquad \beta_4 = 1, \qquad \beta_5 = 1/2.$$

We add that in every case represented in Table 1 proper multi-level algorithms turn out to be superior to single-level algorithms, as far as our upper bounds are concerned. We do not achieve better upper bounds if we restrict considerations to path-independent functionals, i.e., $Y = \varphi(X_T)$ with $\varphi : \mathbb{R}^d \to \mathbb{R}$ being Lipschitz continuous; here, however, the situation changes if the Euler scheme is replaced by the Milstein scheme (in dimension $d = 1$ because of assumption (ii)), which yields $\theta = 2 + 1/(r+1)$, $\theta = 2 + 3/r$, and $\theta = 2$ for the approximation of F, ρ, and $F(s)$, respectively.

	F	ρ	F(s)
smooth functional	$2+\frac{2}{r+1}$	$2+\frac{4}{r}$	$2+\frac{1}{r+1}$
stopped exit time	$3+\frac{2}{r+1}$	$3+\frac{5}{r}$	$3+\frac{2}{r+1}$

Table 1: Orders of convergence of the multi-level algorithm

Corresponding results are available for the approximation of the expectation of $\varphi(X)$ by means of multi-level Euler algorithms. It is well known that $\theta = 2$, if φ is Lipschitz continuous, and $\theta = 3$ holds for stopped exit times φ, see [35]. In the limit $r \to \infty$ we achieve the same values of θ for the approximation of the distribution function or the density of $\varphi(X)$.

Multi-level algorithms, which have been introduced in [34] and [25], see also [42] for the two-level construction, are meanwhile applied to rather different computational problems. The approximation of distribution functions and densities seems to be a new application, which exhibits, in particular, the following features: a singularity, which is due to the presence of the indicator function or the Dirac functional, and the fact that we approximate elements of function spaces instead of just real numbers. The first issue is also investigated, without smoothing, in [5] and [27], and with implicit smoothing through the use of conditional expectations in [26] and [28]. Furthermore, Altmayer, Neuenkirch in [2] combine smoothing by Malliavin integration by parts with the multi-level approach to approximate expectations of discontinuous payoffs in the Heston model. The second issue has already been worked out by Heinrich (see [34]) in the general setting of algorithms taking values in Banach spaces.

We stress that a two-level construction for the approximation of densities in the SDE setting with $Y = X_T$ has already been proposed and analyzed by Kebaier, Kohatsu-Higa (see [43]) in the case $r = \infty$, and their analysis yields $\theta = 5/2$.

Optimality results, which do not just concern upper bounds for the error and cost of specific families of algorithms, seem to be unknown for the problems studied in the present work. The situation is different for the approximation of expectations of Lipschitz continuous functionals, and here suitable multi-level algorithms are almost worst case optimal in the class of all randomized algorithms, see [19].

The rest of the chapter is organized as follows. In Sections 6.1–6.3 we provide the general analysis of the three approximation problem, namely, for distribution functions and densities on compact intervals and for distribution functions at a single point. The structure and the approach in each of these sections is similar: we discuss, in particular, the assumptions on the weak and the strong convergence, and we construct and analyze the respective multi-level algorithms. Section 7.1 contains, in particular, the application of the results from Sections 6.1–6.3 to functionals of the solutions of SDEs, which is complemented by numerical experiments for simple test cases in Section 7.2.

6.1 APPROXIMATION OF DISTRIBUTION FUNCTIONS ON COMPACT INTERVALS

We consider a random variable Y, and we study the approximation of its distribution function F on a compact interval $[S_0, S_1]$, with $S_0 < S_1$ being fixed throughout this section. We do not assume that the distribution of Y can be simulated exactly. Instead, we assume that the simulation is feasible for random variables $Y^{(\ell)}$ that converge to Y in a suitable way.

6.1.1 *Smoothing*

For the approximation of F a straight-forward application of the multi-level Monte Carlo approach based on

$$F(s) = E(1_{]-\infty,s]}(Y))$$

could suffer from the discontinuity of $1_{]-\infty,s]}$, see Remark 7 below. This can be avoided by a smoothing step, provided that a density exists and is sufficiently smooth. Specifically, we assume that

(A1) the random variable Y has a density ρ on $\mathbb{R}$ that is r-times continuously differentiable on $[S_0 - \delta_0, S_1 + \delta_0]$ for some $r \in \mathbb{N}_0$ and $\delta_0 > 0$.

The smoothing is based on rescaled translates of a function $g : \mathbb{R} \to \mathbb{R}$ with the following properties:

(S1) The cost of computing $g(s)$ is bounded by a constant, uniformly in $s \in \mathbb{R}$.

(S2) g is Lipschitz continuous.

(S3) $g(s) = 1$ for $s < -1$ and $g(s) = 0$ for $s > 1$.

(S4) $\int_{-1}^{1} s^j \cdot (1_{]-\infty,0]}(s) - g(s))\, ds = 0$ for $j = 0, \dots, r-1$.

Obviously, g is bounded due to (S2) and (S3).

Remark 1 *Such a function* g *is easily constructed as follows. There exists a uniquely determined polynomial* p *of degree at most* $r+1$ *such that*

$$\int_{-1}^{1} s^j \cdot p(s)\, ds = (-1)^j/(j+1), \quad j = 0, \dots, r-1,$$

as well as $p(1) = 0$ *and* $p(-1) = 1$*. The extension* g *of* p *with* $g(s) = 1$ *for* $s < -1$ *and* $g(s) = 0$ *for* $s > 1$ *has the properties as claimed. Since* $g - 1/2$ *is an odd function, the same function* g *arises in this way for* r *and* $r+1$*, if* r *is even.*

For example, consider $r = 5$ *and put*

$$p(s) = p_0 + p_1 s + p_2 s^2 + p_3 s^3 + p_4 s^4 + p_5 s^5 + p_6 s^6.$$

To find the coefficients p_i*, we need 7 equations. From the conditions* $p(1) = 0$ *and* $p(-1) = 1$ *we immediately obtain*

$$\sum_{i=0}^{6} p_i = 0, \text{ and } \sum_{i=0}^{6} (-1)^i p_i = 1.$$

Futher, using

$$\int_{-1}^{1} s^j p_i s^i\, ds = \frac{1}{1+i+j}\left(1 - (-1)^{i+j+1}\right) \tag{23}$$

we arrive to the system of equations for p_i*:*

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 \\ 2 & 0 & \frac{2}{3} & 0 & \frac{2}{5} & 0 & \frac{2}{7} \\ 0 & \frac{2}{3} & 0 & \frac{2}{5} & 0 & \frac{2}{7} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} & 0 & \frac{2}{7} & 0 & \frac{2}{9} \\ 0 & \frac{2}{5} & 0 & \frac{2}{7} & 0 & \frac{2}{9} & 0 \\ \frac{2}{5} & 0 & \frac{2}{7} & 0 & \frac{2}{9} & 0 & \frac{2}{11} \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \end{pmatrix} = \begin{pmatrix} c0 \\ 1 \\ 1 \\ -\frac{1}{2} \\ \frac{1}{3} \\ -\frac{1}{4} \\ \frac{1}{5} \end{pmatrix}$$

and obtain the solution vector

$$\begin{pmatrix} \frac{1}{2} & -\frac{225}{128} & 0 & \frac{175}{64} & 0 & -\frac{189}{128} & 0 \end{pmatrix}.$$

We have the following estimate for the bias that is induced by smoothing with parameter δ, i.e., by approximation of $1_{]-\infty,s]}$ by $g((\cdot - s)/\delta)$.

Lemma 6.1.1 *There exists a constant* $c > 0$ *such that*

$$\sup_{s\in[S_0,S_1]} |F(s) - E(g((Y-s)/\delta))| \leqslant c \cdot \delta^{r+1}$$

holds for all $\delta \in \,]0, \delta_0]$.

Proof Clearly

$$\begin{aligned} F(s) - E(g((Y-s)/\delta)) &= \int_{-\infty}^{\infty} \rho(u) \cdot (1_{]-\infty,s]}(u) - g((u-s)/\delta))\, du \\ &= \delta \cdot \int_{-1}^{1} \rho(\delta u + s) \cdot (1_{]-\infty,0]}(u) - g(u))\, du, \end{aligned}$$

so that the statement follows in the case $r = 0$. For $r \geqslant 1$ the Taylor expansion

$$\rho(\delta u + s) = \sum_{j=0}^{r-1} \rho^{(j)}(s) \cdot (\delta u)^j / j! + R(\delta u, s)$$

yields

$$|F(s) - E(g((Y-s)/\delta))| \leqslant \delta \cdot \int_{-1}^{1} |R(\delta u, s)| \cdot |1_{]-\infty,0]}(u) - g(u)|\, du \leqslant c \cdot \delta^{r+1}.$$

6.1.2 *Assumptions on Weak and Strong Convergence*

Our multi-level Monte Carlo construction is based on a sequence $(Y^{(\ell)})_{\ell\in\mathbb{N}_0}$ of random variables, defined on a common probability space together with Y, with the following properties for some constant $c > 0$:

(A2) There exists a constant $M > 1$ such that the simulation of the joint distribution of $Y^{(\ell)}$ and $Y^{(\ell-1)}$ is possible at cost at most $c \cdot M^{\ell}$ for every $\ell \in \mathbb{N}$.

(A3) There exist constants $\alpha_1 \geqslant 0$, $\alpha_2 > 0$, and $\alpha_2 \geqslant \alpha_3 \geqslant 0$ such that the weak error estimate

$$\begin{aligned} \sup_{s\in[S_0,S_1]} \left| E\left(g((Y-s)/\delta) - g((Y^{(\ell)}-s)/\delta)\right) \right| &\leqslant \\ &\leqslant c \cdot \min\left(\delta^{-\alpha_1} \cdot M^{-\ell\cdot\alpha_2}, M^{-\ell\cdot\alpha_3}\right) \end{aligned}$$

holds for all $\delta \in \,]0, \delta_0]$ and $\ell \in \mathbb{N}_0$.

(A4) There exist constants $\beta_4 \geqslant 0$ and $\beta_5 > 0$ such that the strong error estimate

$$E \min((Y - Y^{(\ell)})^2/\delta^2, 1) \leqslant c \cdot \delta^{-\beta_4} \cdot M^{-\ell\cdot\beta_5}$$

holds for all $\delta \in \,]0, \delta_0]$ and $\ell \in \mathbb{N}_0$.

For specific applications we present suitable approximations $Y^{(\ell)}$ and corresponding values of the parameters M, α_i and β_i in Section 7.1. Here we proceed with a general discussion of (A3) and (A4).

Lemma 6.1.2 *Let* g *satisfy (S2) and (S3). Then (A4) implies (A3) with* $\alpha_1 = \beta_4/2$, $\alpha_2 = \beta_5/2$, *and* $\alpha_3 = 0$.

Proof

$$\sup_{s\in[S_0,S_1]}\left|E\left(g((Y-s)/\delta)-g((Y^{(\ell)}-s)/\delta)\right)\right|\leqslant$$

$$\leqslant\sup_{s\in[S_0,S_1]}E\left|g\left((Y-s)/\delta\right)-g\left((Y^{(\ell)}-s)/\delta\right)\right|$$

Due to (S2) the following is true:

$$\left|g((Y-s)/\delta)-g((Y^{(\ell)}-s)/\delta)\right|\leqslant c_L\left|\frac{Y-s}{\delta}-\frac{Y^{(\ell)}-s}{\delta}\right|=c_L\left|\frac{Y-Y^{(\ell)}}{\delta}\right|.$$

At the same time, (S3) and (S2) together yield boundedness of g, and the same expression can be estimated in the following way:

$$\left|g((Y-s)/\delta)-g((Y^{(\ell)}-s)/\delta)\right|\leqslant\left|g((Y-s)/\delta)\right|+\left|g((Y^{(\ell)}-s)/\delta)\right|\leqslant 2c_B.$$

Both estimating expressions do not depend on s, thus

$$\sup_{s\in[S_0,S_1]}E\left|g((Y-s)/\delta)-g((Y^{(\ell)}-s)/\delta)\right|\leqslant$$

$$\leqslant c_L\,E\min\left(\frac{Y-Y^{(\ell)}}{\delta},2\frac{c_B}{c_L}\right)\leqslant\left\{\text{Jensen's inequality }\left(E(x^2)\right)^{\frac{1}{2}}\leqslant E(x^2)^{\frac{1}{2}}\right\}$$

$$\leqslant E\min\left(\frac{(Y-Y^{(\ell)})^2}{\delta^2},4\frac{c_B^2}{c_L^2}\right).$$

The second argument of the minimum is a constant and can be replaced by one. We thus arrive to

$$\sup_{s\in[S_0,S_1]}\left|E\left(g((Y-s)/\delta)-g((Y^{(\ell)}-s)/\delta)\right)\right|\leqslant$$

$$\leqslant E\min\left(\frac{(Y-Y^{(\ell)})^2}{\delta^2},1\right)\leqslant\{(A4)\}\leqslant c\cdot\delta^{-\beta_4}\cdot M^{-\ell\cdot\beta_5}$$

With $\alpha_1=\beta_4/2$, $\alpha_2=\beta_5/2$, and $\alpha_3=0$ the latter yields (A3).

Often better estimates for the weak error are known, see Sections 6.3.2 and 7.1. The presence of α_1 and β_4 in these assumptions is motivated by weak and strong error estimates for SDEs or SPDEs, which often scale with some power of δ. See, however, Sections 7.1.1 and 7.1.2

Let $\|Z\|_p=(E|Z|^p)^{1/p}$ for any random variable Z and $1\leqslant p<\infty$. Typically, strong error estimates for $Y-Y^{(\ell)}$ instead of $\min(|Y-Y^{(\ell)}|,\delta)$ are available in the literature. Straightforward relations to (A3) and (A4) are provided by

$$\sup_{s\in[S_0,S_1]}\left|E\left(g((Y-s)/\delta)-g((Y^{(\ell)}-s)/\delta)\right)\right|\leqslant c_L\cdot\delta^{-1}\cdot\|Y-Y^{(\ell)}\|_1,\tag{24}$$

where c_L denotes a Lipschitz constant for g, as well as

$$E\min((Y-Y^{(\ell)})^2,\delta^2)\leqslant\min(\|Y-Y^{(\ell)}\|_2^2,\delta^2)\tag{25}$$

and

$$E\min((Y-Y^{(\ell)})^2,\delta^2)\leqslant E(\delta\cdot\min(|Y-Y^{(\ell)}|,\delta))\leqslant\min(\delta\cdot\|Y-Y^{(\ell)}\|_1,\delta^2).\tag{26}$$

In the following case of equivalence of norms the upper bound in (25) is sharp, and then we have $\beta_4=2$ in (A4), while the optimal value of β_5 is determined by the asymptotic behavior of $\|Y-Y^{(\ell)}\|_2^2$. See Sections 7.1.1 and 7.1.2 for examples.

Lemma 6.1.3 *Suppose that there exist* $c_1 > 0$ *and* $p > 2$ *such that*

$$0 < \|Y - Y^{(\ell)}\|_p \leqslant c_1 \cdot \|Y - Y^{(\ell)}\|_2$$

for all $\ell \in \mathbb{N}_0$. *Then there exists* $c_2 > 0$ *such that*

$$\mathrm{E} \min((Y - Y^{(\ell)})^2, \delta^2) \geqslant c_2 \cdot \min(\|Y - Y^{(\ell)}\|_2^2, \delta^2)$$

for all $\delta \in]0, \delta_0]$ *and* $\ell \in \mathbb{N}_0$.

Proof Put

$$Z_\ell = \frac{(Y - Y^{(\ell)})^2}{\|Y - Y^{(\ell)}\|_2^2}.$$

We show that there exists a constant $c_2 > 0$ such that

$$\mathrm{E} \min(Z_\ell, \delta) \geqslant c_2 \cdot \min(1, \delta)$$

for all $\ell \in \mathbb{N}_0$ and $\delta > 0$.

Clearly $\mathrm{E}(Z_\ell) = 1$ and $\mathrm{E}(Z_\ell^{p/2}) \leqslant c_1^p$. It follows that

$$P(\{Z_\ell > u\}) \leqslant \frac{c_1^p}{u^{p/2}}.$$

Put

$$d_\ell = P(\{Z_\ell > 1/2\}).$$

We claim that

$$d = \inf_{\ell \in \mathbb{N}_0} d_\ell > 0.$$

Assume that $d = 0$. Use

$$1 = \mathrm{E}(Z_\ell) = \int_0^\infty P(\{Z_\ell > u\})\, du \leqslant 1/2 + \int_{1/2}^\infty \min(d_\ell, c_1^p / u^{p/2})\, du$$

and dominated convergence to conclude that, for a minimizing subsequence,

$$\lim_{k \to \infty} \int_{1/2}^\infty \min(d_{\ell_k}, c_1^p / u^{p/2})\, du = 0,$$

which leads to a contradiction. Therefore

$$\mathrm{E} \min(Z_\ell, \delta) = \int_0^\delta P(\{Z_\ell > u\})\, du \geqslant \min(\delta, 1/2) \cdot d \geqslant d/2 \cdot \min(1, \delta).$$

On the other hand, if $\|Y - Y^{(\ell)}\|_2^2$ and $\|Y - Y^{(\ell)}\|_1$ are asymptotically equivalent, then (26) is preferable to (25). See Section 7.1.3 for examples.

Assumption (A4) and the Lipschitz continuity and boundedness of g immediately yield the following fact.

Lemma 6.1.4 *There exists a constant* $c > 0$ *such that*

$$\mathrm{E} \sup_{s \in [S_0, S_1]} \big(g((Y - s)/\delta) - g((Y^{(\ell)} - s)/\delta)\big)^2 \leqslant c \cdot \min(\delta^{-\beta_4} \cdot M^{-\ell \cdot \beta_5}, 1)$$

holds for all $\delta \in]0, \delta_0]$ *and* $\ell \in \mathbb{N}_0$.

Proof See the proof of Lemma 6.1.2.

6.1.3 *The Multi-level Algorithm*

The approximation of F on the interval $[S_0, S_1]$ is based on its approximation at finitely many points

$$S_0 \leqslant s_1 < \cdots < s_k \leqslant S_1, \tag{27}$$

followed by a suitable extension to $[S_0, S_1]$.

For notational convenience we put

$$g^{k,\delta}(t) = (g((t-s_1)/\delta), \ldots, g((t-s_k)/\delta)) \in \mathbb{R}^k, \qquad t \in \mathbb{R},$$

as well as $Z_i^{(0)} = Y^{(-1)} = 0$.

We choose $L_0, L_1 \in \mathbb{N}_0$ with $L_0 \leqslant L_1$ as the minimal and the maximal level, respectively, and we choose replication numbers $N_\ell \in \mathbb{N}$ for all levels $\ell = L_0, \ldots, L_1$, as well as $k \in \mathbb{N}$ and $\delta \in \,]0, \delta_0]$. The corresponding multi-level algorithm for the approximation at the points s_i is defined by

$$\mathcal{M}_{N_{L_0},\ldots,N_{L_1}}^{k,\delta,L_0,L_1} = \frac{1}{N_{L_0}} \cdot \sum_{i=1}^{N_{L_0}} g^{k,\delta}(Y_i^{(L_0)}) + \sum_{\ell=L_0+1}^{L_1} \frac{1}{N_\ell} \cdot \sum_{i=1}^{N_\ell} \left(g^{k,\delta}(Y_i^{(\ell)}) - g^{k,\delta}(Z_i^{(\ell)}) \right) \tag{28}$$

with an independent family of $\mathbb{R}^2$-valued random variables $(Y_i^{(\ell)}, Z_i^{(\ell)})$ for $\ell = L_0, \ldots, L_1$ and $i = 1, \ldots, N_\ell$ such that equality in distribution holds for $(Y_i^{(\ell)}, Z_i^{(\ell)})$ and $(Y^{(\ell)}, Y^{(\ell-1)})$.

Remark 2 *In the particular case* $L = L_0 = L_1$, *i.e., in the single-level case, we actually have a classical Monte Carlo algorithm, based on independent copies of* $Y^{(L)}$ *only. In addition to*

$$\mathcal{M}_N^{k,\delta,L,L} = \frac{1}{N} \cdot \sum_{i=1}^{N} g^{k,\delta}(Y_i^{(L)})$$

with $\delta > 0$, *we also consider the single-level algorithm without smoothing. Hence we put*

$$g^{k,0}(t) = \left(1_{]-,\infty,s_1]}(t), \ldots, 1_{]-,\infty,s_k]}(t)\right) \in \mathbb{R}^k, \qquad t \in \mathbb{R},$$

to obtain

$$\mathcal{M}_N^{k,0,L,L} = \frac{1}{N} \cdot \sum_{i=1}^{N} g^{k,0}(Y_i^{(L)}).$$

For the analysis of the single-level algorithm it suffices to assume that the simulation of the distribution of $Y^{(\ell)}$ *is possible at cost at most* $c \cdot M^\ell$ *for every* $\ell \in \mathbb{N}$, *cf. (A2). Furthermore, there is no need for a strong error estimate like (A4), and if we do not employ smoothing, then (A3) may be replaced by the following assumption. There exist a constant* $\alpha > 0$ *such that the weak error estimate*

$$\sup_{s \in [S_0, S_1]} \left| E\left(1_{]-\infty,s]}(Y) - 1_{]-\infty,s]}(Y^{(\ell)})\right) \right| \leqslant c \cdot M^{-\ell \cdot \alpha} \tag{29}$$

holds for all $\ell \in \mathbb{N}_0$. *It turns out that the analysis of single-level algorithms without smoothing is formally reduced to the case* $\delta > 0$ *if we take*

$$\alpha_1 = 0, \qquad \alpha_2 = \alpha, \qquad \alpha_3 = \alpha. \tag{30}$$

In the sequel $\|\cdot\|_\infty$ denotes the supremum norm on $C([S_0, S_1])$ and $|\cdot|_\infty$ denotes the ℓ_∞-norm on $\mathbb{R}^k$.

For the extension we take a sequence of linear mappings $Q_k^r : \mathbb{R}^k \to C([S_0, S_1])$ with the following properties for some constant $c > 0$:

(E1) For all $k \in \mathbb{N}$ and $x \in \mathbb{R}^k$ the cost for computing $Q_k^r(x)$ is bounded by $c \cdot k$.

(E2) For all $k \in \mathbb{N}$ and $x \in \mathbb{R}^k$

$$\|Q_k^r(x)\|_\infty \leqslant c \cdot |x|_\infty.$$

(E3) For all $k \in \mathbb{N}$

$$\|F - Q_k^r(F(s_1), \ldots, F(s_k))\|_\infty \leqslant c \cdot k^{-(r+1)}.$$

These properties are achieved, e.g., by piecewise polynomial interpolation with degree $\max(r, 1)$ at equidistant points $s_i = S_0 + (i-1) \cdot (S_1 - S_0)/(k-1)$ with $k \geqslant 2$.

We employ $Q_k^r(\mathcal{M})$ with $\mathcal{M} = \mathcal{M}^{k,\delta,L_0,L_1}_{N_{L_0},\ldots,N_{L_1}}$ as a randomized algorithm for the approximation of F on $[S_0, S_1]$. Observe that $\mathcal{M}$ is square-integrable, since g is bounded, so that (E2) yields $E\,\|Q_k^r(\mathcal{M})\|_\infty^2 < \infty$. The error of $Q_k^r(\mathcal{M})$ is defined by

$$\operatorname{error}(Q_k^r(\mathcal{M})) = \left(E\,\|F - Q_k^r(\mathcal{M})\|_\infty^2\right)^{1/2}.$$

Since the error is based on the supremum norm, $\operatorname{error}(Q_k^r(\mathcal{M}))$ does not increase if we replace $Q_k^r(x)$ by $s \mapsto \sup_{u \in [S_0, s]} (Q_k^r(x))(u)$ to get a non-decreasing approximation on $[S_0, S_1]$.

The variance of any square-integrable $\mathbb{R}^k$-valued random variable $\mathcal{M}$ is defined by

$$\operatorname{Var}(\mathcal{M}) = E\,|\mathcal{M} - E(\mathcal{M})|_\infty^2,$$

and

$$E\,|x - \mathcal{M}|_\infty^2 \leqslant 2 \cdot (|x - E(\mathcal{M})|_\infty^2 + \operatorname{Var}(\mathcal{M}))$$

holds for $x \in \mathbb{R}^k$. Furthermore,

$$\operatorname{Var}(\mathcal{M}) \leqslant 4 \cdot E(|\mathcal{M}|_\infty^2).$$

The Bienaymé formula for real-valued random variables turns into the inequality

$$\operatorname{Var}(\mathcal{M}) \leqslant c \cdot \log k \cdot \sum_{i=1}^{n} \operatorname{Var}(\mathcal{M}_i), \tag{31}$$

if $\mathcal{M} = \sum_{i=1}^{n} \mathcal{M}_i$ with independent square-integrable random variables $\mathcal{M}_i$ taking values in $\mathbb{R}^k$. Here c is a universal constant. In the context of multi-level algorithms this is exploited in [34].

We put

$$q = \min\left(\frac{r+1+\alpha_1}{\alpha_2}, \frac{r+1}{\alpha_3}\right). \tag{32}$$

Theorem 6.1.5 *The following order, with* $\eta = 1$, *is achieved by algorithms* $Q_k^r(\mathcal{M}^{k,\delta,L_0,L_1}_{N_{L_0},\ldots,N_{L_1}})$ *with suitably chosen parameters:*

$$q \leqslant \max(1, \beta_4/\beta_5) \quad \Rightarrow \quad \gamma = 2 + \frac{\max(1, q)}{r+1}, \tag{33}$$

$$q > \max(1, \beta_4/\beta_5) \wedge \beta_5 > 1 \quad \Rightarrow \quad \gamma = 2 + \frac{\max(1, \beta_4/\beta_5)}{r+1}, \tag{34}$$

$$q > 1 > \beta_4 \wedge \beta_5 = 1 \quad \Rightarrow \quad \gamma = 2 + \frac{1}{r+1}, \tag{35}$$

$$q > \max(1, \beta_4/\beta_5) \wedge \beta_5 < 1 \quad \Rightarrow \quad \gamma = 2 + \frac{\max(1, \beta_4 + (1-\beta_5) \cdot q)}{r+1}. \tag{36}$$

Moreover, with $\eta = 3$,

$$q > \beta_4 \geqslant 1 \wedge \beta_5 = 1 \quad \Rightarrow \quad \gamma = 2 + \frac{\beta_4}{r+1}. \tag{37}$$

Proof Let $\mathcal{M}$ denote any square-integrable random variable with values in $\mathbb{R}^k$. For the error of $Q^r_k(\mathcal{M})$ we have

$$\begin{aligned}
\operatorname{error}(Q^r_k(\mathcal{M})) &\leqslant \|F - Q^r_k(F(s_1), \dots, F(s_k))\|_\infty + \\
&\quad + \left(E\,\|Q^r_k((F(s_1), \dots F(s_k)) - \mathcal{M})\|^2_\infty\right)^{1/2} \\
&\leqslant c \cdot \left(k^{-(r+1)} + \left(E\,|(F(s_1), \dots F(s_k)) - \mathcal{M}|^2_\infty\right)^{1/2}\right) \\
&\leqslant 2c \cdot \left(k^{-2(r+1)} + |(F(s_1), \dots F(s_k)) - E(\mathcal{M})|^2_\infty + \operatorname{Var}(\mathcal{M})\right)^{1/2}
\end{aligned}$$

with a constant $c > 0$ according to (E2) and (E3).

Now we consider the algorithm $\mathcal{M} = \mathcal{M}^{k,\delta,L_0,L_1}_{N_{L_0},\dots,N_{L_1}}$ with $\delta > 0$. We write $a \preceq b$ if there exists a constant $c > 0$ that does not depend on the parameters k, δ, L_0, L_1, $N_{L_0}, \dots, N_{L_1}$ such that $a \leqslant c \cdot b$. Moreover, $a \succeq b$ means $b \preceq a$, and $a \asymp b$ stands for $a \preceq b$ and $a \succeq b$.

Note that $E(\mathcal{M}) = E(g^{k,\delta}(Y^{(L_1)}))$. Hence the bias term is estimated by

$$\begin{aligned}
|(F(s_1), \dots, F(s_k)) - E(\mathcal{M})|_\infty &= \sup_{i=1,\dots,k} |F(s_i) - E(g((Y^{(L_1)} - s_i)/\delta))| \\
&\preceq \delta^{r+1} + \min\left(\delta^{-\alpha_1} \cdot M^{-L_1 \cdot \alpha_2}, M^{-L_1 \cdot \alpha_3}\right),
\end{aligned}$$

see Lemma 6.1.1 and (A3).

The variance of $\mathcal{M}$ is estimated as follows. From (31) we obtain

$$\operatorname{Var}(\mathcal{M}) \preceq \log k \cdot \left(\frac{\operatorname{Var}\left(g^{k,\delta}(Y^{(L_0)})\right)}{N_{L_0}} + \sum_{\ell=L_0+1}^{L_1} \frac{\operatorname{Var}\left(g^{k,\delta}(Y^{(\ell)}) - g^{k,\delta}(Y^{(\ell-1)})\right)}{N_\ell} \right).$$

Moreover,

$$\begin{aligned}
&\operatorname{Var}\left(g^{k,\delta}(Y^{(\ell)}) - g^{k,\delta}(Y^{(\ell-1)})\right) \leqslant \\
&\qquad \leqslant 4 \cdot E \sup_{i=1,\dots,k} \left(g((Y^{(\ell)} - s_i)/\delta) - g((Y^{(\ell-1)} - s_i)/\delta)\right)^2 \preceq \\
&\qquad\qquad \preceq \min(\delta^{-\beta_4} \cdot M^{-\ell \cdot \beta_5}, 1)
\end{aligned}$$

for $\ell = L_0 + 1, \dots, L_1$, see Lemma 6.1.4, and

$$\operatorname{Var}(g^{k,\delta}(Y^{(L_0)})) \preceq 1,$$

since g is bounded. Therefore

$$\operatorname{Var}(\mathcal{M}) \preceq \log k \cdot \left(\frac{1}{N_{L_0}} + \sum_{\ell=L_0+1}^{L_1} \frac{\min(\delta^{-\beta_4} \cdot M^{-\ell \cdot \beta_5}, 1)}{N_\ell} \right).$$

Combining these estimates we finally get

$$\begin{aligned}
\operatorname{error}^2(Q^r_k(\mathcal{M})) &\preceq k^{-2(r+1)} + \delta^{2(r+1)} + \min\left(\delta^{-2\alpha_1} \cdot M^{-L_1 \cdot 2\alpha_2}, M^{-L_1 \cdot 2\alpha_3}\right) \\
&\quad + \log k \cdot \left(\frac{1}{N_{L_0}} + \sum_{\ell=L_0+1}^{L_1} \frac{\min(\delta^{-\beta_4} \cdot M^{-\ell \cdot \beta_5}, 1)}{N_\ell} \right).
\end{aligned} \tag{38}$$

Now we analyze the computational cost of the algorithm $\mathcal{M}$. For $\ell = L_0, \dots, L_1$ and $i = 1, \dots, N_\ell$ the cost of computing $g^{k,\delta}(Y_i^{(\ell)})$ or $g^{k,\delta}(Y_i^{(\ell)}) - g^{k,\delta}(Z_i^{(\ell)})$ is bounded by $M^\ell + k$, up to a constant, see (S1) and (A2). Use (E1) to obtain

$$\operatorname{cost}(Q_k^r(\mathcal{M})) \preceq c(k, L_0, L_1, N_{L_0}, \dots, N_{L_1}) \tag{39}$$

with

$$c(k, L_0, L_1, N_{L_0}, \dots, N_{L_1}) = \sum_{\ell=L_0}^{L_1} N_\ell \cdot (M^\ell + k). \tag{40}$$

Note that for every k the cost per replication is essentially constant on all levels $\ell \leqslant L^*$, where

$$L^* = \log_M k. \tag{41}$$

Observe that the estimates (38) and (39) are valid, too, for single-level algorithms without smoothing, i.e., for $L_0 = L_1$ and $\delta = 0$, if we formally define the parameters α_i by (30), which leads to $q = (r+1)/\alpha$.

We determine parameters of the algorithm $Q_k^r(\mathcal{M})$ such that an error of about $\varepsilon \in \left]0, \min(1, \delta_0^{r+1})\right[$ is achieved at a small cost. More precisely, we minimize the upper bound (39) for the cost, subject to the constraint that the upper bound (38) for the squared error is at most ε^2, up to multiplicative constants for both quantities.

First of all we consider the case $\delta > 0$, and we choose

$$\delta = \varepsilon^{1/(r+1)} \tag{42}$$

and, up to integer rounding,

$$k = \varepsilon^{-1/(r+1)} \tag{43}$$

and

$$N_{L_0} = \varepsilon^{-2} \cdot \log_M \varepsilon^{-1}. \tag{44}$$

This yields

$$\operatorname{error}^2(Q_k^r(\mathcal{M})) \preceq \varepsilon^2 + a^2(L_1) + \log \varepsilon^{-1} \cdot \sum_{\ell=L_0+1}^{L_1} \frac{\min(\delta^{-\beta_4} \cdot M^{-\ell \cdot \beta_5}, 1)}{N_\ell}$$

with

$$a(L_1) = \min\left(\delta^{-\alpha_1} \cdot M^{-L_1 \cdot \alpha_2}, M^{-L_1 \cdot \alpha_3}\right). \tag{45}$$

Furthermore,

$$L^* = \frac{1}{r+1} \cdot \log_M \varepsilon^{-1}. \tag{46}$$

Due to the dependence of (39) on k and the decay of $a(L_1)$ and $\min(\delta^{-\beta_4} \cdot M^{-\ell \cdot \beta_5}, 1)$ as functions of L_1 and ℓ, respectively, it suffices to study

$$L_0 \geqslant L^*. \tag{47}$$

Moreover, $a(L_1) \leqslant \varepsilon$ requires $L_1 \geqslant q \cdot L^*$. Consequently, we choose

$$L_1 = \max(1, q) \cdot L^*, \tag{48}$$

up to integer rounding.

For a single-level algorithm with smoothing, i.e., for $L_0 = L_1$ and $\delta > 0$, all parameters have thus been determined, and we obtain $\operatorname{error}(Q_k^r(\mathcal{M})) \preceq \varepsilon$ as well as

$$c(k, L_1, L_1, N_{L_1}) \asymp \varepsilon^{-2} \cdot \log \varepsilon^{-1} \cdot M^{L^*} = \varepsilon^{-2-1/(r+1)} \cdot \log \varepsilon^{-1} \tag{49}$$

if $q \leqslant 1$, and

$$c(k, L_1, L_1, N_{L_1}) \asymp \varepsilon^{-2} \cdot \log \varepsilon^{-1} \cdot M^{q \cdot L^*} = \varepsilon^{-2-q/(r+1)} \cdot \log \varepsilon^{-1}, \tag{50}$$

if $q > 1$. For a single-level algorithm without smoothing we obtain the same result.

For a proper multi-level algorithm with

$$L^* \leqslant L_0 < L_1$$

we obtain

$$\operatorname{error}^2(Q_k^r(\mathcal{M})) \preceq \varepsilon^2 + \log \varepsilon^{-1} \cdot \sum_{\ell=L_0+1}^{L_1} \frac{v_\ell}{N_\ell}$$

with

$$v_\ell = \min(M^{L^* \cdot \beta_4} \cdot M^{-\ell \cdot \beta_5}, 1)$$

as well as

$$c(k, L_0, L_1, N_{L_0}, \dots, N_{L_1}) \asymp \varepsilon^{-2} \cdot \log \varepsilon^{-1} \cdot M^{L_0} + \sum_{\ell=L_0+1}^{L_1} N_\ell \cdot M^\ell.$$

Observing

$$c(k, L_0, L_1, N_{L_0}, \dots, N_{L_1}) \succeq \varepsilon^{-2} \cdot \log \varepsilon^{-1} \cdot M^{L^*}$$

and (49), we get (33) in the case $q \leqslant 1$ already by single-level algorithms.

To establish the theorem in the case

$$q > 1$$

we fix L_0 for the moment, and we minimize

$$h(L_0, N_{L_0+1}, \dots, N_{L_1}) = \varepsilon^{-2} \cdot \log \varepsilon^{-1} \cdot M^{L_0} + \sum_{\ell=L_0+1}^{L_1} N_\ell \cdot M^\ell$$

subject to

$$\sum_{\ell=L_0+1}^{L_1} \frac{v_\ell}{N_\ell} \leqslant \varepsilon^2 / \log \varepsilon^{-1}.$$

A Lagrange multiplier leads to

$$N_\ell = \varepsilon^{-2} \cdot \log \varepsilon^{-1} \cdot G(L_0) \cdot \left(v_\ell \cdot M^{-\ell}\right)^{1/2}, \tag{51}$$

up to integer rounding, which satisfies the constraint with

$$G(L_0) = \sum_{\ell=L_0+1}^{L_1} \left(v_\ell \cdot M^\ell\right)^{1/2} = \sum_{\ell=L_0+1}^{L_1} \left(\min(M^{L^* \cdot \beta_4} \cdot M^{-\ell \cdot \beta_5}, 1) \cdot M^\ell\right)^{1/2}.$$

Moreover, this choice of $N_{L_0+1}, \dots, N_{L_1}$ yields

$$h(L_0, N_{L_0+1}, \dots, N_{L_1}) = \varepsilon^{-2} \cdot \log \varepsilon^{-1} \cdot \left(M^{L_0} + G^2(L_0)\right). \tag{52}$$

Put

$$L^\dagger = \frac{\beta_4}{\beta_5} \cdot L^*.$$

Consider the case

$$1 < q \leqslant \beta_4 / \beta_5.$$

Then we have $L_1 \leqslant L^\dagger$, and therefore

$$M^{L_0} + G^2(L_0) = M^{L_0} + \left(\sum_{\ell=L_0+1}^{L_1} M^{\ell/2} \right)^2 \asymp M^{L_0} + M^{L_1} \asymp M^{L^* \cdot q}.$$

Observing (50) we get (33) in the present case already by single-level algorithms.

From now on we consider the case

$$q > \max(1, \beta_4/\beta_5).$$

Suppose that $L_0 < L^\dagger$, which requires $\beta_4/\beta_5 > 1$ to hold. Then we get

$$\begin{aligned} M^{L_0} + G^2(L_0) &\asymp M^{L_0} + \left(\sum_{\ell=L_0+1}^{L^\dagger} M^{\ell/2} \right)^2 + M^{L^* \cdot \beta_4} \cdot \left(\sum_{\ell=L^\dagger+1}^{L_1} M^{\ell \cdot (1-\beta_5)/2} \right)^2 \\ &\asymp M^{L^\dagger} + M^{L^* \cdot \beta_4} \cdot \left(\sum_{\ell=L^\dagger+1}^{L_1} M^{\ell \cdot (1-\beta_5)/2} \right)^2 \asymp M^{L^\dagger} + G^2(L^\dagger). \end{aligned}$$

It therefore suffices to study the case

$$L_0 \geqslant L^\dagger,$$

where we have

$$M^{L_0} + G^2(L_0) = M^{L_0} + M^{L^* \cdot \beta_4} \cdot \left(\sum_{\ell=L_0+1}^{L_1} M^{\ell \cdot (1-\beta_5)/2} \right)^2 .$$

If $\beta_5 = 1$ then

$$M^{L_0} + G^2(L_0) \asymp M^{L_0} + M^{L^* \cdot \beta_4} \cdot (L_1 - L_0)^2.$$

If $\beta_5 > 1$ then

$$M^{L_0} + G^2(L_0) \asymp M^{L_0} + M^{L^* \cdot \beta_4} \cdot M^{L_0 \cdot (1-\beta_5)} \asymp M^{L_0}.$$

If $\beta_5 < 1$ then

$$M^{L_0} + G^2(L_0) \asymp M^{L_0} + M^{L^* \cdot \beta_4} \cdot M^{L_1 \cdot (1-\beta_5)}.$$

Hence we choose

$$L_0 = \max(1, \beta_4/\beta_5) \cdot L^* \tag{53}$$

in all these cases. Hereby we obtain

$$M^{L_0} + G^2(L_0) \asymp M^{L^* \cdot \max(1,\beta_4/\beta_5)} \cdot \begin{cases} (L^*)^2, & \text{if } \beta_5 = 1 \text{ and } \beta_4 \geqslant 1, \\ 1, & \text{if } \beta_5 > 1 \text{ or } \beta_5 = 1 \text{ and } \beta_4 < 1, \end{cases}$$

as well as

$$M^{L_0} + G^2(L_0) \asymp M^{\max(1,\beta_4/\beta_5,\beta_4+(1-\beta_5)\cdot q)\cdot L^*}$$

if $\beta_5 < 1$. In any case these estimates are superior to $M^{L^* \cdot q}$, cf. (50). Use (52) and $M^{L^*} = \varepsilon^{-1/(r+1)}$ to derive (34)–(37).

Remark 3 *Theorem 6.1.5 is based on the upper bounds* (38) *and* (39) *for the error and the cost, respectively, of the algorithms* $Q_k^r(\mathcal{M}^{k,\delta,L_0,L_1}_{N_{L_0},\dots,N_{L_1}})$. *The parameters that we have determined in the proof of Theorem 6.1.5 are optimal in the following sense: they minimize the upper bound for the cost, subject to the constraint that the upper bound for the error is at most* ε*, up to multiplicative constants for both quantities.*

Obviously, this optimality holds true for the choice of δ, k, N_{L_0}*, and* L_1 *according to* (42), (43), (44)*, and* (48)*. Moreover, the constraint* (47) *is without loss of generality, so that the minimal level* L_0 *slowly increases with decreasing* ε.

This completes, in particular, the optimization of the parameters of single-level algorithms, where $L_0 = L_1$*. For proper multi-level algorithms the optimal values of* N_ℓ *for* $\ell = L_0 + 1, \ldots, L_1$ *are presented in* (51) *and the optimal value of* L_0 *is presented in* (53)*, if* $q > \max(1, \beta_4/\beta_5)$*. It is straightforward to verify*

$$N_\ell = \varepsilon^{-2-\beta_4/(r+1)} \cdot \log \varepsilon^{-1} \cdot M^{-\ell \cdot (1+\beta_5)/2} \cdot \mathcal{K}, \tag{54}$$

where

$$\mathcal{K} = \begin{cases} L^*, & \text{if } \beta_5 = 1, \\ M^{L^* \cdot \max(1,\beta_4/\beta_5) \cdot (1-\beta_5)/2} & \text{if } \beta_5 > 1, \\ M^{L^* \cdot q \cdot (1-\beta_5)/2}, & \text{if } \beta_5 < 1. \end{cases}$$

Furthermore, we have carried out the comparison of multi-level and single-level algorithms in the proof of Theorem 6.1.5. This comparison, too, is merely based on the upper bounds for the error and the cost, and on the assumption that $\alpha = \alpha_3$*. In this sense we have a superiority of proper multi-level algorithms over single-level algorithms if and only if*

$$q > \max(1, \beta_4/\beta_5), \tag{55}$$

which corresponds to (34)–(37) *in Theorem 6.1.5. The lack of superiority, which is present in* (33) *in Theorem 6.1.5, is explained as follows. For* $q \leqslant 1$ *the maximal level can be chosen so small that the computational cost on all levels is dominated by the number* k *of discretization points that is needed to achieve a good approximation of* F *even from exact data* $F(s_1), \ldots, F(s_k)$*. For* $1 < q \leqslant \beta_4/\beta_5$ *the negative impact of smoothing on the variances leads to variances* $\min(\delta^{-\beta_4} \cdot M^{-\ell \cdot \beta_5}, 1)$ *of order one on all levels* $\ell = L_0 + 1, \ldots, L_1$.

Single-level algorithms with smoothing are never inferior to single-level algorithms without smoothing, and they are superior if and only if

$$\frac{r+1}{\alpha_3} > \max(1, q). \tag{56}$$

For large values of r *the latter holds true if and only if* $\alpha_2 > \alpha_3$*; see Section 7.1.3 for an example.*

6.2 APPROXIMATION OF DENSITIES ON COMPACT INTERVALS

In this section we study the approximation of the density ρ of Y on an interval $[S_0, S_1]$ for some fixed $S_0 < S_1$. The construction and analysis closely follows the approach from Section 6.1.

6.2.1 *Smoothing*

We employ assumption (A1) with $r \geqslant 1$, and $g : \mathbb{R} \to \mathbb{R}$ is assumed to satisfy the properties (S1) and (S2), while (S3) and (S4) are replaced by:

(S5) $g(s) = 0$ if $|s| > 1$.

(S6) $\int_{-1}^{1} g(s)\, ds = 1$ and $\int_{-1}^{1} s^j \cdot g(s)\, ds = 0$ for $j = 1, \ldots, r-1$.

Obviously, g is bounded due to (S2) and (S5). Moreover, if $g \in C^1(\mathbb{R})$ satisfies (S3) and (S4) and g' is Lipschitz continuous, then $-g'$, instead of g, satisfies (S5) and (S6). In kernel density estimation, a function g with integral one and vanishing moments up to order $r-1$ is called a kernel of order at least r.

Remark 4 *We modify the construction from Remark 1 as follows. There exists a uniquely determined polynomial* p *of degree at most* $r+1$ *such that*

$$\int_{-1}^{1} p(s)\,ds = 1$$

and

$$\int_{-1}^{1} s^j \cdot p(s)\,ds = 0, \qquad j = 0, \dots, r-1,$$

as well as $p(1) = p(-1) = 0$*. Extend* p *by zero to obtain* g *with the properties as claimed. Since* g *is an even function, the same function* g *arises in this way for* r *and* $r+1$*, if* r *is odd.*

We have the following estimate for the bias that is induced by smoothing with parameter δ, i.e., by approximation of the Dirac functional at s by $1/\delta \cdot g((\cdot - s)/\delta)$.

Lemma 6.2.1 *There exists a constant* $c > 0$ *such that*

$$\sup_{s\in[S_0,S_1]} |\rho(s) - 1/\delta \cdot E(g((Y-s)/\delta))| \leqslant c \cdot \delta^r$$

holds for all $\delta \in \,]0, \delta_0]$.

Proof Due to

$$\rho(s) = \int_{-1}^{1} \rho(s)g(u)\,du, \text{ and } \frac{1}{\delta}E(g((Y-s)/\delta)) = \frac{1}{\delta}\int_{-\delta}^{\delta} g((x-s)/\delta)\rho(x)\,dx =$$
$$= \left\{\text{change of variable } u = \frac{x-s}{\delta}, x = u\delta + s\right\} = \int_{-1}^{1} g(u)\rho(\delta u + s)\,du$$

the following is true:

$$\rho(s) - 1/\delta \cdot E(g((Y-s)/\delta)) = \int_{-1}^{1} g(u)\cdot(\rho(s) - \rho(\delta u + s))\,du = \{\text{ expand } \rho(\delta u + s)\} =$$
$$= \int_{-1}^{1} g(u)\left(-\sum_{i=1}^{r-1}\frac{\rho^{(i)}(s)}{i!}(\delta u)^i + O((\delta u)^r)\right)du = \{\text{properties of } g \text{ from Remark 4}\} =$$
$$= \int_{-1}^{1} g(u)O(\delta^r)u^r\,du \leqslant c\delta^r.$$

6.2.2 *Assumptions on Weak and Strong Convergence*

We employ the assumptions (A2)–(A4) from Section 6.1.2 with possibly different values of α_i in the weak error estimate (A3). We make use of Lemma 6.1.4, and we refer to Section 7.1 for specific examples with corresponding values of α_i.

6.2.3 *The Multi-level Algorithm*

The definition (28) of the algorithms $\mathcal{M}^{k,\delta,L_0,L_1}_{N_{L_0},\dots,N_{L_1}}$ also applies for the approximation of densities, except for $g^{k,\delta}$, which is now defined by

$$g^{k,\delta}(t) = \frac{1}{\delta}\cdot(g((t-s_1)/\delta), \dots, g((t-s_k)/\delta)) \in \mathbb{R}^k, \quad t \in \mathbb{R}.$$

In the present setting we have $\delta > 0$ also for single-level algorithms.

Hereby we obtain approximations to ρ at the points (27), which are extended to functions on $[S_0, S_1]$ by means of linear mappings $Q^r_k : \mathbb{R}^k \to C([S_0, S_1])$. We assume that (E1) and (E2) are satisfied, but instead of (E3) the following property is assumed to hold with some constant $c > 0$:

(E4) For all $k \in \mathbb{N}$

$$\|\rho - Q_k^r(\rho(s_1), \dots, \rho(s_k))\|_\infty \leqslant c \cdot k^{-r}.$$

As before, piecewise polynomial interpolation at equidistant points, now of degree $\max(r-1, 1)$, might be used for this purpose.

We employ $Q_k^r(\mathcal{M})$ with $\mathcal{M} = \mathcal{M}^{k,\delta,L_0,L_1}_{N_{L_0},\dots,N_{L_1}}$ as a randomized algorithm for the approximation of ρ on $[S_0, S_1]$, and the error of $Q_k^r(\mathcal{M})$ is defined by

$$\mathrm{error}(Q_k^r(\mathcal{M})) = \left(\mathrm{E}\,\|\rho - Q_k^r(\mathcal{M})\|_\infty^2\right)^{1/2}.$$

Clearly the error does not increase if we replace $Q_k^r(x)$ by $\max(Q_k^r(x), 0)$.

Recall the definition of q from (32).

Theorem 6.2.2 *The following order, with* $\eta = 1$, *is achieved by algorithms* $Q_k^r(\mathcal{M}^{k,\delta,L_0,L_1}_{N_{L_0},\dots,N_{L_1}})$ *with suitably chosen parameters:*

$$q \leqslant \max(1, \beta_4/\beta_5) \quad \Rightarrow \quad \gamma = 2 + \frac{\max(1,q)+2}{r}, \tag{57}$$

$$q > \max(1, \beta_4/\beta_5) \wedge \beta_5 > 1 \quad \Rightarrow \quad \gamma = 2 + \frac{\max(1,\beta_4/\beta_5)+2}{r}, \tag{58}$$

$$q > 1 > \beta_4 \wedge \beta_5 = 1 \quad \Rightarrow \quad \gamma = 2 + \frac{3}{r}, \tag{59}$$

$$q > \max(1, \beta_4/\beta_5) \wedge \beta_5 < 1 \quad \Rightarrow \quad \gamma = 2 + \frac{\max(1,\beta_4 + (1-\beta_5)\cdot q)+2}{r}. \tag{60}$$

Moreover, with $\eta = 3$,

$$q > \beta_4 \geqslant 1 \wedge \beta_5 = 1 \quad \Rightarrow \quad \gamma = 2 + \frac{\beta_4 + 2}{r}. \tag{61}$$

Proof We mimic the proof of Theorem 6.1.5. We use (A3), (E2) and (E4), Lemma 6.1.4 and Lemma 6.2.1, and the boundedness of g to obtain

$$\begin{aligned}\mathrm{error}^2(Q_k^r(\mathcal{M})) \preceq{}& k^{-2r} + \delta^{2r} + 1/\delta^2 \cdot \min\left(\delta^{-2\alpha_1}\cdot M^{-L_1\cdot 2\alpha_2}, M^{-L_1\cdot 2\alpha_3}\right) \\ &+ \log k/\delta^2 \cdot \left(\frac{1}{N_{L_0}} + \sum_{\ell=L_0+1}^{L_1} \frac{\min(\delta^{-\beta_4}\cdot M^{-\ell\cdot\beta_5}, 1)}{N_\ell}\right),\end{aligned} \tag{62}$$

where $\mathcal{M} = \mathcal{M}^{k,\delta,L_0,L_1}_{N_{L_0},\dots,N_{L_1}}$. The upper bound (39) for the computational cost is also valid in the present case. We minimize (39), subject to the constraint that the upper bound (62) for the squared error is at most ε^2, up to multiplicative constants for both quantities.

Put

$$\tilde{\varepsilon} = \varepsilon^{1+1/r}.$$

First of all we choose

$$\delta = \varepsilon^{1/r} = \tilde{\varepsilon}^{1/(r+1)} \tag{63}$$

and, up to integer rounding,

$$k = \varepsilon^{-1/r} = \tilde{\varepsilon}^{-1/(r+1)} \tag{64}$$

and

$$N_{L_0} = \varepsilon^{-2-2/r}\cdot \log_M \varepsilon^{-1} \asymp \tilde{\varepsilon}^{-2}\cdot\log_M \tilde{\varepsilon}^{-1}. \tag{65}$$

This yields

$$\operatorname{error}^2(Q_k^r(\mathcal{M})) \preceq$$
$$\preceq \varepsilon^2 + 1/\delta^2 \cdot \left(a^2(L_1) + \log \tilde{\varepsilon}^{-1} \cdot \sum_{\ell=L_0+1}^{L_1} \frac{\min(\delta^{-\beta_4} \cdot M^{-\ell \cdot \beta_5}, 1)}{N_\ell} \right),$$

where $a(L_1)$ is given by (45). Furthermore,

$$L^* = \frac{1}{r} \cdot \log_M \varepsilon^{-1} = \frac{1}{r+1} \cdot \log_M \tilde{\varepsilon}^{-1}, \tag{66}$$

see (41), and it suffices to study $L_0 \geqslant L^*$.

Since $\delta \cdot \varepsilon = \tilde{\varepsilon}$, the proof of Theorem 6.1.5 is applicable with ε being replaced by $\tilde{\varepsilon}$. We obtain the same values for η, but γ must be replaced by $\gamma \cdot (1 + 1/r)$.

Remark 5 *The following comments on optimality etc. are meant in the sense of Remark 3. We have a superiority of proper multi-level algorithms over single-level algorithms if and only if (55) holds true. Moreover, the optimal values of* δ, k, *and* N_{L_0}, *and* L_1 *are given by (63), (64), (65), and*

$$L_1 = \frac{\max(1, q)}{r} \cdot \log_M \varepsilon^{-1},$$

see (48). In particular, this completes the optimization of the parameters of single-level algorithms, where $L_0 = L_1$.

Suppose that $q > \max(1, \beta_4/\beta_5)$, *so that we consider proper multi-level algorithms. The optimal value of* L_0 *is given by*

$$L_0 = \frac{\max(1, \beta_4/\beta_5)}{r} \cdot \log_M \varepsilon^{-1},$$

see (53), The optimality of

$$N_\ell = \varepsilon^{-2-(\beta_4+2)/r} \cdot \log \varepsilon^{-1} \cdot M^{-\ell \cdot (1+\beta_5)/2} \cdot \mathcal{K},$$

where

$$\mathcal{K} = \begin{cases} L^*, & \text{if } \beta_5 = 1, \\ M^{L^* \cdot \max(1, \beta_4/\beta_5) \cdot (1-\beta_5)/2} & \text{if } \beta_5 > 1, \\ M^{L^* \cdot q \cdot (1-\beta_5)/2}, & \text{if } \beta_5 < 1. \end{cases}$$

for $\ell = L_0 + 1, \ldots, L_1$, *with* L^* *given by (66), is derived from (51) in a straightforward way.*

6.3 APPROXIMATION OF DISTRIBUTION FUNCTIONS AT A SINGLE POINT

Now we study the approximation of the distribution function F of Y at a single fixed point $s \in [S_0, S_1]$.

6.3.1 *Smoothing*

We employ assumption (A1) and the smoothing approach from Section 6.1.1, which involves the assumptions (S1)–(S4). In particular, we make use of Lemma 6.1.1.

6.3.2 *Assumptions on Weak and Strong Convergence*

We consider the setting from Section 6.1.2, and we assume (A2) and (A3) while, instead of (A4), the following property is assumed to hold with a constant $c > 0$:

(A5) There exist constants $\beta_1 \geqslant 0$ and $\beta_2 > \beta_3 \geqslant 0$ such that the strong error estimate

$$\sup_{s\in[S_0,S_1]} \mathrm{E}\left(g((Y-s)/\delta) - g((Y^{(\ell)}-s)/\delta)\right)^2 \leqslant \leqslant c \cdot \min\left(\delta^{-\beta_1} \cdot M^{-\ell\cdot\beta_2}, M^{-\ell\cdot\beta_3}\right)$$

holds for all $\delta \in \,]0,\delta_0]$ and $\ell \in \mathbb{N}_0$.

See Section 7.1 for specific applications and approximations $Y^{(\ell)}$ with corresponding values of the parameters β_i.

We use different assumptions on the strong error for approximation of F on compact intervals and at a single point, namely (A4) with Lemma 6.1.4 as an immediate consequence in the first case and (A5) in the second case. Clearly, (A4) implies (A5) for every bounded and Lipschitz continuous function g with

$$\beta_1 = \beta_4, \qquad \beta_2 = \beta_5, \qquad \beta_3 = 0, \tag{67}$$

which is used in Section 7.1.3, but better values of β_1, β_2, and β_3 may be available. See Section 7.1 for examples where $\beta_1 < \beta_4$ and $\beta_3 > 0$. Note that (A5) corresponds directly to the weak error estimate (A3), and it yields the latter for every bounded and measurable function g with

$$\alpha_i = \beta_i/2 \tag{68}$$

for $i = 1,2,3$. See Section 7.1 for applications.

Strong error estimates for $Y - Y^{(\ell)}$ or $1_{]-\infty,s]}(Y) - 1_{]-\infty,s]}(Y^{(\ell)})$ may be used to establish (A5) and (A3). From the Lipschitz continuity of g we immediately get (A5) with $\beta_1 = 2$ and $\beta_3 = 0$, while the value of β_2 is determined by the asymptotic behavior of $\|Y - Y^{(\ell)}\|_2^2$. A refined analysis, which merely requires Y to have a bounded density, yields the following results, which are applicable under the assumptions (S2) and (S3) or (S2) and (S5) on g.

Lemma 6.3.1 (Avikainen (2009)) *There exists a constant* $c > 0$ *such that*

$$\sup_{s\in[S_0,S_1]} \|g((Y-s)/\delta) - g((Y^{(\ell)}-s)/\delta)\|_q^q \leqslant \leqslant c^q \cdot \sup_{s\in[S_0-\delta_0,S_1+\delta_0]} \|1_{]-\infty,s]}(Y) - 1_{]-\infty,s]}(Y^{(\ell)})\|_1$$

and

$$\sup_{s\in[S_0-\delta,S_1+\delta]} \|1_{]-\infty,s]}(Y) - 1_{]-\infty,s]}(Y^{(\ell)})\|_1 \leqslant c \cdot \|Y - Y^{(\ell)}\|_p^{p/(p+1)}$$

holds for all $p, q \geqslant 1$, $\delta \in \,]0,\delta_0]$, *and* $\ell \in \mathbb{N}_0$.

Proof See in [5] at p.387 for the proof of the first estimate and Lemma 3.4 for the second estimate.

Lemma 6.3.2 *For every* $1 \leqslant q \leqslant p < \infty$ *there exists a constant* $c > 0$ *such that*

$$\sup_{s\in[S_0,S_1]} \|g((Y-s)/\delta) - g((Y^{(\ell)}-s)/\delta)\|_q^q \leqslant c \cdot \delta^{1-q-q/p} \cdot \|Y - Y^{(\ell)}\|_p^q$$

holds for all $\delta \in \,]0,\delta_0/2]$ *and* $\ell \in \mathbb{N}_0$.

Proof Put

$$\Delta = g((Y-s)/\delta) - g((Y^{(\ell)}-s)/\delta).$$

In the sequel, we adopt the notation $\preceq$ from the proof of Theorem 6.1.5, where now the hidden constant must not depend on δ, ℓ or s.

Because of assumption (A1), the density ρ of Y is bounded on $[S_0 - \delta_0, S_1 + \delta_0]$. By Lemma 6.3.1,

$$\mathrm{E}\,\Delta^q \preceq \|Y - Y^{(\ell)}\|_p^{p/(p+1)},$$

so all that remains is to establish

$$\mathrm{E}\,\Delta^q \preceq \delta^{1-q-q/p} \cdot \|Y - Y^{(\ell)}\|_p^q$$

in the case $\delta^{1-q-q/p} \cdot \|Y - Y^{(\ell)}\|_p^q \leqslant \|Y - Y^{(\ell)}\|_p^{p/(p+1)}$, i.e., for

$$\|Y - Y^{(\ell)}\|_p \leqslant \delta^{1+1/p}. \tag{69}$$

If $|Y - s| > 2\delta$ and $|Y - Y^{(\ell)}| < \delta$, then $|Y^{(\ell)} - s| > \delta$ and hence $\Delta = 0$ follows, since g is constant on $]-\infty, -1[$ as well as on $]1, \infty[$. Accordingly, we consider

$$\begin{aligned}
A_1 &= \{|Y - s| \leqslant 2\delta\}, \\
A_2 &= \{|Y - s| > 2\delta\} \cap \left\{|Y - Y^{(\ell)}| \geqslant \delta\right\}, \\
A_3 &= \{|Y - s| > 2\delta\} \cap \left\{|Y - Y^{(\ell)}| < \delta\right\},
\end{aligned}$$

and we then have

$$\mathrm{E}\,\Delta^q = \mathrm{E}(\Delta^q \cdot 1_{A_1}) + \mathrm{E}(\Delta^q \cdot 1_{A_2}).$$

Provided that $p_1 = P(A_1) > 0$,

$$\begin{aligned}
\mathrm{E}(\Delta^q \,|\, A_1) &\leqslant \left\{\text{Jensen's inequality in the form } \mathrm{E}\,(\Delta^p)^{\frac{q}{p}} \leqslant (\mathrm{E}\,\Delta^p)^{\frac{q}{p}} \text{ due to } q < p\right\} \\
&\leqslant (\mathrm{E}(\Delta^p \,|\, A_1))^{q/p} \preceq \left\{\left|g((Y - s)/\delta) - g((Y^{(l)} - s)/\delta)\right| \leqslant \frac{c_L}{\delta}\left|Y - Y^{(l)}\right|\right\} \\
&\preceq \delta^{-q}\, p_1^{-q/p} \cdot \|Y - Y^{(\ell)}\|_p^q.
\end{aligned}$$

Conditional probability formula delivers

$$\mathrm{E}(\Delta^q \cdot 1_{A_1}) = p_1\, \mathrm{E}(\Delta^q \,|\, 1_{A_1}) \preceq \delta^{-q}\, p_1^{1-q/p} \cdot \|Y - Y^{(\ell)}\|_p^q.$$

From the boundedness of the density of Y we have

$$p_1 = \int_{s-2\delta}^{s+2\delta} \rho(\xi)\, d\xi \leqslant 4c_\rho \delta.$$

From the latter we finally obtain

$$\delta^{-q} p_1^{1-q/p} \cdot \|Y - Y^{(\ell)}\|_p^q \preceq \delta^{1-q-q/p} \cdot \|Y - Y^{(\ell)}\|_p^q.$$

Turning now to A_2, Markov's inequality gives the estimate for p_2:

$$p_2 \leqslant P\left(\left\{|Y - Y^{(\ell)}| \geqslant \delta\right\}\right) \leqslant \delta^{-p} \cdot \|Y - Y^{(\ell)}\|_p^p.$$

and hence, using the boundedness of g,

$$\begin{gathered}
\mathrm{E}(\Delta^q \cdot 1_{A_2}) \preceq p_2 \leqslant \delta^{-p} \|Y - Y^{(\ell)}\|_p^p = \delta^{-p} \|Y - Y^{(\ell)}\|_p^{p-q} \|Y - Y^{(\ell)}\|_p^q \leqslant \\
\{\text{apply (69) to the term with power } p - q\} \leqslant \delta^{1-q-q/p} \|Y - Y^{(\ell)}\|_p^q.
\end{gathered}$$

If $\|Y - Y^{(\ell)}\|_p$ and $\|Y - Y^{(\ell)}\|_1$ are asymptotically equivalent for every $1 \leqslant p < \infty$, then Lemma 6.3.1 and Lemma 6.3.2 should be applied with large values of p, and this yields (A5) with β_1 arbitrarily close to 1 and (A3) with α_1 arbitrarily close to 0. See Sections 7.1.1 and 7.1.2 for examples.

6.3.3 *The Multi-level Algorithm*

We study multi-level algorithms

$$\mathcal{M}^{\delta,L_0,L_1}_{N_{L_0},\dots,N_{L_1}} = \frac{1}{N_{L_0}} \cdot \sum_{i=1}^{N_{L_0}} g^\delta(Y_i^{(L_0)}) + \sum_{\ell=L_0+1}^{L_1} \frac{1}{N_\ell} \cdot \sum_{i=1}^{N_\ell} \left(g^\delta(Y_i^{(\ell)}) - g^\delta(Z_i^{(\ell)}) \right)$$

with

$$g^\delta(t) = g((t-s)/\delta), \qquad t \in \mathbb{R},$$

which form a particular instance of (28). The error of $\mathcal{M} = \mathcal{M}^{\delta,L_0,L_1}_{N_{L_0},\dots,N_{L_1}}$ is defined by

$$\operatorname{error}(\mathcal{M}) = \left(\mathrm{E}\,|F(s) - \mathcal{M}|^2 \right)^{1/2},$$

and Remark 2 applies to single-level algorithms.

Put

$$\beta^\dagger = \frac{\beta_1}{\beta_2 - \beta_3},$$

and recall the definition of q from (32).

Theorem 6.3.3 *The following order, with* $\eta = 0$, *is achieved by algorithms* $\mathcal{M}^{\delta,L_0,L_1}_{N_{L_0},\dots,N_{L_1}}$ *with suitably chosen parameters:*

$$q \leqslant \beta^\dagger \wedge \beta_3 \neq 1 \quad \Rightarrow \quad \gamma = 2 + \frac{(1-\beta_3)_+ \cdot q}{r+1}, \tag{70}$$

$$q > \beta^\dagger \wedge \beta_3 \neq 1 \wedge \beta_2 > 1 \quad \Rightarrow \quad \gamma = 2 + \frac{(1-\beta_3)_+ \cdot \beta^\dagger}{r+1}, \tag{71}$$

$$q > \beta^\dagger \wedge \beta_2 < 1 \quad \Rightarrow \quad \gamma = 2 + \frac{\beta_1 + (1-\beta_2) \cdot q}{r+1}. \tag{72}$$

Moreover, with $\eta = 2$,

$$\beta_3 = 1 \quad \Rightarrow \quad \gamma = 2, \tag{73}$$

$$q > \beta^\dagger \wedge \beta_2 = 1 \quad \Rightarrow \quad \gamma = 2 + \frac{\beta_1}{r+1}. \tag{74}$$

Proof We proceed analogously to the proof of Theorem 6.1.5. Use Lemma 6.1.1, the assumptions (A3) and (A5), and the boundedness of g to obtain

$$\begin{aligned}\operatorname{error}^2(\mathcal{M}) \preceq \delta^{2(r+1)} &+ \min\left(\delta^{-2\alpha_1} \cdot M^{-L_1 \cdot 2\alpha_2}, M^{-L_1 \cdot 2\alpha_3}\right) \\ &+ \frac{1}{N_{L_0}} + \sum_{\ell=L_0+1}^{L_1} \frac{\min\left(\delta^{-\beta_1} \cdot M^{-\ell \cdot \beta_2}, M^{-\ell \cdot \beta_3}\right)}{N_\ell \cdot \delta^2}\end{aligned} \tag{75}$$

for $\mathcal{M} = \mathcal{M}^{\delta,L_0,L_1}_{N_{L_0},\dots,N_{L_1}}$. Furthermore, by (S1) and (A2),

$$\operatorname{cost}(\mathcal{M}) \preceq c(L_0, L_1, N_{L_0}, \dots, N_{L_1}) \tag{76}$$

with

$$c(L_0, L_1, N_{L_0}, \dots, N_{L_1}) = \sum_{\ell=L_0}^{L_1} N_\ell \cdot M^\ell.$$

We minimize the upper bound (76) for the cost, subject to the constraint that the upper bound (75) for the squared error is at most ε^2, up to multiplicative constants for both quantities.

To this end we choose δ according to (42), and, up to integer rounding,

$$N_{L_0} = \varepsilon^{-2} \tag{77}$$

as well as

$$L_1 = q \cdot L^* \tag{78}$$

with L^* given by (46).

For a single-level algorithm, i.e., $L_0 = L_1$, this yields $\operatorname{error}(\mathcal{M}) \preceq \varepsilon$ and

$$c(L_1, L_1, N_{L_1}) \asymp \varepsilon^{-2-q/(r+1)}. \tag{79}$$

For a proper multi-level algorithm, i.e., $L_0 < L_1$, we obtain

$$\operatorname{error}^2(\mathcal{M}) \preceq \varepsilon^2 + \sum_{\ell=L_0+1}^{L_1} \frac{\nu_\ell}{N_\ell}$$

with

$$\nu_\ell = \min\left(M^{L^* \cdot \beta_1} \cdot M^{-\ell \cdot \beta_2}, M^{-\ell \cdot \beta_3}\right)$$

as well as

$$c(L_0, L_1, N_{L_0}, \dots, N_{L_1}) \asymp \varepsilon^{-2} \cdot M^{L_0} + \sum_{\ell=L_0+1}^{L_1} N_\ell \cdot M^\ell.$$

Fix L_0 for the moment. We minimize

$$h(L_0, N_{L_0+1}, \dots, N_{L_1}) = \varepsilon^{-2} \cdot M^{L_0} + \sum_{\ell=L_0+1}^{L_1} N_\ell \cdot M^\ell$$

subject to

$$\sum_{\ell=L_0+1}^{L_1} \frac{\nu_\ell}{N_\ell} \leqslant \varepsilon^2.$$

A Lagrange multiplier method leads to

$$N_\ell = \varepsilon^{-2} \cdot G(L_0) \cdot \left(\nu_\ell \cdot M^{-\ell}\right)^{1/2}, \tag{80}$$

up to integer rounding, which satisfies the constraint with

$$G(L_0) = \sum_{\ell=L_0+1}^{L_1} \left(\nu_\ell \cdot M^\ell\right)^{1/2} = \sum_{\ell=L_0+1}^{L_1} \left(\min(M^{L^* \cdot \beta_1} \cdot M^{-\ell \cdot \beta_2}, M^{-\ell \cdot \beta_3}) \cdot M^\ell\right)^{1/2}.$$

Moreover, this choice of $N_{L_0+1}, \dots, N_{L_1}$ yields

$$h(L_0, N_{L_0+1}, \dots, N_{L_1}) = \varepsilon^{-2} \cdot \left(M^{L_0} + G^2(L_0)\right).$$

Put

$$L^\dagger = \beta^\dagger \cdot L^*.$$

In the case $q \leqslant \beta^\dagger$ we have $L_1 \leqslant L^\dagger$, and therefore

$$M^{L_0} + G^2(L_0) = M^{L_0} + \left(\sum_{\ell=L_0+1}^{L_1} M^{\ell \cdot (1-\beta_3)/2}\right)^2.$$

In the case $q > \beta^\dagger$ we have $L^\dagger < L_1$, and therefore

$$M^{L_0} + G^2(L_0) = M^{L_0} + \left(\sum_{\ell=L_0+1}^{L^\dagger} M^{\ell\cdot(1-\beta_3)/2} + M^{L^*\cdot\beta_1/2} \cdot \sum_{\ell=L^\dagger+1}^{L_1} M^{\ell\cdot(1-\beta_2)/2} \right)^2 .$$

Since

$$M^{L_0} + \left(\sum_{\ell=L_0+1}^{L} M^{\ell\cdot(1-\beta_3)/2} \right)^2 \asymp \begin{cases} M^{L_0}, & \text{if } \beta_3 > 1, \\ M^{L_0} + (L-L_0)^2, & \text{if } \beta_3 = 1, \\ M^{L_0} + M^{L\cdot(1-\beta_3)}, & \text{if } \beta_3 < 1, \end{cases}$$

for $L = L_1$ and $L = L^\dagger$, we take

$$L_0 = 0$$

in both cases.

This leads to

$$M^{L_0} + G^2(L_0) \asymp \begin{cases} 1, & \text{if } \beta_3 > 1, \\ L_1^2, & \text{if } \beta_3 = 1, \\ M^{L_1\cdot(1-\beta_3)}, & \text{if } \beta_3 < 1, \end{cases}$$

if $q \leqslant \beta^\dagger$. Moreover, it is straightforward to verify

$$M^{L_0} + G^2(L_0) \asymp \begin{cases} 1, & \text{if } \beta_3 > 1, \\ (L^\dagger)^2, & \text{if } \beta_3 = 1, \\ M^{L^\dagger\cdot(1-\beta_3)}, & \text{if } \beta_3 < 1 \text{ and } \beta_2 > 1, \\ M^{L^*\cdot\beta_1} \cdot (L_1 - L^\dagger)^2, & \text{if } \beta_2 = 1, \\ M^{L^*\cdot(\beta_1 + q(1-\beta_2))}, & \text{if } \beta_2 < 1, \end{cases}$$

if $q > \beta^\dagger$. Except for the case $\beta_3 = 0$ and $q \leqslant \beta^\dagger$ these estimates are superior to M^{L_1}, which corresponds to (79).

Remark 6 *The following comments on optimality etc. are meant in the sense of Remark 3. The optimal values of* δ, N_{L_0}, *and* L_1 *are given by (42), (77), and (78), which completes the optimization of the parameters of single-level algorithms. For proper multi-level algorithms,* $L_0 = 0$ *is optimal, and the optimal replication numbers* $N_{L_0+1}, \dots, N_{L_1}$ *and* L_0 *can be easily derived from (80).*

Proper multi-level algorithms are superior to single-level algorithms if and only if

$$\beta_3 \neq 0 \vee q > \beta_1/\beta_2.$$

In the case $\beta_3 = 0$ *and* $q \leqslant \beta_1/\beta_2$ *the lack of superiority is caused by the negative impact of smoothing, which leads to variances of order one on all levels level* $\ell = 0, \dots, L_1$.

Single-level algorithms with smoothing are superior to single-level algorithms without smoothing if and only if

$$\frac{r+1}{\alpha_3} > \frac{r+1+\alpha_1}{\alpha_2}. \tag{81}$$

7

APPLICATIONS AND NUMERICAL EXPERIMENTS

7.1 APPLICATIONS

At first we consider a general situation, where all we have at hand is (A1), (A2), and an upper bound on the order of the strong error of $Y - Y^{(\ell)}$, which does not depend on p. Specifically, we assume that there exists a constant

$$0 < \beta \leqslant 2$$

with the following property. For every $1 \leqslant p < \infty$ there exists a constant $c_p > 0$ such that

$$\|Y - Y^{(\ell)}\|_p \leqslant c_p \cdot M^{-\ell \cdot \beta/2} \tag{82}$$

for every $\ell \in \mathbb{N}$. In the sequel $\varepsilon > 0$ may be chosen arbitrarily small.

From (82) we obtain (A4) with

$$\beta_4 = 2, \qquad \beta_5 = \beta, \tag{83}$$

see (25), and Lemma 6.3.1 and Lemma 6.3.2 yield (A5) with

$$\beta_1 = 1 + \varepsilon, \qquad \beta_2 = \beta, \qquad \beta_3 = \beta/2 - \varepsilon \tag{84}$$

under the assumptions (S2) and (S3) or (S2) and (S5). Using Lemma 6.3.1 and Lemma 6.3.2 again we get (A3) under both sets of assumptions on g with

$$\alpha_1 = \varepsilon, \qquad \alpha_2 = \beta/2, \qquad \alpha_3 = \beta/2 - \varepsilon, \tag{85}$$

and (29) holds with

$$\alpha = \beta/2 - \varepsilon. \tag{86}$$

It follows that

$$q = \frac{2 \cdot (r+1)}{\beta} + \varepsilon$$

and

$$\max(1, \beta_4/\beta_5) = 2/\beta,$$

so that (34), (36), and (37) in Theorem 6.1.5 yield

$$1 \leqslant \beta \leqslant 2 \qquad \Rightarrow \qquad \gamma = 2 + \frac{2}{\beta \cdot (r+1)}, \tag{87}$$

$$0 < \beta < 1 \qquad \Rightarrow \qquad \gamma = \frac{2}{\beta} + \frac{2}{r+1} + \varepsilon \tag{88}$$

for the approximation of F on $[S_0, S_1]$. Likewise, (58), (60), and (61) in Theorem 6.2.2 yield

$$1 \leqslant \beta \leqslant 2 \qquad \Rightarrow \qquad \gamma = 2 + \frac{2 \cdot (1+\beta)}{\beta \cdot r}, \tag{89}$$

$$0 < \beta < 1 \qquad \Rightarrow \qquad \gamma = \frac{2}{\beta} + \frac{2 \cdot (1+\beta)}{\beta \cdot r} + \varepsilon, \tag{90}$$

for the approximation of ρ on $[S_0, S_1]$. Moreover,

$$\beta^\dagger = 2/\beta + \varepsilon,$$

so that (71), (72), and (74) in Theorem 6.3.3 yield

$$1 \leqslant \beta \leqslant 2 \qquad \Rightarrow \qquad \gamma = 2 + \frac{2-\beta}{\beta \cdot (r+1)} + \varepsilon, \tag{91}$$

$$0 < \beta < 1 \qquad \Rightarrow \qquad \gamma = \frac{2}{\beta} + \frac{1}{r+1} + \varepsilon \tag{92}$$

for the approximation of F at a single point $s \in [S_0, S_1]$. For all three problems we get $\gamma = \max(2, 2/\beta)$ in the limit $r \to \infty$, and proper multi-level algorithms are always superior to single-level algorithms, see Remarks 3, 5, and 6.

Remark 7 *We compare the smoothing approach for the approximation of F at a single point with a direct approach, which only requires that Y has a bounded density ρ (see [5]), see Lemma 6.3.1.*

We study multi-level algorithms

$$\mathcal{M}^{L_0,L_1}_{N_{L_0},\ldots,N_{L_1}} =$$

$$= \frac{1}{N_{L_0}} \cdot \sum_{i=1}^{N_{L_0}} 1_{]-\infty,s]}(Y_i^{(L_0)}) + \sum_{\ell=L_0+1}^{L_1} \frac{1}{N_\ell} \cdot \sum_{i=1}^{N_\ell} \left(1_{]-\infty,s]}(Y_i^{(\ell)}) - 1_{]-\infty,s]}(Z_i^{(\ell)}) \right)$$

for the approximation of $F(s)$. As previously, we assume that (82) with $0 < \beta \leqslant 2$ is all we have at hand. The analysis from Theorem 6.3.3 directly applies, if we take

$$\beta_1 = 0, \qquad \beta_2 = \beta/2 - \varepsilon, \qquad \beta_3 = \beta/2 - \varepsilon,$$

and

$$\alpha_1 = 0, \qquad \alpha_2 = \beta/2 - \varepsilon, \qquad \alpha_3 = \beta/2 - \varepsilon.$$

We achieve the order (γ', η') with

$$\gamma' = \frac{2+\beta}{\beta} + \varepsilon,$$

so that the smoothing approach is superior to the direct approach iff $\beta < 2$ and $r \geqslant 1$.

In the sequel we consider three specific settings in the context of stochastic differential equations (SDEs). We let X denote the solution process of the SDE, which is supposed to take values in $\mathbb{R}^d$. For simplicity, we alway take the Euler scheme with equidistant time-steps for approximation of X, and we do not discuss results on the existence and smoothness of densities. As previously, $\varepsilon > 0$ may be chosen arbitrarily small.

7.1.1 *Smooth Path-independent Functionals for SDEs*

Let

$$Y = \varphi(X_T),$$

where $\varphi : \mathbb{R}^d \to \mathbb{R}$ is Lipschitz continuous. We assume that the cost of computing $\varphi(x)$ is uniformly bounded for $x \in \mathbb{R}^d$, and for approximation of Y we use $Y^{(\ell)} = \varphi(X_T^{(\ell)})$, where $X^{(\ell)}$ denotes the Euler scheme with 2^ℓ equidistant time-steps. Obviously, (A2) holds with $M = 2$. For weak error estimates we refer to [7]. Hereby we obtain (A3) with

$$\alpha_1 = 0, \qquad \alpha_2 = 1, \qquad \alpha_3 = 1 \tag{93}$$

under the assumptions (S2) and (S3) or (S2) and (S5) on g and the smoothness and non-degeneracy assumptions (C) and (UH) on the coefficients of the SDE. Furthermore, (29) holds with

$$\alpha = 1.$$

It is well-known that (82) holds with

$$\beta = 1$$

already under standard assumptions on the coefficients of the SDE. Hence we get (A4) with

$$\beta_4 = 2, \qquad \beta_5 = 1, \tag{94}$$

see (83), and (A5) with

$$\beta_1 = 1 + \varepsilon, \qquad \beta_2 = 1, \qquad \beta_3 = 1/2 - \varepsilon,$$

see (84).

We therefore have $q = r + 1$ and $\max(1, \beta_4/\beta_5) = 2$, and (33) and (37) in Theorem 6.1.5 yield

$$(\gamma, \eta) = \begin{cases} (3, 1), & \text{if } r \leqslant 1, \\ (2 + 2/(r+1), 3), & \text{if } r \geqslant 2, \end{cases}$$

for the approximation of F on $[S_0, S_1]$. Likewise, (57) and (61) in Theorem 6.2.2 yield

$$(\gamma, \eta) = \begin{cases} (6, 1), & \text{if } r = 1, \\ (2 + 4/r, 3), & \text{if } r \geqslant 2, \end{cases}$$

for the approximation of ρ on $[S_0, S_1]$. For both problems, proper multi-level algorithms are superior to single-level algorithms if and only if $r \geqslant 2$, see Remarks 3 and 5. Moreover, $\beta^\dagger = 2 + \varepsilon$, so that (70) and (74) in Theorem 6.3.3 yield

$$\gamma = \begin{cases} 5/2 + \varepsilon, & \text{if } r = 0, \\ 2 + 1/(r+1) + \varepsilon, & \text{if } r \geqslant 1, \end{cases}$$

for the approximation of F at a single point $s \in [S_0, S_1]$. For this problem, proper multi-level algorithms are superior to single-level algorithms for every $r \in \mathbb{N}_0$, see Remark 6. For all three problems we get $\gamma = 2$ in the limit $r \to \infty$.

If the coefficients of the SDE merely satisfy the standard assumptions, instead of (C) and (UH) from [7], we may apply (85) to obtain $\alpha_1 = \varepsilon$, $\alpha_2 = 1/2$, and $\alpha_3 = 1/2 - \varepsilon$, see also Section 2.2 in [42]. While the latter is inferior to (93), it leads to essentially the same orders of convergence for approximation of densities or distribution functions if $r \geqslant 1$, see (87), (89), and (91).

Remark 8 *A two-level construction of the form*

$$\mathcal{M}^{\delta, L_0, L_1}_{N_{L_0}, N_{L_1}} = \frac{1}{N_{L_0}} \cdot \sum_{i=1}^{N_{L_0}} g^\delta(Y_i^{(L_0)}) + \frac{1}{N_{L_1}} \cdot \sum_{i=1}^{N_{L_1}} \left(g^\delta(Y_i^{(L_1)}) - g^\delta(Z_i^{(L_1)}) \right),$$

which is the counterpart of the two-level construction from [42] for the approximation of $\mathrm{E}(\varphi(X_T))$, *is employed in [43] for the approximation of the density* ρ *of* $Y = X_T$ *at a single point* s. *Here the sequence* $(Y^{(\ell)})_{\ell \in \mathbb{N}}$ *consists of suitably regularized Euler schemes with* ℓ *equidistant time-steps. By assumption,* $\rho \in C_b^\infty(\mathbb{R}^d, \mathbb{R})$, *i.e., the multi-dimensional counterpart to (A1) is satisfied for every* $r \in \mathbb{N}_0$. *Using Malliavin calculus techniques, the authors derive a central limit theorem for* $L_1 \cdot (\mathcal{M}^{\delta, L_0, L_1}_{N_{L_0}, N_{L_1}} - \rho(x))$ *with properly chosen parameters* L_0, N_{L_0}, N_{L_1}, *and* δ *as* L_1 *tends to infinity. For every dimension* d *the order* $\gamma = 5/2 + \varepsilon$ *is achieved in this way, while the multi-level approach achieves the order* $\gamma = 2 + \varepsilon$ *(at least for* $d = 1$*).*

Remark 9 *Consider the problem of approximating a quantile of* Y*, which is studied in [56] in the particular case of a projection* $\varphi(x) = x_i$*. By assumption,* $\rho \in C_b^\infty(\mathbb{R}, \mathbb{R})$*. The authors employ a single-level algorithm that is based on a suitably regularized Euler scheme, cf. Remark 8. The approximation to the quantile is given as the corresponding empirical quantile, and an error of order* $\gamma = 3$ *is achieved, if* ρ *is bounded away from zero in a neighborhood of the quantile.*

Under the latter assumption, the order of approximation to F *in the supremum norm and to the quantile coincide, and given (A1) for every* $r \in \mathbb{N}_0$ *we expect our multi-level algorithm to achieve the order* $\gamma = 2 + \varepsilon$ *also for quantile approximation and every Lipschitz continuous function* φ*. Furthermore, the multi-level algorithm may be used to approximate the distribution function* F *and the density* ρ *in parallel, which allows to control the impact of inverting the approximation to* F*.*

Remark 10 *We comment on the optimality of the parameters* α_i *and* β_i *according to (93) and (94) in (A3) and (A4). Due to [7], the estimate (A3) with (93) is sharp under the assumptions (C) and (UH). Under standard assumptions,* $2^{\ell/2} \cdot (X - X^{(\ell)})$ *converges in distribution to a stochastic process* U *with* U_T *being non-degenerate in general, see [41]. In the latter case we have a projection* $\varphi(x) = x_i$ *such that (82) with* $M = 2$ *and* $p = 1$ *does not hold for any* $\beta > 1$*. A slight generalization of Lemma 6.1.3 shows that (A4) does not hold for any* $\beta_4 < 2$ *or* $\beta_5 > 1$*. Hence the estimate (A4) with (94) cannot be improved in general for the Euler scheme.*

The approximation of marginal densities of SDE in studied in a number of papers under different aspects. The convergence rate of the density of the Euler approximation $X_T^{(\ell)}$ towards ρ is studied in, e.g., [6] and [31]. In the paper [47], authors construct a forward-reverse kernel estimator and provide an upper bound for its variance.

7.1.2 *Smooth Path-dependent Functionals for SDEs*

Let

$$Y = \varphi(X)$$

with $\varphi : C([0,T], \mathbb{R}^d) \to \mathbb{R}$ being Lipschitz continuous. We assume that the cost of computing $\varphi(x)$ for a piecewise linear path $x \in C([0,T], \mathbb{R}^d)$ with m breakpoints is bounded by a constant times m, and for approximation of Y we use $Y^{(\ell)} = \varphi(X^{(\ell)})$, where $X^{(\ell)}$ denotes the Euler scheme with 2^ℓ equidistant time-steps and piecewise linear interpolation. Then (A2) holds with $M = 2$, and the following fact is well-known under standard assumptions on the coefficients of the SDE. For every $1 \leqslant p < \infty$ there exists a constant $c_p > 0$ such that

$$\|Y - Y^{(\ell)}\|_p \leqslant c_p \cdot \left(\ell \cdot M^{-\ell}\right)^{1/2}$$

for every $\ell \in \mathbb{N}$. Consequently, (82) holds with

$$\beta = 1 - \varepsilon,$$

and we get (A4) with

$$\beta_4 = 2, \qquad \beta_5 = 1 - \varepsilon \tag{95}$$

see (83), (A5) with

$$\beta_1 = 1 + \varepsilon, \qquad \beta_2 = 1 - \varepsilon, \qquad \beta_3 = 1/2 - \varepsilon$$

under the assumptions (S2) and (S3) or (S2) and (S5), see (84), as well as (A3) with

$$\alpha_1 = 0, \qquad \alpha_2 = 1/2 - \varepsilon, \qquad \alpha_3 = 1/2 - \varepsilon, \tag{96}$$

see (85). Furthermore, (29) holds with

$$\alpha = 1/2 - \varepsilon,$$

see (86).

We therefore have $q = 2 \cdot (r+1) + \varepsilon$ and $\max(1, \beta_4/\beta_5) = 2 + \varepsilon$, and (36) in Theorem 6.1.5 yields

$$\gamma = 2 + 2/(r+1) + \varepsilon$$

for the approximation of F on $[S_0, S_1]$. Likewise, (60) in Theorem 6.2.2 yields

$$\gamma = 2 + 4/r + \varepsilon$$

for the approximation of ρ on $[S_0, S_1]$. Moreover, $\beta^\dagger = 2 + \varepsilon$, so that (72) in Theorem 6.3.3 yields

$$\gamma = 2 + 1/(r+1) + \varepsilon$$

for the approximation of F at a single point $s \in [S_0, S_1]$. For all three problems proper multi-level algorithms are always superior to single-level algorithms, see Remarks 3, 5, and 6.

Note that Section 7.1.1 is dealing with a particular instance of the functionals studied here. We achieve essentially the same order of convergence for the problems studied in Sections 7.1.1 and 7.1.2, if $r \geqslant 1$, and we always get $\gamma = 2$ in the limit $r \to \infty$.

Remark 11 *We comment on the optimality of the parameters α_i and β_i according to (96) and (95) in (A3) and (A4). Due to Remark 10 the estimate (A4) with (95) cannot be improved in general for the Euler scheme. Concerning (A3) we are not aware of an optimality result. We refer, however, to [1], where authors study processes $Y^{(\ell)}$ that coincide with the Euler scheme $X^{(\ell)}$ at the discretization points, but instead of 2^ℓ Brownian increments the whole trajectory of the Brownian motion is employed. They provide an upper bound of the order $2/3 - \varepsilon$ for Wasserstein distance of X and $Y^{(\ell)}$ in the case $d = 1$.*

7.1.3 *Stopped Exit Times for SDEs*

Consider a bounded domain $D \subset \mathbb{R}^d$ such that $X_0 \in D$, and let

$$Y = \varphi(X)$$

be the corresponding exit time, stopped at $T > 0$, i.e.,

$$\varphi(x) = \inf\{t \geqslant 0 : x(t) \in \partial D\} \wedge T$$

for $x \in C([0,T], \mathbb{R}^d)$. We assume that the cost of computing $\varphi(x)$ for a piecewise linear path $x \in C([0,T], \mathbb{R}^d)$ with m breakpoints is bounded by a constant times m, and as in the previous section $Y^{(\ell)}$ is the Euler scheme $X^{(\ell)}$ composed with φ. Then (A2) holds with $M = 2$. For every $1 \leqslant p < \infty$ there exists a constant $c_p > 0$ such that

$$\|Y - Y^{(\ell)}\|_p \leqslant c_p \cdot M^{-\ell/(2p)} \tag{97}$$

for every $\ell \in \mathbb{N}$, see [14]. From (26) we get (A4) with

$$\beta_4 = 1, \qquad \beta_5 = 1/2,$$

and (67) and Lemma 6.3.1 yield (A5) with

$$\beta_1 = 1, \qquad \beta_2 = 1/2, \qquad \beta_3 = 1/4.$$

Furthermore, (24) and Lemma 6.3.1 yield (A3) with

$$\alpha_1 = 1, \qquad \alpha_2 = 1/2, \qquad \alpha_3 = 1/4$$

under the assumptions (S2) and (S3) or (S2) and (S5), while (29) holds with

$$\alpha = 1/4.$$

We therefore have $q = 2r + 4$ and $\max(1, \beta_4/\beta_5) = 2$, and (36) in Theorem 6.1.5 yields

$$(\gamma, \eta) = (3 + 2/(r+1), 1)$$

for the approximation of F on $[S_0, S_1]$. Likewise, (60) in Theorem 6.2.2 yields

$$(\gamma, \eta) = (3 + 5/r, 1)$$

for the approximation of ρ on $[S_0, S_1]$. Moreover, $\beta^\dagger = 3$, so that (72) in Theorem 6.3.3 yields

$$(\gamma, \eta) = (3 + 2/(r+1), 0) \tag{98}$$

for the approximation of F at a single point $s \in [S_0, S_1]$. For all three problems, proper multi-level algorithms are superior to single-level algorithms for every $r \in \mathbb{N}_0$, see Remarks 3, 5, and 6, but we only get $\gamma = 3$ in the limit $r \to \infty$. The latter is in contrast to the results from Sections 7.1.1 and 7.1.2, and it is basically due to the fact that the upper bound (97) for strong approximation of Y by $Y^{(\ell)}$ depends on p in the most unfavorable way. We add that numerical experiments suggest that the upper bound (97) cannot be improved, in general. Furthermore, observe that for stopped exit times the same order γ is achieved for the approximation of F on a compact interval and at a single point.

We add that (56) and (81) are satisfied for every $r \geqslant 1$, and therefore smoothing already help for the single-level algorithm to approximate the distribution function of the stopped exit time.

Remark 12 *For the approximation of the mean* $E(Y)$ *of the stopped exit time a multi-level Euler algorithm has been constructed and analyzed in [35]. It is shown that the order* $\gamma = 3 + \varepsilon$ *is achieved under standard smoothness assumptions on the coefficients of the SDE and on the domain* D.

7.2 NUMERICAL EXPERIMENTS

The main goal of our numerical experiments is to demonstrate the potential of the new multi-level algorithm. We consider three benchmark problems according to Sections 7.1.1–7.1.3 for a simple, scalar SDE, where the solutions are known analytically. We present results only for the approximation of distribution functions on a compact interval $[S_0, S_1]$, as the main numerical difference to the other two problems studied in this paper is in the deterministic interpolation part. Our numerical experiments show the computational gain in terms of upper bounds, achieved by the multi-level Monte Carlo approach with smoothing in comparison to the single-level Monte Carlo approach without smoothing. Furthermore, we compare the error of the multi-level algorithm with the accuracy demand ε, which serves as an input to the algorithm. An extensive numerical study of our algorithm is out of the scope of the current paper and will be presented in a subsequent paper.

Consider a geometric Brownian motion X, given by

$$\begin{aligned} dX_t &= \mu \cdot X_t \, dt + \sigma \cdot X_t \, dW_t, \qquad t \in [0, T], \\ X_0 &= 1, \end{aligned}$$

where W denotes a scalar Brownian motion. For the approximation of X we use the Euler scheme with equidistant time-steps, so that $M = 2$. The corresponding values of the parameters α_i and β_i are presented in Sections 7.1.1–7.1.3.

In the examples from this section, the assumption (A1) holds for every $r \in \mathbb{N}$, but typically we think of r being unknown. Hence we choose $\tilde{r} \in \mathbb{N}_0$, instead, and a particular purpose of the numerical experiments is to illustrate the impact of $\tilde{r}$. In all our experiments we take

$$\tilde{r} = 3, 5, 7, 9, 11,$$

and the corresponding smoothing polynomials g according to Remark 1 can be seen in Figure 17.

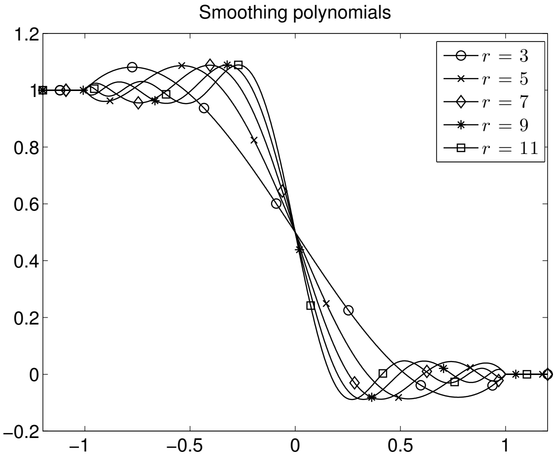


Figure 17: Smoothing polynomials g.

Given ε and $\tilde{r}$, we basically choose the remaining parameters of the multi-level (single-level) algorithm such that all four (three) terms in the upper bound (38) are of the order ε^2. For the multi-level algorithm with smoothing we choose the parameters L_0, L_1, and N_ℓ according to (53), (48), (44), and (54), with r replaced by $\tilde{r}$, while

$$\delta = 2^{-1/(\tilde{r}+1)} \cdot \varepsilon^{1/(\tilde{r}+1)},$$

cf. (42). For the single-level algorithm without smoothing, see Remark 2, we choose $L = L_0 = L_1$ and N_L according to (48) and (44), too, however observing (30), which leads to $q = (\tilde{r}+1)/\alpha$.

In the second stage of the algorithm we employ piecewise polynomial interpolation Q_k^3 of degree 3 with equidistant knots for any $\tilde{r}$. Due to the Lebesgue constants involved, this is preferable to $Q_k^{\tilde{r}}$ with a large value of $\tilde{r}$ if the overall number k of interpolation points is comparatively small. Furthermore, it is convenient if $k-1$ is a multiple of 3 and proportional to the length of the interval $[S_0, S_1]$. In both cases, single-level and multi-level, we therefore take

$$k = 3 \cdot \left\lceil 5 \cdot \varepsilon^{-1/(\tilde{r}+1)} \cdot (S_1 - S_0)/3 \right\rceil + 1, \tag{99}$$

cf. (43).

To specify the computational gain we compare the upper bound (40) for the cost of the multi-level Monte Carlo algorithm with smoothing and the corresponding upper bound

$$c(k, L, N) = N \cdot (2^L + k),$$

for the cost of the single-level algorithm. The ratio $c(k, L_0, L_1, N_{L_0}, \dots, N_{L_1})/c(k, L, N)$, which is a function of the desired accuracy ε, is used to describe the computational gain.

To assess the accuracy of the multi-level algorithm, $\mathrm{error}(Q_k^3(\mathcal{M}))$, which depends on ε and $\tilde{r}$, should be compared with the desired accuracy ε. Since $\mathrm{error}(Q_k^3(\mathcal{M}))$ is not known exactly, we employ a simple Monte Carlo experiment with 25 independent replications for each of the values of $\tilde{r}$ and each of the values $\varepsilon = 2^{-i}$ for $i = 3, \dots, 11$. The estimate is denoted by $\mathrm{RMSE}(\varepsilon, \tilde{r})$.

7.2.1 *Smooth Path-independent Functionals for SDEs*

In this section we set

$$\mu = 0.05,\ \sigma = 0.2,\ T = 1,$$

and we approximate the distribution function $F(s) = E(1_{]\infty,s]}(Y))$ of

$$Y = X_T$$

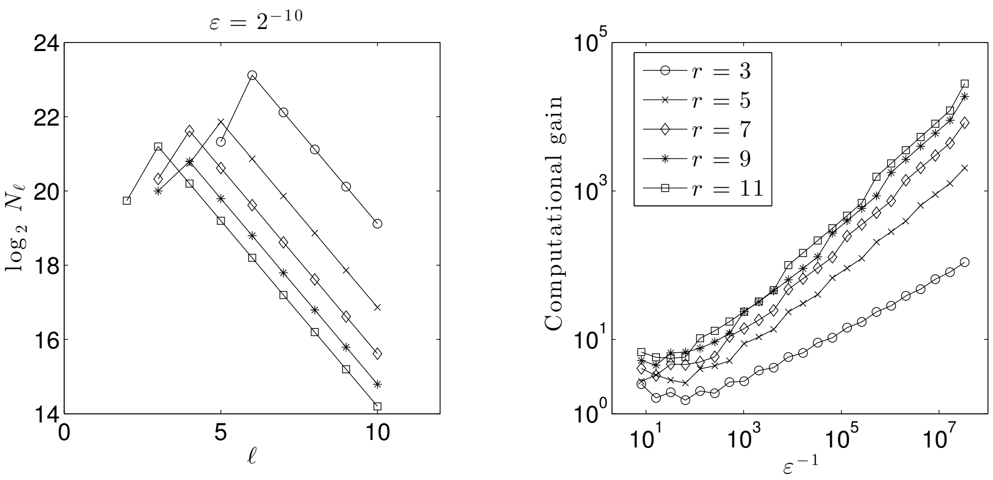


Figure 18: Replication numbers per level (left) and computational gain (right).

on the interval

$$[S_0, S_1] = [0, 2].$$

Note that Y is lognormally distributed with parameters $\mu - \sigma^2/2$ and σ^2.

The computational gain as well as the replication numbers N_ℓ for the multi-level algorithm with $\varepsilon = 2^{-11}$ are presented in Figure 18. The maximal level L_1 of the multi-level algorithm coincides with the level chosen by the single-level algorithm, and this level does not depend on $\tilde{r}$. For smaller values of $\tilde{r}$ the multi-level algorithm start on a higher level L_0, and therefore the computational gain in the case $\tilde{r} = 3$ is only about a factor two. For large values of $\tilde{r}$ we observe a reasonable computational gain already for moderate values of ε. In Figure 19 we compare the estimate $\text{RMSE}(\varepsilon, \tilde{r})$ for the error of the multi-level algorithm and the accuracy demand ε. Note that $\text{RMSE}(\varepsilon, \tilde{r})$ is in the range of ε; actually, it is less that ε in almost all cases.

7.2.2 *Smooth Path-dependent Functionals for SDEs*

For this test case we use the same parameters for the SDE and the same interval $[S_0, S_1]$ as in Section 7.2.1. We approximate the function

$$F(s) = E\left(e^{-\mu \cdot T} \cdot \max(X_T - X_0, 0) \cdot 1_{]-\infty, s]}(Y)\right),$$

where

$$Y = \max_{t \in [0,T]} X_t.$$

See p.207 in [54] for the analytical solution. Note that this problem does not exactly fit into our framework, due the the presence of $\max(X_T - X_0, 0)$ in the definition of the functional. Still, the multi-level smoothing approach is applicable.

See Figures 20, with replication numbers for $\varepsilon = 2^{-10}$, and 21 for the results. As the main difference, compared to the previous section, the computational gain is substantially larger for the path-dependent functional. This is due to the following facts. The orders of strong convergence are essentially the same for both problems. However, the maximal level, which once more coincide with the level chosen by the single-level algorithm, is essentially twice as large as in the previous case, due to the slower decay of the bias. This results in a larger value of $L_1 - L_0$, which provides an advantage to the multi-level approach.

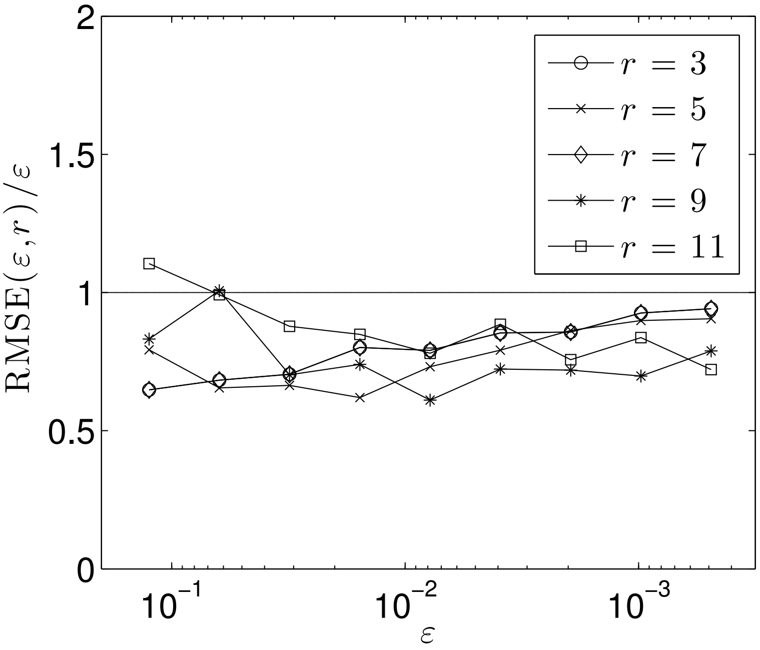


Figure 19: Error vs. accuracy demand ε.

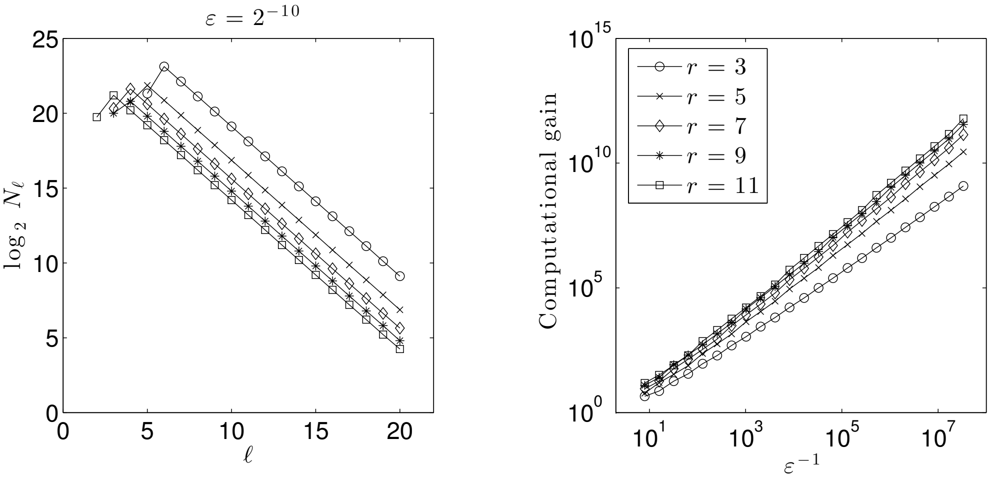


Figure 20: Replication numbers per level (left) and computational gain (right).

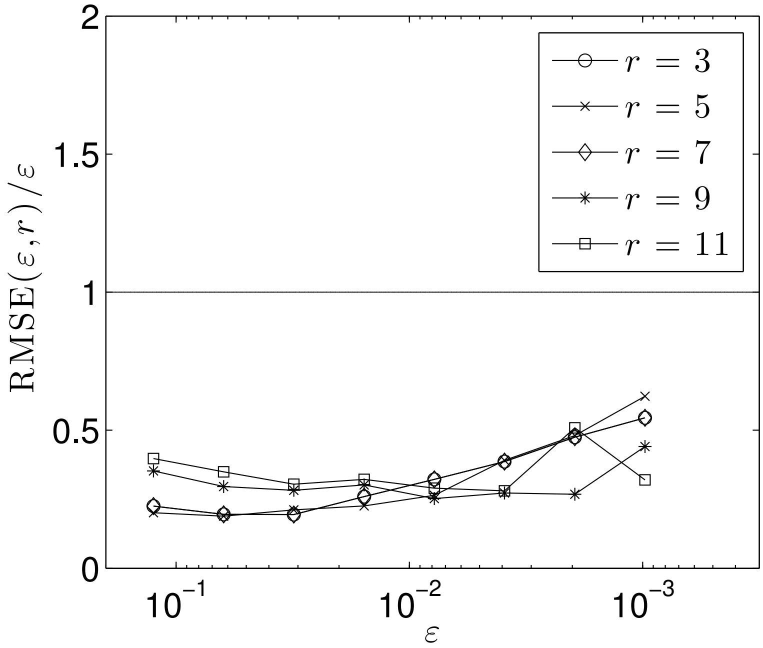


Figure 21: Error vs. accuracy demand ε.

7.2.3 *Stopped Exit Times for SDEs*

In this section we set

$$\mu = 0.01,\ \sigma = 0.2,\ T = 2,$$

and we approximate the distribution function $F(s) = E(1_{]\infty,s]}(Y))$ of

$$Y = \inf\{t \geqslant 0 : X_t = b\} \wedge T$$

with $b = 0.8$ on the interval

$$[S_0, S_1] = [0, 1].$$

The distribution of $\inf\{t \geqslant 0 : X_t = b\}$ is an inverse Gaussian distribution with parameters $\ln b/(\mu - \sigma^2/2)$ and $(\ln b)^2/\sigma^2$, and this yields an explicit formula for F since $T > S_1$.

See Figures 22, with replication numbers for $\varepsilon = 2^{-9}$, and 23 for the results. Observe that the computational gain is even larger than in the previous section. This difference is due to the fact that smoothing already yields an improved weak error estimate for the present problem. Consequently,

$$L_1 = \left(2 + \frac{2}{\tilde{r} + 1}\right) \cdot \log_2 \varepsilon^{-1}$$

is the maximal level for the multi-level algorithm, up to integer rounding, but for the single-level algorithm without smoothing we have to take

$$L = 4 \cdot \log_2 \varepsilon^{-1}.$$

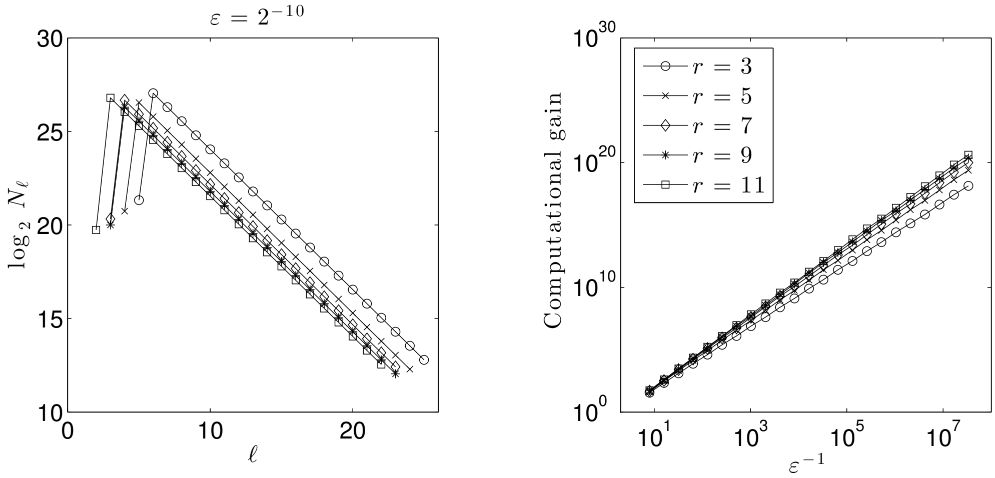


Figure 22: Replication numbers per level (left) and computational gain (right).

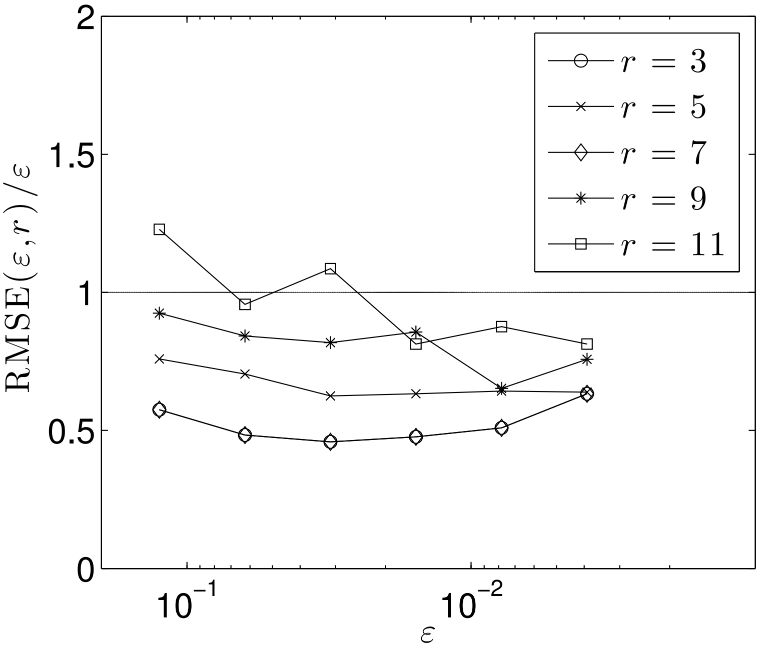


Figure 23: Error vs. accuracy demand ε.

Part IV

MODELLING ASYMMETRIC FLOW FIELD FLOW FRACTIONATION WITH STOCHASTIC DIFFERENTIAL EQUATIONS WITH REFLECTION

8

STOCHASTIC DIFFERENTIAL EQUATIONS WITH REFLECTION

We will start with a short description of mathematical tools, used during the modeling of AF4. As mentioned in the Section 1.3, we assume that the particles do not interact with each other and do not change their hydrodynamic radius. And their drift at any point in the channel is equal to the velocity of the solvent at that point. Based on these assumptions we will build an SDE based model for the particles motion in the channel.

We first present a stochastic model that describes the motion of a particle in a bounded convex domain $D \subset \mathbb{R}^d$ with a C^2-boundary. In the sequel, $n(x)$ denotes the unique inward normal vector at any point $x \in \partial D$, and $\pi(x)$ denotes the projection of any point $x \in \mathbb{R}^d$ onto $\overline{D}$ w.r.t. the Euclidean distance.

Consider a d-dimensional standard Brownian motion $(W(t))_{t\geqslant 0}$ on a filtered probability space that satisfies the usual conditions, an initial value $x_0 \in D$, and Lipschitz continuous mappings $\mu : \overline{D} \to \mathbb{R}^d$ and $\sigma : \overline{D} \to \mathbb{R}^{d\times d}$. A strong solution to the associated *Skorohod problem* is a pair of adapted continuous processes $X = (X(t))_{t\geqslant 0}$ and $\phi = (\phi(t))_{t\geqslant 0}$ taking values in $\overline{D}$ and $\mathbb{R}^d$, respectively, with the following properties. The process ϕ is locally of bounded variation with $\phi(0) = 0$, and with $|\phi|(t)$ denoting the total variation of ϕ on $[0, t]$ we have

$$|\phi|(t) = \int_0^t 1_{\partial D}(X(s))\, d|\phi|(s), \tag{100}$$

$$\phi(t) = \int_0^t n(X(s))\, d|\phi|(s) = \int_0^t n(X(s))\, d\xi(s), \tag{101}$$

$$X(t) = x_0 + \int_0^t \mu(X(s))\, ds + \int_0^t \sigma(X(s))\, dW(s) + \phi(t) \tag{102}$$

for $t \geqslant 0$.

Definition 2 *A pair* $(X, \phi) = (X(t), \phi(t))_{t\in[0,\infty[}$ *of continuous adapted processes with values in* $\overline{D}$ *and* $\mathbb{R}^d$*, respectively, is a strong solutions of the Skorohod equation* (100)-(101)*, if* $X(0) \in \overline{D}$, $\phi(0) = 0$ *and* (100)-(101) *hold true.*

For stochastic Skorohod problems the existence and uniqueness of solutions is studied under various assumptions on the domain D and on the driving process. We will state the standard result from Saisho (1979) for an existence and uniqueness result, which covers in particular the present case. In addition, we will state that X is a strong Markov process, which actually holds for more general driving processes, see Kohatsu-Higa (2001).

8.0.4 *Boundary*

Define the set $\mathbb{N}_x$ of inward normal unit vectors at $x \in \partial D$ by

$$\mathbb{N}_x = \cup_{r>0}\mathbb{N}_{x,r},$$

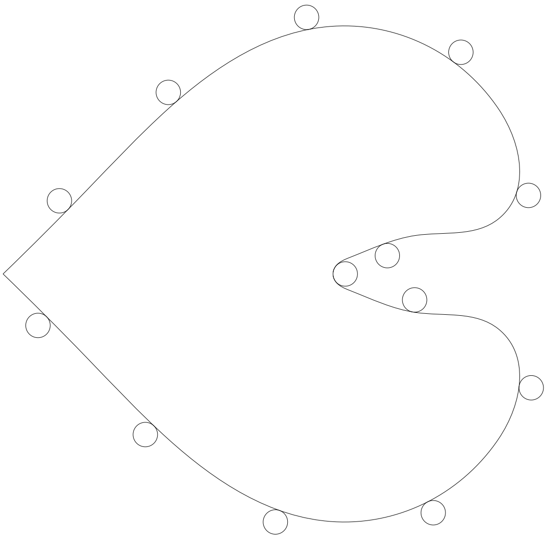


Figure 24: Example of the domain, satisfying Conditions 1 and 2.

$$\mathbb{N}_{x,r} = \left\{ n \in \mathbb{R}^2 : |n| = 1,\ B(x - rn,\ r) \cap D = \emptyset \right\},$$

where $B(x,\ r) = \left\{ y \in \mathbb{R}^2 : |x - y| < r \right\}$, $r > 0$.

Remark 8.0.1 N_x *can be empty, for example in the case of "non convex angle".*

Condition 1 (Uniform exterior sphere condition) *Following Lions and Sznitman [44], we introduce two conditions on the domain* D*.*

There exists a constant r_0 *such that*

$$\mathbb{N}_x = \mathbb{N}_{x,r_0} \neq \emptyset,\ \forall x \in \partial D$$

Remark 8.0.2 *Some general comments on Condition 1*

- *Condition 1 means that we can touch every* $x \in \partial D$ *from outside with a ball with radius* r_0 *without intersection with domain* D *(see Figure 24)*
- *Condition 1 does not necessarily mean, that* ∂D *is smooth (see Figure 24)*
- *If* D *is convex, then* r_0 *can be set equal to* ∞*.*
- *If* $\operatorname{dist}(x, \overline{D}) < r_0$, $x \notin \overline{D}$ *then there exists the unique projection* $\pi(x)$ *such that*

$$|x - \pi(x)| = \operatorname{dist}(x, \overline{D})$$

Condition 2 *There exist constants* $\delta > 0$ *and* $\beta \in [1, \infty)$ *with the following property: for any* $x \in \partial D$ *there exists a unique vector* l_x *such that*

$$\langle l_x,\ n \rangle \geqslant \beta^{-1}, \quad \forall n \in \cup_{y \in B(x,\delta) \cap \partial D} \mathbb{N}_y$$

Remark 8.0.3 *Condition 2 means that the normal vectors do not fluctuate widely.*

Remark 8.0.4 *If* D *is convex or has a* $\mathcal{C}^2$ *boundary, then Conditions 1 and 2 are verified.*

8.0.5 *Existence and Uniqueness*

Theorem 8.0.5 (Saisho [53]) *If Conditions 1, 2 for domain* D *are verified and* b *and* σ *are uniformly Lipschitz continuous in space functions, then there exists a unique strong solution of the Skorohod problem* (100)-(101).

Remark 13 *By* (100)*, the random measure corresponding to* $|\phi|$ *is concentrated on the set of time instances* $s \in [0, t]$ *such that* $X(s) \in \partial D$*. Moreover, for any continuous function* f *of bounded variation on* $[0, t]$ *we have*

$$f(t) = \int_0^t m(s)\, d|f|(s)$$

for some function m *with values in the unit sphere in* $\mathbb{R}^d$*, and* m *is uniquely determined almost everywhere w.r.t. the measure associated with the total variation* $|f|$ *of* f*. By* (101)*, the corresponding representation of the process* ϕ *is given by the inward normal vectors at the points* $X(s) \in \partial D$*. The formal counterpart*

$$dX(t) = \mu(X(t))\, dt + \sigma(X(t))\, dW(t) + d\phi(t) \tag{103}$$

of (102) *is therefore called a* stochastic differential equation with normal reflection *at the boundary of* D.

Instead of reflection we actually have *absorption* in a part of the boundary, which is denoted by $\partial_a D$ and which is assumed to be an open set

$$\emptyset \neq \partial_a D \subsetneq \partial D,$$

and we consider the *hitting time*

$$\tau = \inf\{t \geqslant 0 : X(t) \in \partial_a D\}.$$

The computational problem is to determine the distribution function

$$F(t) = P(1_{(\tau \leqslant t)}), \qquad t \geqslant 0,$$

on a compact time interval $[0, T]$. Put

$$\partial_r D = \partial D \setminus \partial_a D.$$

In the application described below, the Skorohod problem together with τ is used to model the motion of a particle that is reflected at $\partial_r D$, while it leaves the domain D via $\partial_a D$.

Remark 14 *Numerical problems for SDEs with reflection and for exit times of solutions of SDEs have been studied in a number of papers, a few of which will be presented here and used subsequently.*

For an SDE without reflection let τ *denote the exit time from the domain* $D \subset \mathbb{R}^d$*. [32] and [35] study weak and strong approximation, respectively, of the so called* stopped hitting time $\tau \wedge T$ *via the Euler scheme. Furthermore, the multi-level Monte Carlo approach is used in [35] to compute the expectation* $E(\tau \wedge T)$.

For SDEs with normal reflection, variants of the Euler scheme are employed, for instance, in the following papers. For strong approximation of $(X(t))_{t \in [0,T]}$ *[55] establishes upper bounds of order almost* $1/4$ *for convex domains and of order almost* $1/2$ *for convex polyhedra, where the error is based on the* L_∞*-distance of corresponding paths. Weak approximation for a class of functionals of* X *and* ϕ *is studied in [49], where they derive upper bounds of order (almost)* $1/2$ *under suitable regularity assumptions. For recent work on weak approximation on stopped and reflected diffusions and for further references, see [8]; in particular, an adaptive weak approximation scheme is constructed there, and the application of the multi-level approach is proposed in this context.*

8.1 PDE REPRESENTATION

Further everywhere where it's important all vectors are by default meant to be columns.

We consider d-dimensional stochastic equation

$$dX(t) = b(X(t))\,dt + \sigma(X(t))\,dW_t + n(t)\,d\xi_t \tag{104}$$

with initial value $X_0 = x \in D$, $\xi_0 = 0 \in \mathbb{R}$. $n(t)$ is assumed to be an adapted process taking values in the set of inward pointing normal vectors of $\partial D = \overline{D} \setminus D$ at $X(t)$, provided that $X(t) \in \partial D$.

Let L denote the infinitesimal generator of the SDE (104):

$$Lv(x) = \langle \mu(x), \nabla v(x) \rangle + \frac{1}{2} \mathrm{tr} \left(\sigma\sigma^T H(v(x)) \right).$$

$H(v(x))$ stands for Hessian matrix of function v. We consider the following boundary value problem:

$$\begin{cases} \dfrac{\partial}{\partial t} u(t,x) + Lu(t,x) = 0, & x \in \overline{D}, \\ u(T,x) = f(x), & x \in \overline{D}, \\ u(t,x) = g(x), & x \in \partial_a D, \\ \dfrac{\partial}{\partial n} u(t,x) = h(x), & x \in \partial_r D, \end{cases} \tag{105}$$

where $t \in [0,T]$, $f\colon \overline{D} \to \mathbb{R}$, $g\colon \partial_a D \to \mathbb{R}$ and $h\colon \partial_r D \to \mathbb{R}$ are sufficiently regular functions, $n(x)$ is assumed to be unique inward normal vector at $x \in \partial D$ and we denote

$$\frac{\partial}{\partial n} u(t,x)$$

the normal derivative of $u(t,x)$ at $x \in \partial D$.

The following theorem is the special case of Theorem II.5.1 in [21].

Theorem 8.1.1 *Assume that the solution* $u(t,x)$ *of problem* (105) *has bounded time-derivative, gradient and Hessian matrix for all* $(t,x) \in [0,T] \times \overline{D}$. *Then we have the stochastic representation.*

$$u(t,x) = E\left[f(X_T)\mathbb{1}_{\tau \geqslant T} + g(X_\tau)\mathbb{1}_{\tau < T} - \int_t^{\tau \wedge T} h(X(s))\,d\xi_s \,\middle|\, X(t) = x \right]$$

where $\tau = \inf\{s \geqslant t | X(s) \in \partial D_s\}$ *and* $x \wedge y = \min(x,y)$.

Proof

$$\begin{aligned} du(t,X(t)) = &\left(\frac{\partial u}{\partial t} + \langle \mu, \nabla u \rangle + \frac{1}{2}\mathrm{tr}\left(\sigma\sigma^T H(u)\right) \right) dt + \\ &+ (\nabla u)^T \sigma dW + (\nabla u)^T n(t) d\xi \end{aligned}$$

Fixing $T > t > 0$ and the stopping time τ, we apply Itô's formula to $u(s, X(s))$ and obtain

$$\begin{aligned} u(T\wedge\tau, X_{T\wedge\tau}) =& u(t,X(t)) + \int_t^{T\wedge\tau} \left(Lu(s,X(s)) + \frac{\partial u}{\partial t}(s,X(s)) \right) ds \\ &+ \int_t^{T\wedge\tau} (\nabla u(s,X(s)))^T \sigma\, dW_s \\ &+ \int_t^{T\wedge\tau} \langle \nabla u(s,X(s)), n(X(s)) \rangle\, d\xi_s. \end{aligned}$$

Taking conditional expectation on $X(t) = x$ yields

$$\begin{aligned} E\left[u(T\wedge\tau, X_{T\wedge\tau})|X(t)=x\right] &= E\left[u(t,X(t))|X(t)=x\right] + \\ &+ E\left[\int_t^{T\wedge\tau}\left(Lu(s,X(s)) + \frac{\partial u}{\partial t}(s,X(s))\right)\,ds\,\middle|\,X(t)=x\right] + \\ &+ E\left[\int_t^{T\wedge\tau}(\nabla u(s,X(s)))^T\sigma\,dW_s\,\middle|\,X(t)=x\right] + \\ &+ E\left[\int_t^{T\wedge\tau}\langle\nabla u(s,X(s)), n(X(s))\rangle\,d\xi_s\,\middle|\,X(t)=x\right]. \end{aligned} \tag{106}$$

The term on the left can be represented as sum:

$$\begin{aligned} &E\left[u(T\wedge\tau, X_{T\wedge\tau})|X(t)=x\right] = \\ &\quad = E\left[\mathbf{1}_{\tau\geqslant T}u(T\wedge\tau, X_{T\wedge\tau}) + \mathbf{1}_{\tau<T}u(T\wedge\tau, X_{T\wedge\tau})|X(t)=x\right] = \\ &\quad\quad = E\left[\mathbf{1}_{\tau\geqslant T}f(X_T) + \mathbf{1}_{\tau<T}g(X_\tau)|X(t)=x\right], \end{aligned}$$

because when $\tau \geqslant T$, $u(T\wedge\tau, X_{T\wedge\tau}) = u(T, X_T) = f(X_T)$, else $X_{T\wedge\tau} = X_\tau$ and this implies $X_\tau \in \partial_a D$ what yields $u(\tau, X_\tau) = g(X_\tau)$.

The first term on the right is actually $u(t,x)$ itself:

$$E\left[u(t,X(t))|X(t)=x\right] = u(t,x).$$

Due to PDE in problem (105)

$$Lu(s,X(s)) + \frac{\partial u}{\partial t}(s,X(s)) = 0$$

and the second term on the right in (106) vanishes.

Under conditions of the theorem the process

$$\int_t^{T\wedge\tau}(\nabla u(s,X(s)))^T\sigma\,dW_s$$

is a martingale, thus

$$E\left[\int_t^{T\wedge\tau}(\nabla u(s,X(s)))^T\sigma\,dW_s\middle|X(t)=x\right] = \int_t^t(\nabla u(s,X(s)))^T\sigma\,dW_s = 0.$$

Since $d\xi_s$ is concentrated on $\{X(s)\in\partial_r D\}$, in the integral

$$\int_t^{T\wedge\tau}\langle\nabla u(s,X(s)), n(X(s))\rangle\,d\xi_s$$

$\langle\nabla u(s,X(s)), n(X(s))\rangle$ can be replaced by $h(X(s))$ and we get

$$E\left[\int_t^{T\wedge\tau}\langle\nabla u(s,X(s)), n(X(s))\rangle\,d\xi_s\,\middle|\,X(t)=x\right] = E\left[\int_t^{T\wedge\tau}h(X(s))\,d\xi_s\,\middle|\,X(t)=x\right].$$

Gathering these things altogether, we finally obtain

$$E\left[\mathbf{1}_{\tau\geqslant T}f(X_T) + \mathbf{1}_{\tau<T}g(X_\tau)|X(t)=x\right] = u(t,x) + E\left[\int_t^{T\wedge\tau}h(X(s))\,d\xi_s\,\middle|\,X(t)=x\right]$$

and this concludes the proof.

8.2 STRONG APPROXIMATION OF STOPPED HITTING TIMES

In this section we consider a non-empty set $D \subseteq \mathbb{R}^d$ as well as a time-homogeneous strong Markov process $(X(t))_{t \geqslant 0}$ with càdlàg paths and a sequence of time-homogeneous Markov processes $(X_n(t_k))_{k \in \mathbb{N}_0}$, where $t_k = k \cdot T/n$ for $T > 0$ and $n \in \mathbb{N}$. The latter processes are extended to continuous-time processes by $X_n(t) = X_n(t_k)$ for $t \in]t_k, t_{k+1}[$. All these processes take values in $\overline{D}$, and they are defined on a common filtered probability space that satisfies the usual conditions. Let

$$\tau = \inf\{t \geqslant 0 : X(t) \in D_a\}$$

and

$$\tau_n = \inf\{t \geqslant 0 : X_n(t) \in D_a\}$$

denote the corresponding hitting times for an open set $D_a \subset \overline{D}$. Throughout this section we assume that

$$P(1_{(\tau<\infty)}) = 1 \tag{107}$$

and

$$P(1_{(\tau_n<\infty)}) = 1 \tag{108}$$

for every $n \in \mathbb{N}$.

To study the strong approximation of $\tau \wedge T$ by $\tau_n \wedge T$ we put

$$a_n = E\left(\sup_{t \in [0,T]} |X(t) - X_n(t)|\right),$$

which is the error of strong approximation of X by X_n on $[0, T]$. Furthermore, we define

$$u(s, x) = E^x((\tau + s) \wedge T) \tag{109}$$

as well as

$$u_n(s, x) = E^x((\tau_n + s) \wedge T)$$

for $s \in [0, T]$ and $x \in \overline{D}$. Thus $u(s, x) - u_n(s, x)$ is the bias of $\tau_n \wedge (T - s)$ as an approximation to $\tau \wedge (T - s)$ with initial value x in both cases. We put

$$b_n = \sup_{s \in [0,T], x \in \overline{D}} |u(s, x) - u_n(s, x)|.$$

We follow the approach from Higham *et al.* (2011) to derive the following result on strong approximation of the stopped hitting time $\tau \wedge T$.

Theorem 8.2.1 *Assume that there exists a constant* $c > 0$ *such that*

$$|u(s, x) - u(s, y)| \leqslant c \cdot |x - y|, \qquad s \in [0, T],\ x, y \in \overline{D}. \tag{110}$$

Then

$$E\,|\tau \wedge T - \tau_n \wedge T|^p \leqslant 2 \cdot T^{p-1} \cdot (c\, a_n + b_n)$$

holds for all $n \in \mathbb{N}$ *and* $1 \leqslant p < \infty$.

Proof At first we study the case $p = 1$. Let $(\mathfrak{F}_t)_{t \in [0,T]}$ denote the underlying filtration, and put $\widetilde{a} = a \wedge T$.

To control the case $\widetilde{\tau} > \widetilde{\tau_n}$, i.e., $\widetilde{\tau} > \tau_n$, we use

$$E((\widetilde{\tau} - \tau_n) \cdot 1_{\{\widetilde{\tau} > \tau_n\}}) = E(E((\widetilde{\tau} - \tau_n) \cdot 1_{\{\widetilde{\tau} > \tau_n\}} \mid \mathfrak{F}_{\tau_n})).$$

Note that $\{\widetilde{\tau} > \tau_n\} \in \mathfrak{F}_{\tau_n}$. The strong Markov property of X yields

$$E((\widetilde{\tau} - \tau_n) \cdot 1_{\{\widetilde{\tau} > \tau_n\}} \mid \mathfrak{F}_{\tau_n}) = 1_{\{\widetilde{\tau} > \tau_n\}} \cdot (u(\widetilde{\tau_n}, X(\tau_n)) - \tau_n).$$

Use $u(s,x) = s$ for $x \in D_a$ and (108) to derive

$$1_{\{\tilde{\tau} > \tau_n\}} \cdot \tau_n = 1_{\{\tilde{\tau} > \tau_n\}} \cdot u(\widetilde{\tau_n}, X_n(\tau_n))$$

and

$$u(\widetilde{\tau_n}, X(\tau_n)) \geqslant u(\widetilde{\tau_n}, X_n(\tau_n)).$$

It follows that

$$\begin{aligned} E((\tilde{\tau} - \tau_n) \cdot 1_{\{\tilde{\tau} > \tau_n\}} \mid \mathfrak{F}_{\tau_n}) &\leqslant u(\widetilde{\tau_n}, X(\tau_n)) - u(\widetilde{\tau_n}, X_n(\tau_n)) \\ &= u(\widetilde{\tau_n}, X(\widetilde{\tau_n})) - u(\widetilde{\tau_n}, X_n(\widetilde{\tau_n})), \end{aligned}$$

and therefore (110) implies

$$E((\tilde{\tau} - \tau_n) \cdot 1_{\{\tilde{\tau} > \tau_n\}}) \leqslant c \cdot E\,|X(\widetilde{\tau_n}) - X_n(\widetilde{\tau_n})| \leqslant c\, a_n. \tag{111}$$

To control the case $\widetilde{\tau_n} > \tilde{\tau}$ we apply the same consideration to X_n. Use $u_n(s,x) = s$ for $x \in D_a$ and (107) to derive

$$\begin{aligned} E((\widetilde{\tau_n} - \tau) \cdot 1_{\{\widetilde{\tau_n} > \tau\}} \mid \mathfrak{F}_{\tau}) &\leqslant u_n(\tilde{\tau}, X_n(\tilde{\tau})) - u_n(\tilde{\tau}, X(\tilde{\tau})) \\ &= A_1 + A_2 + A_3 \end{aligned}$$

with

$$\begin{aligned} A_1 &= u_n(\tilde{\tau}, X_n(\tilde{\tau})) - u(\tilde{\tau}, X_n(\tilde{\tau})), \\ A_2 &= u(\tilde{\tau}, X_n(\tilde{\tau})) - u(\tilde{\tau}, X(\tilde{\tau})), \\ A_3 &= u(\tilde{\tau}, X(\tilde{\tau})) - u_n(\tilde{\tau}, X(\tilde{\tau})). \end{aligned}$$

As previously, $E\,|A_2| \leqslant c\, a_n$ follows from (110), and by definition $|\,E(A_i)| \leqslant b_n$ for $i = 1, 3$. Hence

$$E((\widetilde{\tau_n} - \tau) \cdot 1_{\{\widetilde{\tau_n} > \tau\}}) \leqslant c\, a_n + 2\, b_n. \tag{112}$$

Combine (111) and (112) to complete the proof in the case $p = 1$.
Use $|\tilde{\tau} - \widetilde{\tau_n}|^p \leqslant |\tilde{\tau} - \widetilde{\tau_n}| \cdot T^{p-1}$ to extend this result to the case $p > 1$.

Now we consider the setting of Section 5.1, and we put $D_a = \partial_a D$. For the approximation of X on $[0, T]$ we employ the Euler scheme X_n with a constant step-size T/n and with the projection π to take care of the reflection. By definition, $X_n(0) = x_0$ and

$$X_n(t_{k+1}) = \pi\big(X_n(t_k) + T/n \cdot \mu(X_n(t_k)) + \sigma(X_n(t_k)) \cdot (W(t_{k+1}) - W(t_k))\big).$$

We combine Theorem 8.2.1 with upper bounds for a_n and b_n in the case of the projected Euler scheme, which are due to Słomiński (2001) and Costantini *et al.* (1998).

Corollary 1 *Consider the projected Euler scheme* X_n *for a stochastic differential equation with normal reflection. In addition to* (107) *and* (108)*, assume that*

(A1) $u \in C^{1,2}([0,T] \times \overline{D})$,

(A2) $\frac{\partial}{\partial s} u$ *is Hölder continuous with exponent* $1/2$,

(A3) $\frac{\partial^2}{\partial x_i \partial x_j} u$ *is Lipschitz continuous for all* $1 \leqslant i, j \leqslant d$.

Let $1 \leqslant p < \infty$. *Then there exists a constant* $c > 0$ *such that*

$$(E\,|\tau \wedge T - \tau_n \wedge T|^p)^{1/p} \leqslant c \cdot (\ln n / n)^{1/(4p)}$$

holds for every $n \geqslant 2$.

Proof Clearly (A1) implies (110). Moreover,

$$a_n \leqslant c \cdot (\ln n/n)^{1/4}$$

for every $n \geqslant 2$ with a constant $c > 0$, see Słomiński (2001, Thm. 3.2). Let $0 < \varepsilon < 1/2$. Under the assumptions (A1)-(A3),

$$b_n \leqslant c \cdot 1/n^{1/2-\varepsilon}$$

for every $n \geqslant 2$ with a constant $c > 0$, see Costantini *et al.* (1998, Thm. 3.6). Apply Theorem 8.2.1.

Let us comment on the assumptions that are imposed in Corollary 1.

Remark 15 *A sufficient condition for* (108) *to hold is*

$$\det \sigma(x) \neq 0$$

for every $x \in \overline{D}$ *and*

$$\lambda(\pi^{-1}(\partial_a D)) > 0,$$

where λ *denotes the* d*-dimensional Lebesgue measure. In this case the probability to reach* $\partial_a D$ *in one step of the projected Euler scheme is bounded away from zero, uniformly over all starting points* $x \in \overline{D}$.

Remark 16 *Now we turn to the assumptions (A1)-(A3). Let* L *denote the infinitesimal generator associated to* $dX(t) = \mu(X(t))\,dt + \sigma(X(t))\,dW(t)$, *and let* $\frac{\partial}{\partial n}$ *denote the normal derivative at any point in* $\partial_r D$. *Consider the mixed boundary value problem*

$$\begin{cases} \frac{\partial}{\partial t}u + Lu = 0 & \text{on } [0,T[\times D, \\ u(t,\cdot) = t & \text{on } \partial_a D \text{ for } t \in [0,T[, \\ \frac{\partial}{\partial n}u(t,\cdot) = 0 & \text{on } \partial_r D \text{ for } t \in [0,T[, \\ u(T,\cdot) = T & \text{on } \overline{D}. \end{cases} \tag{113}$$

For any classical solution to (113) *in* $C^{1,2}([0,T] \times \overline{D})$ *we have the representation* (109), *see Costantini* et al. *(1998, Thm. 2.5). We therefore have (A1)-(A3), if a classical solution exists and is sufficiently smooth.*

8.3 MATHEMATICAL MODEL

The Navier-Stokes-Brinkman system is governing the flow in the channel and in the porous membrane. The equations read as follows:

$$\begin{aligned} &\frac{\partial \vec{V}}{\partial t} - \nabla \cdot \left(\nu \nabla \vec{V}\right) + \left(\vec{V}, \nabla\right)\vec{V} + \nu K^{-1} + \frac{1}{\rho}\nabla p = \vec{f} \\ &\nabla \cdot \vec{V} = 0 \end{aligned} \tag{114}$$

Here $\vec{V} = (V_1, V_2)$, p, ν and ρ are the velocity, the pressure, the viscosity and the density of the fluid, respectively. Further on, K is the permeability of the membrane, which is considered as a porous medium here. In the pure fluid region K is set to infinity, which cancels the corresponding term, and (114) reduces to the Navier-Stokes system. On the other hand, due to the slow flow in the domain D, we will consider Stokes system (1) instead of system (114). This simplification does not change the nature of our numerical simulations, but in the future we may consider approach with (114) too.

The particle transport in the pure fluid region $D \subset \mathbb{R}^2$ or $D \subset \mathbb{R}^3$ is governed by the velocity field and by a stochastic term involving a 2-dimensional (3-dimensional) Brownian motion $\vec{W} = (W_1, W_2)$ ($\vec{W} = (W_1, W_2, W_3)$). The particle position $\vec{X} = (X_1, X_2)$ ($\vec{X} = (X_1, X_2, X_3)$) is a stochastic process with continuous paths taking values in $\overline{D}$. This process is described by an SDE (100)-(101), driven by $\vec{W}$, with drift coefficient $\mu = \vec{V}$ and with a constant diffusion coefficient $\hat{\sigma} = \text{diag}(\sigma) \in \mathbb{R}^{2\times 2}$ or $\hat{\sigma} = \text{diag}(\sigma) \in \mathbb{R}^{3\times 3}$ for some diffusion parameter $\sigma > 0$. Since particles cannot penetrate the membrane, we consider all boundaries of the channel to be reflecting ones, except the inflow and the outflow boundaries.

9

NUMERICAL SIMULATIONS

In this chapter we will present results of our numerical experiments for different AF4 devices. From one side, only setting from Section 9.1 satisfy all the requirements from the Part iii and Chapter 8. Also we have analytical results only in the setting of the Section 9.1. On the other hand, the main goal of our experiments in Section 9.2 and 9.3 was to obtain qualitative results, which would allow the reduce the number of the Lab experiments.

9.1 RECTANGULAR GEOMETRY

We will repeat the solution for (1) in a rectangular domain (see Figure 4):

$$V_1 = -6(u_1+u_2)\frac{x_1}{L}\cdot\frac{x_2}{w}\left(1-\frac{x_2}{w}\right)+6u_1\frac{x_2}{w}\left(1-\frac{x_2}{w}\right), \tag{115}$$

$$V_2 = -(u_1+u_2)\left(1-3\left(\frac{x_2}{w}\right)^2+2\left(\frac{x_2}{w}\right)^3\right)\cdot\frac{w}{L} \tag{116}$$

The streamlines of the solution (115)-(116) can be seen on figure 25.

Consider an SDE (100)-(101) with $V(X)$ given by (115)-(116), $\sigma = \sqrt{2\cdot D}$, initial position $X(0) = \left(0.01, \frac{u_1+u_2}{D}\cdot\frac{w}{L}\right)$. Domain D is chosen as a rectangle with length $L = 0.174$ and width $w = 290\cdot 10^{-6}$ meters. Diffusion coefficient is $D = 72.4\cdot 10^{-12}$ meters^2 per second and corresponds to the particles with a hydrodynamic radius 3.2 nanometers.

Now one can derive an estimate for the median of exit times distribution (see [58]) as:

$$t_r = \frac{1-\exp\left\{-\frac{u_0 w}{D}\right\}}{\exp\left\{-\frac{u_0 w}{D}\right\}\left(1+\frac{2D}{u_0 w}\right)+1-\frac{2D}{u_0 w}}\cdot\frac{w^2}{6D}\cdot\ln\left(\frac{\frac{z_0}{L}-\frac{\langle v\rangle w}{u_0 L}}{1-\frac{\langle v\rangle w}{u_0 L}}\right), \tag{117}$$

where $u_0 = (u_1+u_2)\cdot\frac{w}{L}$ and $\langle v\rangle = u_1$. For $u_0 = 2\cdot 10^{-5}$ and $\langle v\rangle = 18.7\cdot 10^{-3}$ we have $t_r = 186.8022$ seconds, while our numerical experiments with $\varepsilon = 10^{-2}$ give us result $t_r = 186.9410$. The cumulative distribution for this case can be seen on Figure 26.

9.2 HOLLOW FIBER DEVICE

Now we will consider a hollow fiber (HF) device. We will estimate the retention time, with respect to the flow parameters and estimate the fractogram.

Focusing-injection stage. An HF geometry and a sketch of flow directions during the focusing-injection stage can be seen in Figure 27. It should be noted that different scales for different parts of the device are used in the sketch. The axisymmetric HF, which is considered here, is 173 mm long itself (see the middle of the sketch), while the left capillary (serving to connect the HF to the sample reservoir) is 1300 mm long. During the focusing-injection stage, the solvent is entering the Hollow Fiber from both sides, and leaves through the membrane (see Figure 27). The ratio between the left and the right volumetric fluxes determines the position of the focusing line. The particles are injected from the left side during

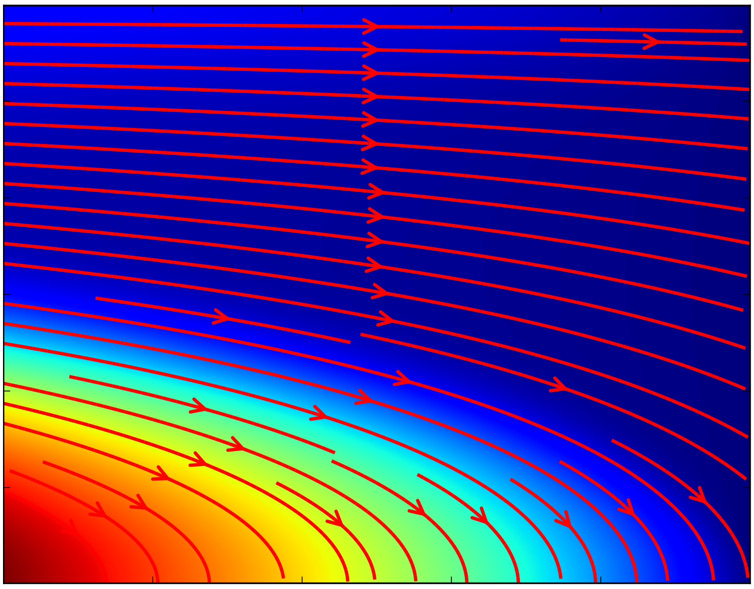


Figure 25: Domain D of the system and the streamlines of the velocity.

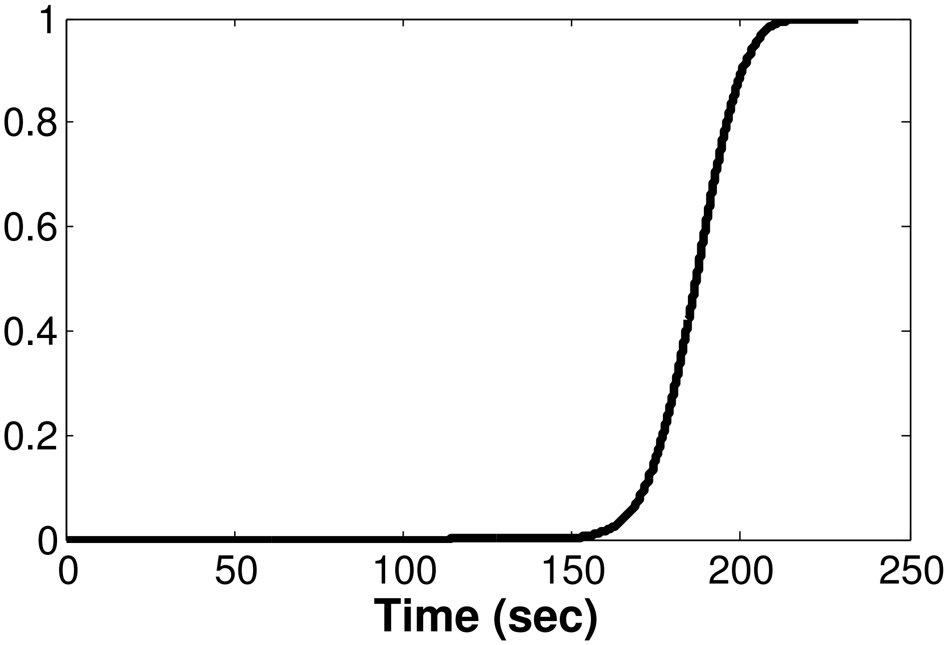


Figure 26: CDF for exit times

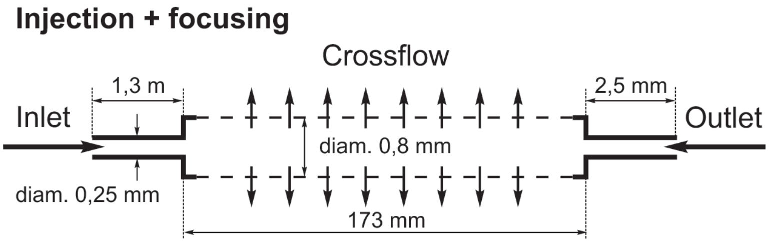

Injection + focusing
Crossflow
1,3 m
2,5 mm
Inlet
Outlet
diam. 0,8 mm
diam. 0,25 mm
173 mm

Figure 27: Hollow Fiber device during the focusing-injection stage for AFFFF.

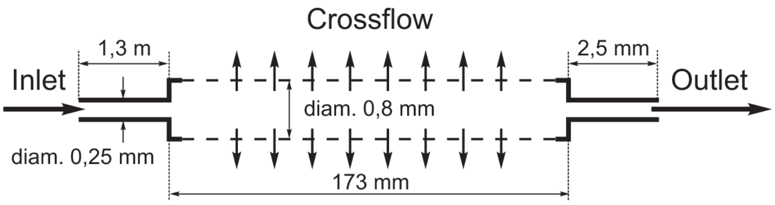


Figure 28: Hollow Fiber device during the Elution Stage for AFFFF.

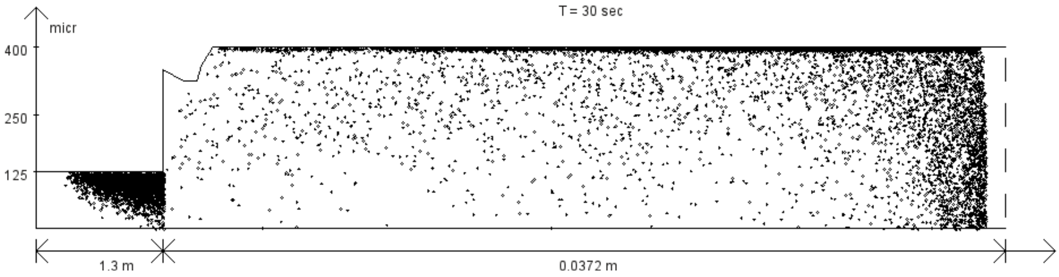


Figure 29: Particle injection, size $r = 3.2$ nm.

a certain time interval (shorter than the total duration of the focusing-injection stage), and they are transported towards the membrane due to the strong cross flow.

Elution stage. A sketch of flow directions during the elution stage can be seen in Figure 28. Due to the parabolic profile of the horizontal velocity, a horizontal separation of the particles is achieved at a certain distance from the focusing line. During the elution stage, smaller particles are transported faster than the bigger particles towards the channel outlet, because they experience a higher tangential flow velocity.

Simulation of the focusing-injection stage. The goal of the simulations, as defined by our industrial partner, is to study the dynamics of the injection of the particles, as well as the influence of the inflow control on the size and the shape of the focusing zone. Two snapshots of the injection process are shown here. Injection of smaller particles ($r = 3.2$ nm, Figure 29) and of larger particles ($r = 32$ nm, Figure 30) is simulated under the same flow conditions, and the snapshots are taken at the same time instance. Figure 29 and Figure 30 only show half of the HF cross section, due to symmetry. The final shape of the focusing zone, which is not shown here, can be controlled by the inflow velocity.

MLMC for restoring the diffusion coefficient at the focusing-elution stage. Let $\hat{X}_2$ denote the distance of a particle from the walls of the channel at the end of the focusing-injection stage. The cross flow $\hat{V}$ near the walls of the channel is almost a constant, so the distribution of the random variable $\hat{X}_2$ is well approximated by an exponential distribution with parameter $\hat{V}/D$ (see [12]), where $D = \sigma^2/2$ with

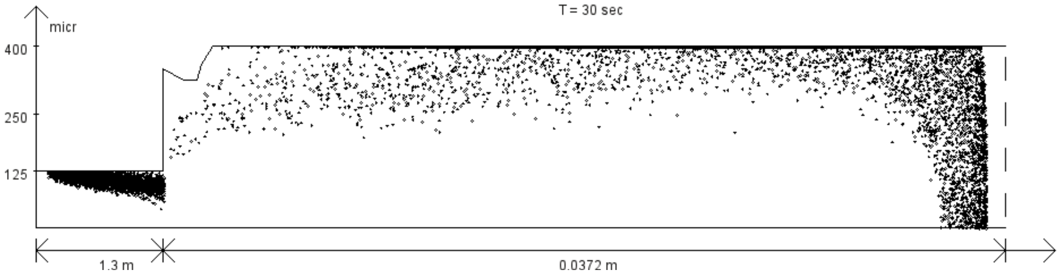


Figure 30: Particle injection, size $r = 32$ nm.

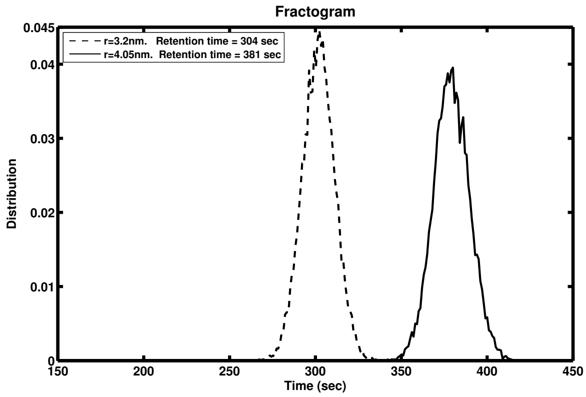


Figure 31: Fractogram for $r = 3.2$ nm and $r = 4.05$ nm, HF device

σ denoting the diffusion parameter of the SDE. Recall that D depends on the particle size. The latter distribution is also known as a barometric distribution.

In order to check the consistency of our approach, we have employed the MLMC algorithm to obtain an empirical distribution of $\hat{X}_2$, which is then used to provide an estimate for D. The estimate is compared to the known value of D, which is calculated based on particles radius (see [12]). Results from numerical experiments with the MLMC algorithm from Section 5.3 are presented in Table 2. *Elution stage.* The horizontal separation of the particles is done actually during the elution stage, and in practice the separation is evaluated from fractograms. Fractograms (see Figure 31) are a measure for the mass of particles that exit the outlet per unit time, and they serve as an approximation to the density of the distribution of the exit time of $\vec{X}$ from the channel. Ideally, the fractograms of particles of different size should be 'well separated'. Additionally, the so-called retention times, which are the medians of the fractograms, should be 'small', as this indicates that the total duration of the separation process is small. We have employed a Monte Carlo algorithm to obtain fractograms via simulation. A particular goal of our simulations is to study the properties of the fractograms depending on different focusing and elution regimes. The results are summarized in Table 3. The retention time is computed for particles with $r = 3.2$ nm (fifth column) and $r = 4.05$ nm (last column) under different focusing regimes (first and third columns) and different elution regimes (fourth column). These results and the respective fractograms, which are not shown here, allow to conclude whether the retention time can be reduced while keeping the fractograms well separated.

Desired accuracy of MLMC simulation	Absolute error for D	Relative error for D
$\varepsilon = 10^{-3}$	$4.0 \cdot 10^{-1}$	$5.7 \cdot 10^{-3}$
$\varepsilon = 10^{-4}$	$9.8 \cdot 10^{-2}$	$1.4 \cdot 10^{-3}$
$\varepsilon = 10^{-5}$	$4.2 \cdot 10^{-2}$	$5.9 \cdot 10^{-4}$

Table 2: Multilevel Monte Carlo method for restoring the coefficient $D = 76.2$ micron^2/sec. Cross flow equal to 30 micron/sec.

Parameter	Focusing [ml/min]	Elution [ml/min]	$t_{r=3.2}$	$t_{r=4.05}$
inlet	0.085	1.2	308	387
outlet	0.765	0.35		
crossflow	0.85	0.85		
inlet	0.17	1.2	291	366
outlet	0.68	0.35		
crossflow	0.85	0.85		
inlet	0.34	1.2	255	322
outlet	0.51	0.35		
crossflow	0.85	0.85		
inlet	0.17	0.775	160	201
outlet	0.255	0.35		
crossflow	0.425	0.425		

Table 3: Parameters of simulation (absolute values). Retention times (sec) of particles with radius 3.2 and 4.05 nanometers after 180 seconds of focusing.

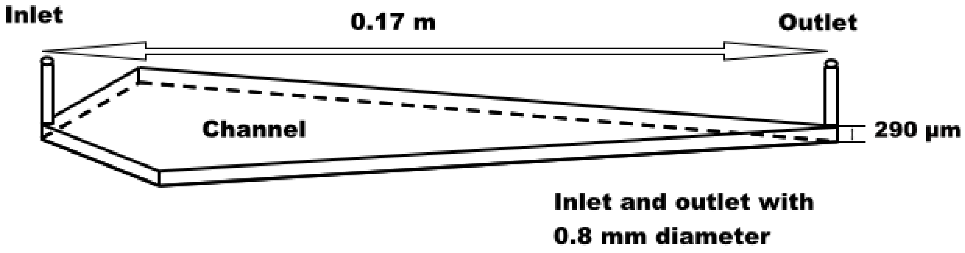


Figure 32: Sketch of the Eclipse device

9.3 ECLIPSE DEVICE

We consider now 3D geometries of the type, like one on the Figures 1 and 32. We intentionally do not present all the parameters from the simulation, except the ones, which are needed for the comparison with the analytical results.

Focusing-injection stage. A sketch of the focusing stage was presented on the Figure 2. In our experiments we assume, that the particles are initially situated in the domain, which is an intersection of the top wall and the inlet pipe. We distributed evenly 12000 particles in that domain.

We start the numerical simulations with injection of the particles, which last for either 4 seconds, either 120 seconds. After all the particles have been injected we continue the focusing stage, until the total time of the Focusing-injection stage is equal to 180 seconds.

Elution stage. Elution stage lasts until all the particles enter the outlet pipe. In general, one expect the average time to leave the channel to be well approximated by the retention time formula (117), which was derived for the simplified geometry. The goal of our experiment is to check if this is true and to check if retention time is affected by the length of the injection stage.

Remark 9.3.1 *We consider also another Eclipse geometry, but for the same experiments for the Focusing-injection and Elution stages. This geometry can be seen on Figure 33, and is very similar to the geometry*

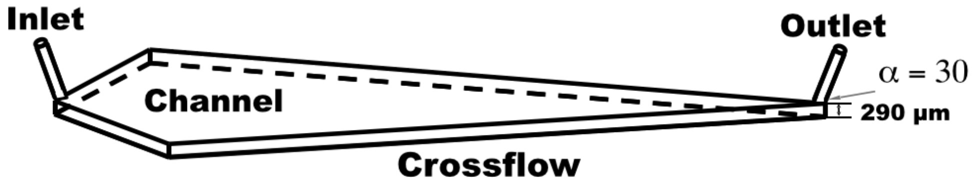


Figure 33: Sketch of the Eclipse device

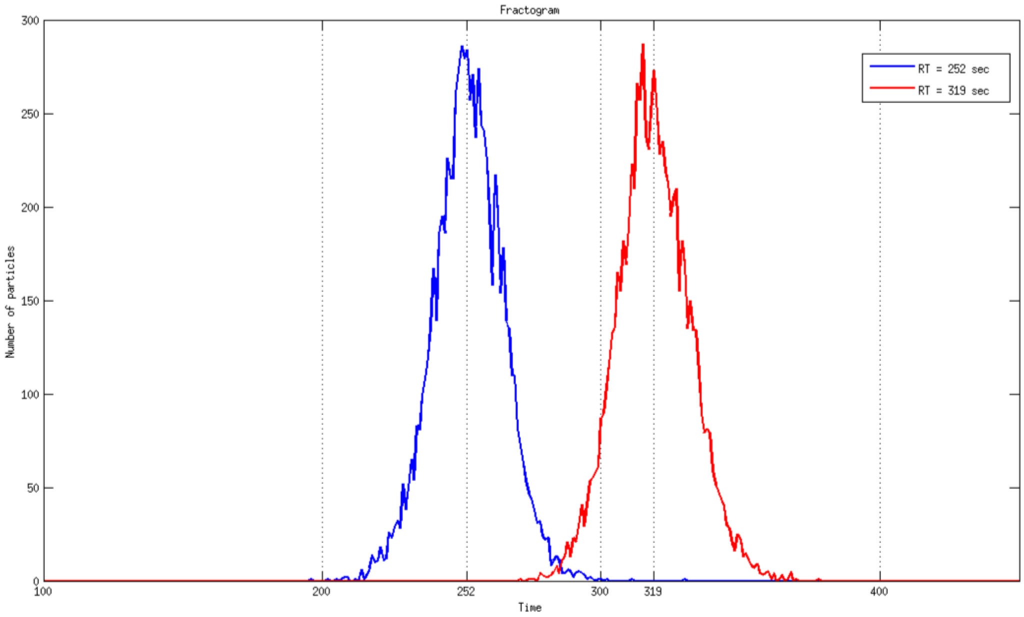


Figure 34: 90 degree geometry. 4 seconds injection.

on the Figure 32, despite the inlet and outlet pipes being under the angle of 30 degrees, with respect to the top wall.

Resulting Fractograms can be seen on Figures 34-37 .

Based on the we can deduce, that

- For a sufficiently long focusing time, the retention time practically does not depend on the angle
- In the case of angle 30 degrees, more particles stay close to side wall according to the current simulations, which results in a heavier tails at the fractograms.

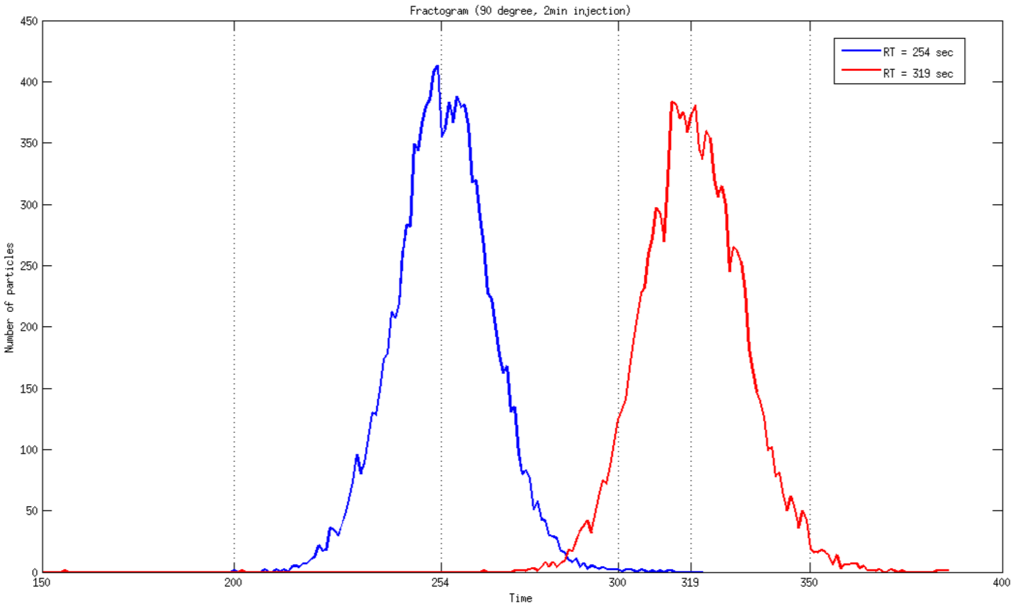


Figure 35: 90 degree geometry. 120 seconds injection.

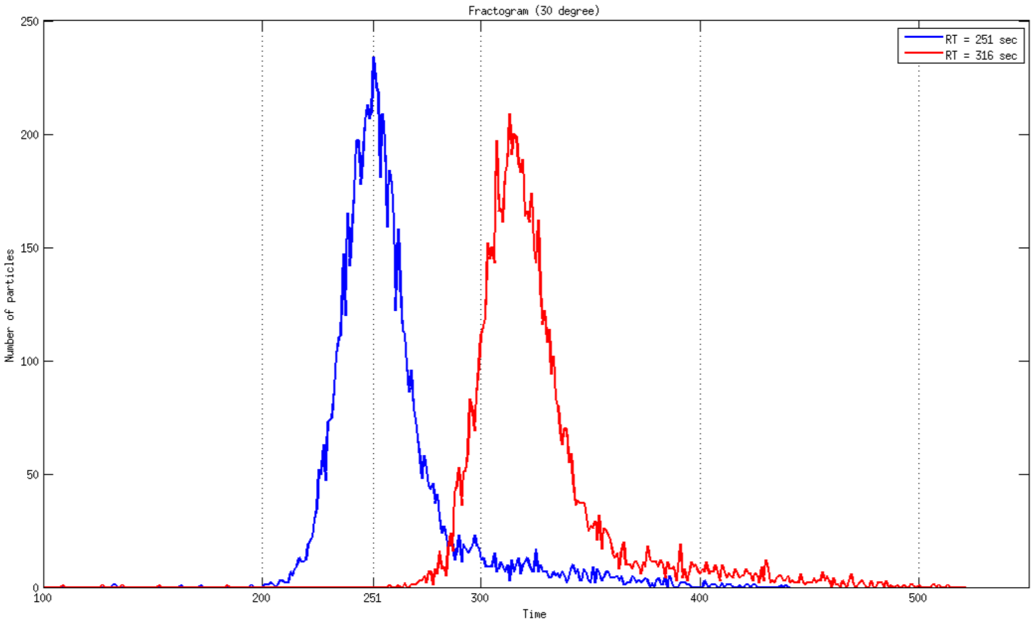


Figure 36: 30 degree geometry. 4 seconds injection.

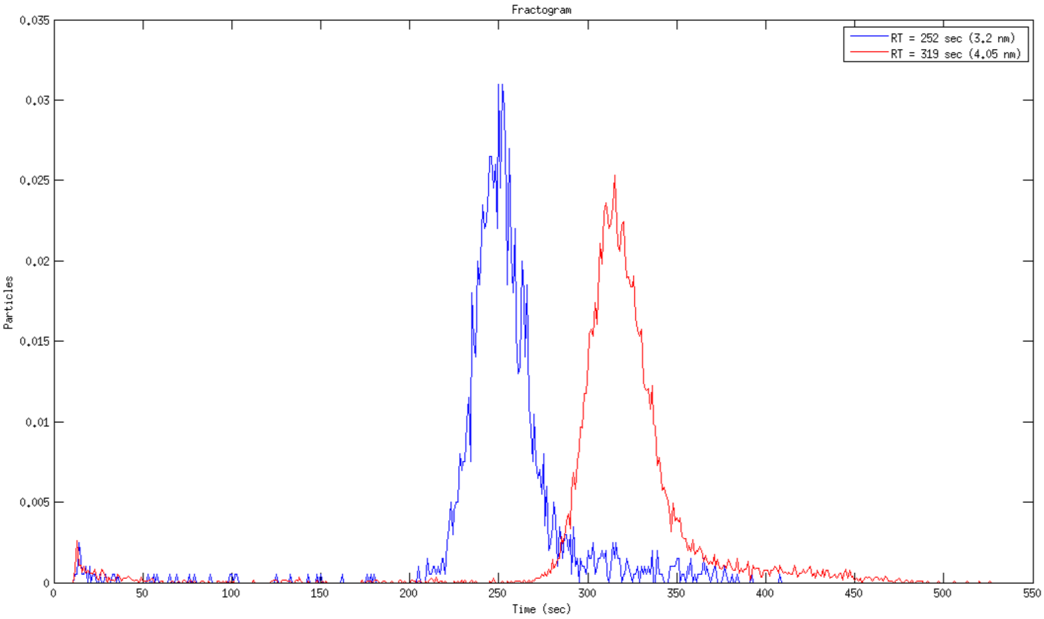


Figure 37: 30 degree geometry. 120 seconds injection.

Part V

REDUCED BASIS APPROACH FOR RETENTION TIME ESTIMATION

In this part of the thesis we present a model for AF4 based on SDE with reflections, or commonly known as Skorohod SDE, and numerical simulations, based on this model.

10

PDE APPROACH FOR RETENTION TIME ESTIMATION

Our long term goal is shape optimization of AFFFF device and optimal control of the fractionation process. In Chapter ii we have presented optimization for the focusing stage. In the future we will do the optimization of the elution stage, according to the requirements of our industrial partner, as well. At the end, the main goal is to have the best separation possible under the given constraints on the flow. The first step in that direction is to have the biggest difference in retention time for different particles. The goal of this chapter is to present important ingredients for optimization algorithms. More precisely, to present efficient two-grid finite volume reduced basis algorithm for retention time estimation.

We choose finite volume discretization, as most of the industrial codes, written for computational flow dynamics problems is written based on it. Another restriction is that most of those codes are locked, so one can not access and modify it, which means that our algorithm should also be of the "black-box" type.

10.1 MEAN EXIT TIME PDE

Let us denote by $u(x)$ the Mean Exit Time (MET) for the particle with respect to its initial position x. Then for a given flow $\vec{V}$ the equation for MET is (see [23])

$$\begin{cases} \mathcal{L}u + 1 = 0, & x \in \mathcal{A} \\ u\left(x\right) = 0, & x \in \partial_a \mathcal{A}, , \\ \dfrac{\partial u}{\partial n} = 0, & x \in \partial_r \mathcal{A}, \end{cases} \tag{118}$$

where $\mathcal{L}u = D\Delta u + \left\langle \vec{V}, \nabla u \right\rangle$, $\mathcal{A}$ is a channel domain (see Figure 1), $\partial_r \mathcal{A}$ is the reflection part of the boundary (blue color on the Figure 38), $\partial_a \mathcal{A}$ is an absorbing part of the boundary (red part on the Figure 38), and $\vec{V}$ is the solution of the Stokes system (1).

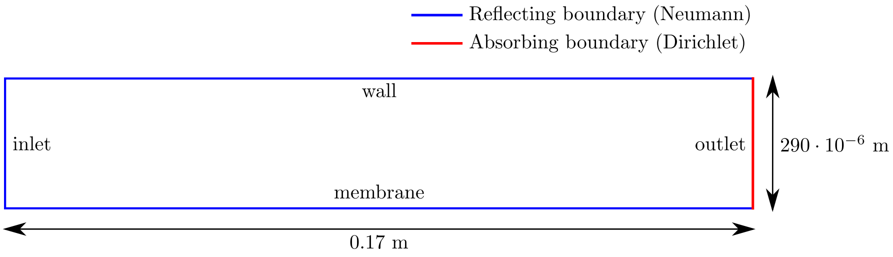


Figure 38: Domain and boundary types

10.2 REDUCED BASIS METHOD

Let $\mathcal{X}$ be a closed subspace of the Sobolev space $H^1(\mathcal{A})$ in a bounded domain $\mathcal{A} \in \mathbb{R}^n$ and B is a set of parameters. We consider the following problem: given $\mu \in B$ find $\eta \in \mathcal{X}$ such that for $\forall v \in \mathcal{X}$

$$a(\eta(\mu), v; \mu) = (f, v), \tag{119}$$

where $f \in L^2(\mathcal{A})$ and a is a bilinear form, corresponding to (118), continuous and coercive over $\mathcal{X}$ that depends additionally on a parameter $\mu \in B$. We assume, that for any μ, we can calculate $\eta(\mu)$ very accurate, but with very high computational cost. So instead of this, we will precalculate very accurately η for $\{\mu_1, \dots, \mu_n\}$, where n is small and then for any other $\hat{\mu} \in B$, we will approximate $\eta(\hat{\mu})$ by the linear combination of $\eta(\mu_1), \dots, \eta(\mu_n)$. Due to the fact, that we want to build low-dimensional subspace K_n to approximate the solutions set $\mathcal{K} \subset \mathcal{X}$, we need to have some regularity of the $\eta(\mu)$ in μ or hope that $\mathcal{K}$ has some kind of simple structure.

For more detailed explanation of Reduced Basis Method idea, we refer to e.g. [17], [46] and [51].

The usual Reduced basis method consists of two stages - offline and online. During the offline stage we prepare a basis — a set of solutions $\eta_1, \dots, \eta_n$ of the equation (119) calculated for a small number of parameters $\mu_1, \dots, \mu_n$, which will be used to approximate $\eta(\mu)$ for $\mu \neq \mu_1, \dots, \mu_n$

Remark 10.2.1 *As we have mentioned before, "black-box" type algorithms are often used in the industrial purposes. Two-grid reduced basis method allows to build such an algorithm, as it does not use the knowledge about the code, but operates only with the solutions.*

10.2.1 *Two-grid Reduced basis*

10.2.1.1 *Offline stage. Pure greedy version.*

Let η^h be the solution of (119) calculated on a fine grid, where h denotes the grid step. Denote by B_{basis} the set of parameters μ, which has been chosen as a basis for future estimations, and by $B_{\text{trial}} \subset B$, with high cardinality ($\#B_{\text{trial}} \gg 1$), the set of parameters, from which B_{basis} will be chosen. At the end, B_{basis} should provide us a possibility to estimate any $\eta(\mu) \in \mathcal{K}$ with accuracy ϵ.

- The basis is constructed iteratively. Assume, that we have already chosen the first k parameters $B^k_{\text{basis}} = \{\mu_1, \dots, \mu_k\}$ from the B_{trial}. Assume, that we have also calculated the fine solutions $\eta^h_1, \dots, \eta^h_k$ corresponding to $\{\mu_1, \dots, \mu_k\}$. Then we orthonormalize[1] the set of vectors $\{\eta^h_1, \dots, \eta^h_k\}$ and denote these elements as $\{\hat{\eta}^h_1, \dots, \hat{\eta}^h_k\}$.
- Then $\forall\, \mu_j \in B_{\text{trial}} \backslash B_{\text{basis}}$, find $\alpha = \{\alpha_1, \dots, \alpha_k\}$ such that:

  $$\left\| \eta^h(\mu_j) - \sum_{i=1}^{k} \alpha_i \cdot \hat{\eta}^h_i \right\|_{L^2} \to \min, \tag{120}$$

  After solving $\#B_{\text{trial}} - k$ minimization problems (120), we have errors $\varepsilon_1, \dots, \varepsilon_{\#B_{\text{trial}}-k}$.
- If $\max\limits_{j=1,\dots,\#B_{\text{trial}}-k} \varepsilon_j \leqslant \epsilon$, then stop. Else add μ_{k+1} to B_{basis} where

  $$\mu_{k+1} = \text{Argmax}\{\varepsilon_1, \dots, \varepsilon_{\#B_{\text{trial}}-k}\}.$$

  Then calculate η^h_{k+1}, corresponding to μ_{k+1}, along with $\hat{\eta}^h_{k+1}$.

---

[1] Note, that we do not solve here an eigenvalue problem as we orthogonalize elements in L^2 norm. This is a key difference from the method, proposed in [17]

10.2.1.2 *Two-grid online stage*

Two-grid online stage has been presented in [17] for the first time, but in the finite element setting. Here we will present it for a nonvariational setting. Let us denote by η^H the solution calculated on a coarse grid. Assume, that we have n basis elements in total: $(\hat{\eta}_1^h, \dots, \hat{\eta}_n^h)$.

- Choose $\mu \in B$ for which a computation is demanded.
- Calculate $\eta^H(\mu)$ and then find α_i, $i = 1, \dots, n$:

$$\alpha(\mu) = (\alpha_1(\mu), \dots, \alpha_n(\mu)) = \operatorname{Argmin} \left\| \eta^H(\mu) - \sum_{i=1}^{n} \alpha_i \cdot \hat{\eta}_i^h \right\|_{L^2}. \tag{121}$$

- Set $\eta_{RB}^h(\mu) = \sum\limits_{i=1}^{n} \alpha_i(\mu) \cdot \hat{\eta}_i^h$

10.2.1.3 *Postprocessing with the rectification matrix*

Let us denote by $\alpha^H(\mu) = (\alpha_1^H(\mu), \dots, \alpha_n^H(\mu))$ a vector of coefficients, which we obtain as a result of two-grid online stage. It is clear that these coefficients differ from the ones we could obtain, if we were mapping our basis directly to the solution on the fine grid.

So let us consider basis elements $\hat{\eta}_1^h, \dots, \hat{\eta}_n^h$ and corresponding them initial vectors $\eta_1^h, \dots, \eta_n^h$. Then we build matrices $A^h, A^H \in \mathbb{R}^{n \times n}$, such that the column i has the mapping coefficients of η_i^h and η_i^H respectively on basis elements. Now we can define a rectification matrix $\mathcal{R} = A^h \left(A^H\right)^{-1}$ and present a postprocessing algorithm:

- Define $\beta = \mathcal{R}\alpha$
- Set $\eta_{RB}^h(\mu) = \sum\limits_{i=1}^{n} \beta_i \cdot \hat{\eta}_i^h$

Remark 10.2.2 *This is so far an empirical approach, proposed at first in [17], which leads to great improvements in practice. One can notice, that the matrix $\mathcal{R}$ is built in the way, that*

$$\mathcal{R}A^H = A^h,$$

and can be seen as a linear approximation of the certain operator $\mathcal{B}$:

$$\mathcal{B}\colon \mathcal{B}\alpha^H(\mu) \to \alpha^h,$$

where α^H are coefficients, obtained as the result of the two-grid online stage, and α^h are the true ones.

Remark 10.2.3 *The problem (121) can be solved in many ways. In this particular paper, we have implemented cell-centered Finite volume discretization. Coarse grids were considered to be embedded into a fine grid, so the problem (121) was solved via aggressive coarsening (see Figure 39) of $\hat{\eta}_i^h$ onto a coarse grid. For our problem this approach has shown very good results (see the Section 10.3). On the other hand, different approach may be preferable for other problems.*

10.3 NUMERICAL RESULTS

In this section we will discuss RB method for (118). Consider the same geometry as in Section 9.1 We have set the following parameters (see Figure 38):

A1. Length - $L = 0.1$ meters

A2. Width - $w = 300 \cdot 10^{-6}$ meters

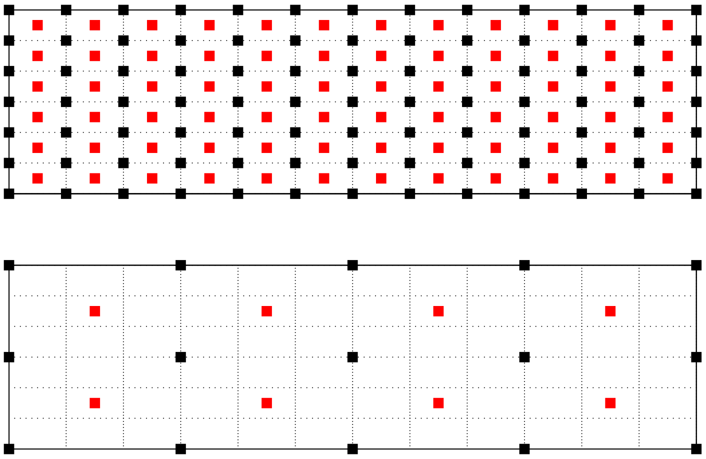


Figure 39: Initial and coarsened grids.

A3. $u_0 \in [2, 2.5] \cdot 10^{-5}$ meters per second

A4. $\langle v \rangle \in [2, 5] \cdot 10^{-2}$ meters per second.

Parameters u_0 and $\langle v \rangle$ are used to define the velocity $\vec{V}$

$$V_1(x_1, x_2) = 6 \cdot \frac{x_2}{w} \cdot \left(1 - \frac{x_2}{w}\right) \cdot \langle v \rangle \cdot \left(1 - \frac{u_0 L}{w \langle v \rangle} \frac{x_1}{L}\right). \tag{122}$$

$$V_2(x_2) = -u_0 \cdot \left(1 - 3 \cdot \left(\frac{x_2}{w}\right)^2 + 2 \cdot \left(\frac{x_2}{w}\right)^3\right). \tag{123}$$

One should notice, that (122)-(123) is the same velocity as (115)-(116), but written in the more convenient for this chapter way. The parameter $\langle v \rangle$ controls the inflow rate, while the parameter u_0, called crossflow, controls the outflow from the bottom of the channel.

In our numerical experiments, the fine grid contains 5625×450 elements.

10.3.1 *Numerical results for the offline stage*

We have chosen 31 uniformly distributed parameters from inflow parameter $\langle v \rangle$ and 16 parameters from outflow rate u_0. So this makes the number of parameters in B_{trial} set equal to 496 elements.

10.3.1.1 *Kolmogorov n-width*

Let $\mathcal{K}$ be a set of solutions $u(\mu)$ of (118), where $\mu \in B$. It is clear that $\mathcal{K}$ is a subset of some normed space $\mathbb{X}$, where $\mathbb{X}$ is L^2, H^1 or our variational space V. Assume, that we want to approximate $\mathcal{K}$ with K_n, which is the n-dimensional subspace of L^2. Then the Kolmogorov n-width (see [15]) is given by

$$d_n(\mathcal{K}, K_n) = \inf_{K_n} \sup_{x \in \mathcal{K}} \inf_{y \in K_n} \|x - y\|_{L^2}$$

and is a measure of the best possible approximation of the $\mathcal{K}$ with some n-dimensional subspace of L^2.

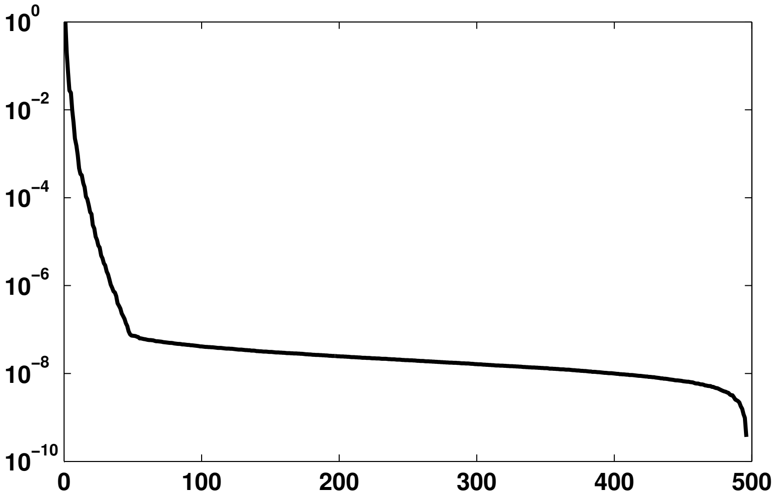


Figure 40: Kolmogorov n-width decay.

In order to estimate the decay of the Kolmogorov n-width we have built matrix $S = \{s_{i,j}\}$, $i, j = 1, \ldots 496$, where $s_{i,j} = \left\langle u\left(\mu_i\right), u\left(\mu_j\right)\right\rangle_{L^2}$ and plotted scaled eigenvalues.

As one can see on Figure 40, the scaled eigenvalues decay very fast, which means that the solution set $\mathcal{K}$ can be approximated very well by the finite dimensional space of low order.

10.3.1.2 *Pure greedy offline stage*

The error decay on B_{trial} depending on number of basis elements in the case of pure greedy offline stage can be seen in the Figure 41. The parameters and chosen according to the greedy algorithm from the Section 10.2.1.1 basis parameters can be seen on Figure 42. Red and green squares stand for the whole B_{trial} parameter space, while green squares correspond to the parameters, chosen for the B_{basis}

10.3.2 *Numerical results for the online stage*

For the tests for online stage we have chosen 30 uniformly distributed values for inflow parameter $\langle v \rangle$ and 15 values for outflow parameter u_0. This gives us in total 450 elements to test, which do not overlap with B_{trial}.

We have tested our 2-grid RB algorithm on 5 different grids: 625×450, 625×150, 375×450, 375×150 and 375×90 points with different number of basis elements.

10.3.3 *Retention time*

As it was explained in the Section 9.2, at the end of the focusing stage we have particles, which form boundary layer at the bottom of the channel. The distance $\hat{X}$ from the bottom of the channel is well approximated by an exponential distribution with parameter u_0/D. In [58] it was shown, that so-called retention time, which is the median for the distribution of exit times for the particle situated at the point

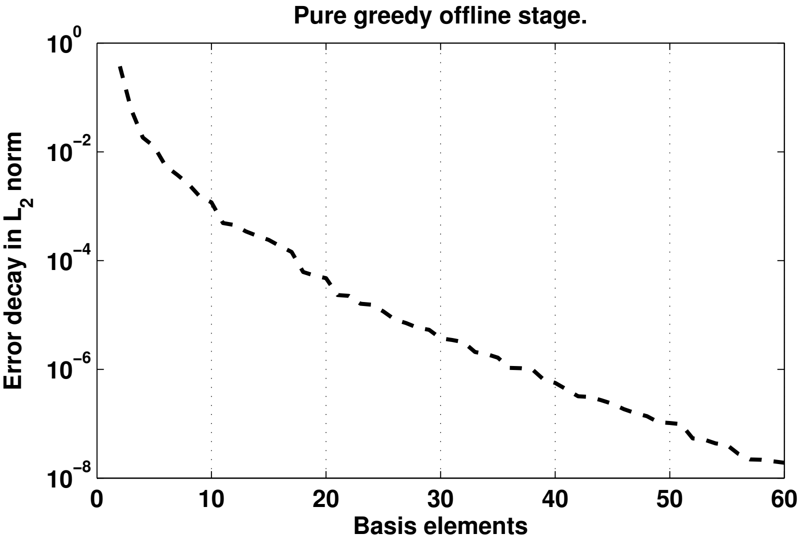


Figure 41: Max L^2 error on B_{trial} in dependence of number of basis elements.

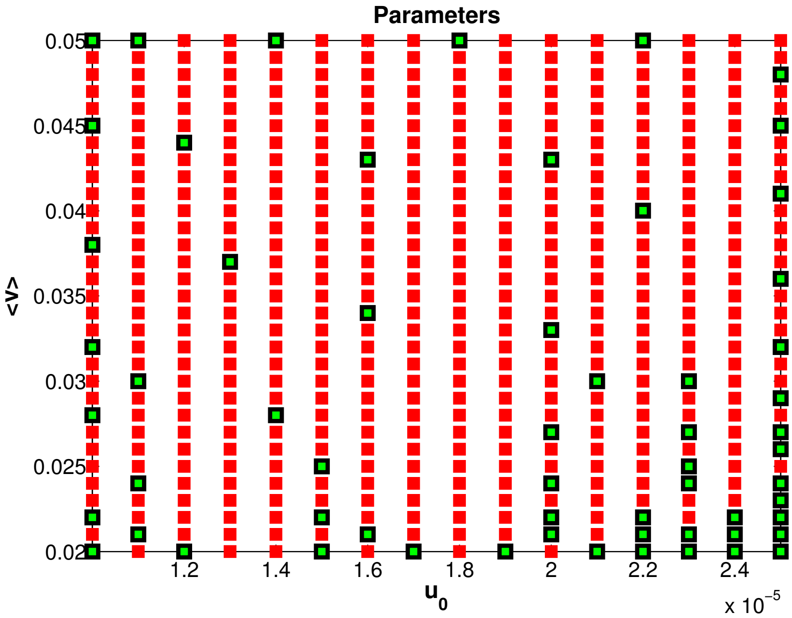


Figure 42: Parameters, chosen for the basis.

# basis	$\left\|u - u^h_{RB,rect}\right\|_{L^2}$ with postprocessing	$\left\|u - u^h_{RB}\right\|_{L^2}$ with no postprocessing	$\frac{\left\|u-u^h_{RB}\right\|_{L^2}}{\left\|u-u^h_{RB,rect}\right\|_{L^2}}$
10	$2.409e-04$	$2.409e-04$	1.0
20	$9.264e-06$	$9.298e-06$	1.005
30	$1.202e-06$	$1.510e-06$	1.378
40	$2.122e-07$	$1.310e-06$	6.25604

Table 4: Online stage test for coarse grid with 625 × 450 points

# basis	$\left\|u - u^h_{RB,rect}\right\|_{L^2}$ with postprocessing	$\left\|u - u^h_{RB}\right\|_{L^2}$ with no postprocessing	$\frac{\left\|u-u^h_{RB}\right\|_{L^2}}{\left\|u-u^h_{RB,rect}\right\|_{L^2}}$
10	$2.412e-04$	$2.837e-04$	1.35
20	$9.338e-06$	$1.233e-04$	18.44
30	$5.917e-07$	$1.232e-04$	225.44
40	$3.386e-07$	$1.232e-04$	376.52

Table 5: Online stage test for coarse grid with 625 × 150 points

$x = \left(z_0, \frac{u_0}{D}\right)$ has the exact representation (117). So, in 2D case in the rectangular geometry, in order to find the retention time from the initial distribution, one should find its center of the mass, which for the well focused concentration has the form $x = \left(z_0, \frac{u_0}{D}\right)$ and then calculate the retention time, according to (117). Due to the fact, that retention time and mean exit time in the 2D case are close to each other, we will compare their values.

First of all, we will compare two different estimations of the mean exit for the center of the mass situated at the point $\left(z_0, \frac{u_0}{D}\right)$, where $z_0 \in [0, L]$.

A1. Let us use the following parameters:
- $L = 0.17$ meters
- $w = 290 \cdot 10^{-6}$ meters
- $u_0 = 23 \cdot 10^{-5}$ meters per second
- $\langle v \rangle = 33 \cdot 10^{-2}$ meters per second
- $D = 72.4 \cdot 10^{-12}$ meters per squared seconds.

A2. The mean exit time for different z_0 is estimated via the reduced basis method, while the retention time is obtained from (117). We will compare those values for any $z_0 \in [0, L]$.

The results can be seen in Figure 43.

We can see, that there is a perfect match between the result of RB calculation and estimator, derived in [58]. On the other hand, we can obtain the same information using MET equation in different geometries (see f.e. a sketch of hollow fiber device on Figure 28) in 2D and in 3D.

# basis	$\left\lvert u - u^h_{RB,rect} \right\rvert_{L^2}$ with postprocessing	$\left\lvert u - u^h_{RB,rect} \right\rvert_{L^2}$ with no postprocessing	$\frac{\lvert u-u^h_{RB} \rvert_{L^2}}{\lvert u-u^h_{RB} \rvert_{L^2}}$
10	$2.409e-04$	$2.409e-04$	1.00008
20	$9.308e-06$	$9.392e-06$	1.01038
30	$1.515e-06$	$1.962e-06$	1.21532
40	$1.054e-06$	$2.357e-06$	1.91325

Table 6: Online stage test for coarse grid with 375 × 450 points

# basis	$\left\lvert u - u^h_{RB,rect} \right\rvert_{L^2}$ with postprocessing	$\lvert u - u^h_{RB} \rvert_{L^2}$ with no postprocessing	$\frac{\lvert u-u^h_{RB} \rvert_{L^2}}{\left\lvert u-u^h_{RB,rect} \right\rvert_{L^2}}$
10	$2.411e-04$	$2.836e-04$	1.351
20	$9.359e-06$	$1.230e-04$	18.056
30	$1.519e-06$	$1.229e-04$	83.657
40	$1.150e-06$	$1.230e-04$	114.416

Table 7: Online stage test for coarse grid with 375 × 150 points

10.3.4 *Future work on the optimization of the elution stage*

Optimization problems for AFFFF are of the great interest and Reduced Basis technique can dramatically reduce numerical complexity of these types of problems. There are different works devoted for optimization of AFFFF technique (see [10], [11]).

The estimation of the MET can also help for optimizing the elution stage, if the flow regime is considered to be constant. Also it can provide the guess for the initial control, if the control is considered to be time-dependent. To illustrate this idea we will consider a three dimensional problem, where we will solve (118), distribute particles in some zone and estimate a distribution of MET.

Assume that we have two types of particles with the same initial distribution with diffusion coefficients D_1 and D_2, such that $D_1 > D_2$. Let us define by $\varkappa_\alpha^{D_i}(u_0, \langle v \rangle)$ a time instance t as the first time, when α percent of the particle with diffusion coefficients D_i will leave the channel, if the velocity has parameters $(u_0, \langle v \rangle)$. It is clear, that for the particles with the same initial distribution and $D_1 > D_2$, we have $\varkappa_\alpha^{D_1}(u_0, \langle v \rangle) < \varkappa_\alpha^{D_2}(u_0, \langle v \rangle)$. So now we can present the following functional with optimization parameter $u = (u_0, \langle v \rangle)$:

Minimize

$$J(u) = (\varkappa_{50}^{D_2}(u) - \varkappa_{50}^{D_1}(u))^2 + \gamma_1 \cdot (\varkappa_{25}^{D_2}(u) - \varkappa_{75}^{D_1}(u))^2 \to \max$$

Subject to

$$\varkappa_5^{D_2}(u) - \varkappa_{95}^{D_1}(u) \geqslant \gamma_2,$$

where γ and γ_2 are certain coefficients.

# basis	$\left\|u - u^h_{RB,rect}\right\|_{L^2}$ with postprocessing	$\left\|u - u^h_{RB}\right\|_{L^2}$ with no postprocessing	$\frac{\left\|u - u^h_{RB}\right\|_{L^2}}{\left\|u - u^h_{RB,rect}\right\|_{L^2}}$
10	$2.429e-04$	$4.618e-04$	2.55709
20	$9.840e-06$	$3.686e-04$	51.0933
30	$1.691e-06$	$3.695e-04$	230.118
40	$2.317e-06$	$3.697e-04$	165.596

Table 8: Online stage test for coarse grid with 375 × 90 points

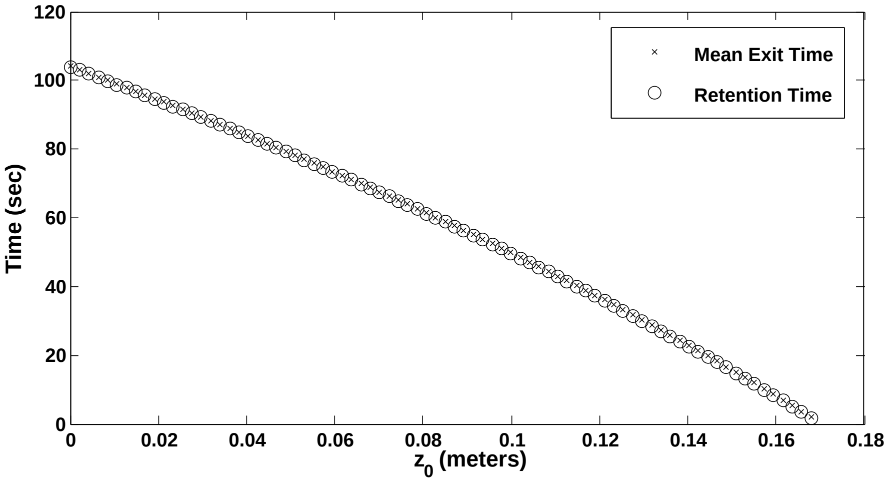


Figure 43: Mean Exit Time obtained via Reduced basis method and Retention time obtained via (117).

CONCLUSION

In Section 1.4 we have stated main problems, to cover in this work. Here is the short overview of the strategies, we used to resolve them.

Optimization of the focusing stage

Part ii of this dissertation is devoted to solving the optimization problem for the focusing stage of AF4. We introduced an optimization algorithm, which takes into the account the specifics of our problem and a goal-oriented functional, which assures, that the concentration is well focused by the end of the focusing process. We used sensitivity formulation instead of the adjoint one due to the special structure of our problem, but provided an explicit PDE for each component of the gradient. Our approach allows us to calculate each component of the gradient independently and in parallel with the solution of the main state equation. Another advantage of this approach is that there is no need to save previous values of the state $c(x, t)$ to compute the gradient, which already leads to great savings in memory in two-dimensional case. The results of this part have been published in [10] and [11].

Estimation of distribution function for exit times

Part iii introduces new Multilevel Monte Carlo algorithm for distribution functions approximation. While till now the standard task for Multilevel Monte Carlo algorithm was to compute the expectation of a real-valued functional, we discuss how to approximate a distribution or density function on a compact interval. In this part a general technique is proposed which is not restricted to the specifics of the AF4, and thus can be useful to a broader range of problems. We give a full description of the algorithm along with error and complexity analysis and numerical experiments. The results of this part will be published in [29].

SDE-based model for particles evolution in the separation channel

In Part iv we present a model for AF4 based on Stochastic differential equation with reflections, or commonly known as Skorohod SDE, and numerical simulations for different geometries. We prove strong convergence result for the exit times of the processes, driven by SDE with reflection. The model at first has been presented in [40]. Numerical experiments have been partially presented in [48]. Results from Chapter 8 will be presented in [39].

PDE based approach for retention time estimation

In Part v we present one-way coupled Stokes equation for the flow and PDE for the Mean Exit time and a modification of the modern two-grid reduced basis approach to the Finite Volume case. This approach allows us to build a black-box type computational scheme, which is important for our application, as in many cases one can not access the code of the industrial software, developed for computational fluid dynamics problems. Our numerical experiments suggest, that two-grid reduced basis method can be a very useful ingredient for the future work on the optimization of the Elution stage, as it is simple, reliable and fast.

BIBLIOGRAPHY

[1] Aurélien Alfonsi, Benjamin Jourdain, and Arturo Kohatsu-Higa. Pathwise optimal transport bounds between a one-dimensional diffusion and its euler scheme. *arXiv preprint arXiv:1209.0576*, 2012. (Cited on page 63.)

[2] M Altmayer and A Neuenkirch. Multilevel monte carlo quadrature of discontinuous payoffs in the generalized heston model using malliavin integration by parts. *Preprint*, 144, 2013. (Cited on page 39.)

[3] LJ Alvarez-Vázquez and A Martinez. Modelling and control of natural convection in canned foods. *IMA journal of applied mathematics*, 63(3):247–265, 1999. (Cited on page 9.)

[4] LJ Alvarez-Vázquez, M Marta, and A MARTinez. Sterilization of canned viscous foods: an optimal control approach. *Mathematical Models and Methods in Applied Sciences*, 14(03):355–374, 2004. (Cited on page 9.)

[5] R. Avikainen. On irregular functionals of SDEs and the Euler scheme. *Finance and Stochastics*, 13 (3):381–401, 2009. (Cited on pages 39, 54, and 60.)

[6] V Bally and D Talay. The law of the euler scheme for stochastic differential equations: Ii. aproximation of the density. *Monte Carlo Methods and Applications*, 2:93–128, 1996. (Cited on page 62.)

[7] Vlad Bally and Denis Talay. The law of the euler scheme for stochastic differential equations. *Probability theory and related fields*, 104(1):43–60, 1996. (Cited on pages 60, 61, and 62.)

[8] Christian Bayer, Anders Szepessy, and Raúl Tempone. Adaptive weak approximation of reflected and stopped diffusions. *Monte Carlo Methods and Applications*, 16(1):1–67, 2010. (Cited on page 75.)

[9] N Bayramov and J Kraus. On the robust solution of convection-diffusion equation coupled with stokes equation. *Computational and Applied Mathematics, (submitted)*, 2013. (Cited on pages 12, 17, 18, and 20.)

[10] Nadir Bayramov, Tigran Nagapetyan, and Rene Pinnau. Fast optimal control of asymmetric flow field flow fractionation processes. In *SIAM Conference on Control and Its Applications*, pages 207–213. Society for Industrial and Applied Mathematics, 2013. (Cited on pages 98 and 100.)

[11] Nadir Bayramov, Tigran Nagapetyan, and Rene Pinnau. Fast optimal control of the focusing stage of the asymmetric flow field flow fractionation processes. Manuscript submitted for publication, 2014. (Cited on pages 98 and 100.)

[12] Richard Becker. *Theorie der Wärme*. Springer, 1961. (Cited on pages 3, 4, 83, and 84.)

[13] Roland Becker and Boris Vexler. Optimal control of the convection-diffusion equation using stabilized finite element methods. *Numerische Mathematik*, 106(3):349–367, 2007. (Cited on page 9.)

[14] Bruno Bouchard, Stefan Geiss, and Emmanuel Gobet. First time to exit of a continuous it\^{o} process: general moment estimates and `l_1`-convergence rate for discrete time approximations. *arXiv preprint arXiv:1307.4247*, 2013. (Cited on pages 38 and 63.)

[15] Annalisa Buffa, Yvon Maday, Anthony T Patera, Christophe Prud?homme, and Gabriel Turinici. A priori convergence of the greedy algorithm for the parametrized reduced basis method. *ESAIM: Mathematical Modelling and Numerical Analysis*, 46(03):595–603, 2012. (Cited on page 94.)

[16] Karin D Caldwell, Steven L Brimhall, Yushu Gao, and J Calvin Giddings. Sample overloading effects in polymer characterization by field-flow fractionation. *Journal of applied polymer science*, 36(3): 703–719, 1988. (Cited on page 5.)

[17] Rachida Chakir and Yvon Maday. Une méthode combinée d'éléments finis à deux grilles/bases réduites pour l'approximation des solutions d'une edp paramétrique. *Comptes Rendus Mathematique*, 347(7):435–440, 2009. (Cited on pages 92 and 93.)

[18] Yap P Chuan, Yuan Y Fan, Linda Lua, and Anton PJ Middelberg. Quantitative analysis of virus-like particle size and distribution by field-flow fractionation. *Biotechnology and bioengineering*, 99(6): 1425–1433, 2008. (Cited on page 3.)

[19] J. Creutzig, S. Dereich, T. Müller-Gronbach, and K. Ritter. Infinite-dimensional quadrature and approximation of distributions. *Foundations of Computational Mathematics*, 9(4):391–429, 2009. (Cited on page 39.)

[20] Wolfgang Fraunhofer and Gerhard Winter. The use of asymmetrical flow field-flow fractionation in pharmaceutics and biopharmaceutics. *European journal of pharmaceutics and biopharmaceutics*, 58(2):369–383, 2004. (Cited on page 3.)

[21] Mark Iosifovich Freidlin. *Functional Integration and Partial Differnetial Equations*, volume 109. Princeton University Press, 1985. (Cited on page 76.)

[22] AV Fursikov, MD Gunzburger, and LS Hou. Optimal dirichlet control and inhomogeneous boundary value problems for the unsteady navier-stokes equations. In *ESAIM: Proceedings*, volume 4, pages 97–116. EDP Sciences, 1998. (Cited on page 9.)

[23] Crispin W Gardiner et al. *Handbook of stochastic methods*, volume 3. Springer Berlin, 1985. (Cited on page 91.)

[24] John Calvin Giddings, Martin E Schimpf, and Karin Caldwell. *Field-flow fractionation handbook*. Wiley-interscience, 2000. (Cited on pages 3 and 5.)

[25] M.B. Giles. Multilevel Monte Carlo path simulation. *Operations Research*, 56(3):607–617, 2008. (Cited on pages 35 and 39.)

[26] M.B. Giles. Improved multilevel Monte Carlo convergence using the Milstein scheme. In A. Keller, S. Heinrich, and H. Niederreiter, editors, *Monte Carlo and Quasi-Monte Carlo Methods 2006*, pages 343–358. Springer-Verlag, 2008. (Cited on page 39.)

[27] M.B. Giles, D.J. Higham, and X. Mao. Analysing multilevel Monte Carlo for options with non-globally Lipschitz payoff. *Finance and Stochastics*, 13(3):403–413, 2009. (Cited on page 39.)

[28] M.B. Giles, K. Debrabant, and A. Rößler. Numerical analysis of multilevel Monte Carlo path simulation using the Milstein discretisation. *ArXiv preprint: 1302.4676*, 2013. (Cited on page 39.)

[29] Mike Giles, Tigran Nagapetyan, and Klaus Ritter. Multilevel monte carlo approximation of distribution functions and densities. Manuscript submitted for publication, 2014. (Cited on page 100.)

[30] Paul Glasserman. *Monte Carlo methods in financial engineering*, volume 53. Springer, 2004. (Cited on page 33.)

[31] Emmanuel Gobet and Celine Labart. Sharp estimates for the convergence of the density of the euler scheme in small time. *Electronic Communications in Probability*, 13:352–363, 2008. (Cited on page 62.)

[32] Emmanuel Gobet and Stéphane Menozzi. Stopped diffusion processes: boundary corrections and overshoot. *Stochastic Processes and their Applications*, 120(2):130–162, 2010. (Cited on page 75.)

[33] Max D Gunzburger. *Perspectives in flow control and optimization*, volume 5. Siam, 2003. (Cited on page 9.)

[34] S. Heinrich. Monte Carlo complexity of global solution of integral equations. *Journal of Complexity*, 14(2):151–175, 1998. (Cited on pages 39 and 45.)

[35] D.J. Higham, X. Mao, M. Roj, Q. Song, and G. Yin. Mean exit times and the multi-level Monte Carlo method. *SIAM Journal on Uncertainty Quantification*, 1(1):2–18, 2013. (Cited on pages 39, 64, and 75.)

[36] M Hinze, R Pinnau, M Ulbrich, and S Ulbrich. *Optimization with PDE Constraints. Mathematical Modelling: Theory and Applications, vol. 23.* Springer Berlin, 2009. (Cited on pages 9, 14, and 15.)

[37] Michael Hinze. *Optimal and instantaneous control of the instationary Navier-Stokes equations.* Citeseer, 2000. (Cited on page 9.)

[38] Oleg Iliev and Vsevolod Laptev. On numerical simulation of flow through oil filters. *Computing and Visualization in Science*, 6(2-3):139–146, 2004. (Cited on page 9.)

[39] Oleg Iliev, Mike Giles, Tigran Nagapetyan, and Klaus Ritter. Multilevel monte carlo method for estimation distribution functions of exit times with application to asymmetric flow field flow fractionation. Manuscript in preparation. (Cited on page 100.)

[40] Oleg Iliev, Tigran Nagapetyan, and Klaus Ritter. Monte carlo simulation of asymmetric flow field flow fractionation. *Monte Carlo Methods and Applications: Proceedings of the 8th IMACS Seminar on Monte Carlo Methods, August 29–September 2, 2011, Borovets, Bulgaria*, 115(123):115, 2012. (Cited on page 100.)

[41] Jean Jacod and Philip Protter. Asymptotic error distributions for the euler method for stochastic differential equations. *The Annals of Probability*, 26(1):267–307, 1998. (Cited on page 62.)

[42] A. Kebaier. Statistical Romberg extrapolation: a new variance reduction method and applications to options pricing. *Annals of Applied Probability*, 14(4):2681–2705, 2005. (Cited on pages 39 and 61.)

[43] A. Kebaier and A. Kohatsu-Higa. An optimal control variance reduction method for density estimation. *Stochastic Processes and their Applications*, 118(2):2143–2180, 2008. (Cited on pages 39 and 61.)

[44] Pierre-Louis Lions and Alain-Sol Sznitman. Stochastic differential equations with reflecting boundary conditions. *Communications on Pure and Applied Mathematics*, 37(4):511–537, 1984. (Cited on page 74.)

[45] Weijiu Liu. Mixing enhancement in enzymatic chemical reactions by optimal tuning of flows. *Systems &amp; Control Letters*, 58(12):834–840, 2009. (Cited on page 9.)

[46] Alf Emil Løvgren, Yvon Maday, and Einar M Rønquist. The reduced basis element method for fluid flows. In *Analysis and simulation of fluid dynamics*, pages 129–154. Springer, 2007. (Cited on page 92.)

[47] Grigori N Milstein, John GM Schoenmakers, and Vladimir Spokoiny. Transition density estimation for stochastic differential equations via forward-reverse representations. *Bernoulli*, 10(2):281–312, 2004. (Cited on page 62.)

[48] Tigran Nagapetyan. Multilevel monte carlo method for approximation of distribution functions and an application to af4. In *Young Researcher Symposium (YRS) 2013*, pages 1–6. Innovationszentrum Applied System Modeling, 2013. (Cited on page 100.)

[49] B Pacchiarotti, C Costantini, and Flavio Sartoretto. Numerical approximation for functionals of reflecting diffusion processes. *SIAM Journal on Applied Mathematics*, 58(1):73–102, 1998. (Cited on page 75.)

[50] Alfio M Quarteroni and Alberto Valli. *Numerical approximation of partial differential equations*, volume 23. Springer, 2008. (Cited on page 12.)

[51] Gianluigi Rozza. Reduced basis approximation and error bounds for potential flows in parametrized geometries. *Communication in Computational Physics*, 9:1–48, 2011. (Cited on page 92.)

[52] Antonio Russo and Giulio Starita. On the existence of steady-state solutions to the navier-stokes system for large fluxes. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze-Serie IV*, 7 (1):171, 2008. (Cited on page 10.)

[53] Yasumasa Saisho. Stochastic differential equations for multi-dimensional domain with reflecting boundary. *Probability Theory and Related Fields*, 74(3):455–477, 1987. (Cited on page 75.)

[54] Steven E Shreve. *Stochastic Calculus for Finance II, Continuous Time Models.* Springer, 2008. (Cited on page 66.)

[55] Leszek Słomiński. Euler's approximations of solutions of sdes with reflecting boundary. *Stochastic processes and their applications*, 94(2):317–337, 2001. (Cited on page 75.)

[56] Denis Talay and Ziyu Zheng. Approximation of quantiles of components of diffusion processes. *Stochastic processes and their applications*, 109(1):23–46, 2004. (Cited on page 62.)

[57] K-G Wahlund and A Litzen. Application of an asymmetrical flow field-flow fractionation channel to the separation and characterization of proteins, plasmids, plasmid fragments, polysaccharides and unicellular algae. *Journal of Chromatography A*, 461:73–87, 1989. (Cited on page 3.)

[58] Karl Gustav Wahlund and J Calvin Giddings. Properties of an asymmetrical flow field-flow fractionation channel having one permeable wall. *Analytical chemistry*, 59(9):1332–1339, 1987. (Cited on pages 4, 5, 9, 81, 95, and 97.)

[59] Jinchao Xu and Ludmil Zikatanov. A monotone finite element scheme for convection-diffusion equations. *Mathematics of Computation of the American Mathematical Society*, 68(228):1429–1446, 1999. (Cited on pages 17, 18, and 20.)

[60] Puhong You, Zhonghai Ding, and Jianxin Zhou. Optimal boundary control of the stokes fluids with point velocity observations. *SIAM journal on control and optimization*, 36(3):981–1004, 1998. (Cited on page 9.)

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

`http://code.google.com/p/classicthesis/`

DECLARATION

I hereby declare and confirm that this thesis represents my own original work, except where it is otherwise indicated in the text. No material in this thesis has been submitted for the award of any degree at any other institution.

Kaiserslautern, February 11, 2014

Tigran Nagapetyan

CURRICULUM VITAE

09/1995-05/2005	Schulausbildung
09/2005-06/2010	Diploma in Applied Mathematics and Computer Science, Staatliche Universität Moskau, Russland
08/2010-01/2011	ProSat Programm, Technische Universität Kaiserslautern, Deutschland
seit 02/2011	Doktorand der Mathematik, Technische Universität Kaiserslautern, Deutschland

09/1995-05/2005	School education
09/2005-06/2010	The specialist degree in Applied Mathematics and Computer Science, Moscow State University, Russian Federation
08/2010-01/2011	ProSat Program, Technical University Kaiserslautern, Germany
from 02/2011	PhD student, Technical University Kaiserslautern, Germany