Communicated by: Prof. Dr. Raúl Tempone, Chair of Mathematics for Uncertainty Quantification

RWTH Aachen University

Master's Thesis

Chance Constrained Optimal Control of Stochastic Differential Equations

---

Author:
Manuel Kleinschmager

Examiners:
Prof. Dr. Raúl Tempone
Dr. Nadhir Ben Rached

Additional Supervisor:
Shyam Mohan Subbiah Pillai, M.Sc.

This thesis was submitted to the ***Chair of Mathematics for Uncertainty Quantification*** *in fulfillment of the requirements for the degree of Master of Science*

in

Mathematics



RWTH AACHEN UNIVERSITY

Abstract

Chair of Mathematics for Uncertainty Quantification

Department of Applied Mathematics

Master of Science

Chance Constrained Optimal Control of Stochastic Differential Equations

by Manuel KLEINSCHMAGER

This thesis addresses chance constrained continuous-time optimal control problems involving controlled stochastic differential equations over a finite time horizon. Chance constraints are used when corresponding almost sure constraints are infeasible or lead to extremely high optimal costs. In particular, we are investigating terminal state chance constraints. To solve these constrained problems, we extend the theory of Lagrangian relaxation to the domain of stochastic optimal control, leading to a relaxed optimization problem requiring the solution of a non-linear PDE with non-smooth boundary conditions. We propose a subgradient-based optimization procedure, which determines the optimal Lagrange multiplier for the dual problem, yielding the dual-optimal solution to the original chance-constrained problem. Additionally, we introduce a novel approach to estimate the relative duality gap in each iteration of the subgradient method using an auxiliary stochastic optimal control problem. This novel approach requires addressing another stochastic optimal control sub-problem similar to the relaxed problem, for which we develop a solution method based on a Taylor-like expansion around the optimal penalization parameter. Additionally, an algorithm that selects the optimal parameters to achieve a relative dual gap of $\mathrm{TOL}_{\mathrm{SG}} > 0$ is developed. To find a suitable initial value for the numerical optimization, we furthermore propose a bisection-based approach for solving SOC problems that deal with a single scalar terminal time chance constraint. We establish theoretical optimal computational work results for the procedures introduced which are verified numerically on two examples. These two examples also illustrate challenges such as solving non-linear PDEs with discontinuous boundary conditions. Using the relative dual gap estimation procedure, we achieve a significant speed-up of the subgradient method when compared to the classical convergence results of subgradient methods. Finally, we detect increasing optimal costs for lower violation probabilities in both examples.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my advisor and first examiner, Prof. Raúl Tempone, for his invaluable support and guidance throughout this thesis. His encouragement and the opportunities he provided allowed me to gain deeper insights into academic research and further develop my skills beyond what is typically available to master's students. I am also deeply grateful to my co-supervisor, Shyam Mohan, for always being available and open to discussions. His willingness to engage in extensive brainstorming sessions and provide continuous guidance played a crucial role in shaping this work. A special thank you goes to my secondary examiner, Dr. Nadhir Ben Rached, for readily agreeing to co-examine this thesis without hesitation, despite being an external examiner. I truly appreciate the time and effort dedicated to reviewing my work. Furthermore, I would like to thank the entire team at Math4UQ for making me feel like a valued member of the group. Despite being *just a student*, I never felt that way—every discussion, every exchange of ideas, and every collaboration made me feel acknowledged and appreciated. The supportive and inspiring environment at the chair has made this journey an incredibly enriching experience. Finally, I extend my appreciation to everyone who has contributed, directly or indirectly, to the completion of this thesis. Your encouragement and insights have been invaluable.

Contents

Abstract iii

Acknowledgements v

1 Introduction 1
- 1.1 Objectives of the thesis . . . 2
- 1.2 Structure of the thesis . . . 3

2 Mathematical Preliminaries 5
- 2.1 Stochastic Differential Equations . . . 5
- 2.2 Optimal Control of SDEs . . . 11
- 2.3 Optimization . . . 13
  - 2.3.1 Lagrangian relaxation . . . 13
  - 2.3.2 Subgradient methods . . . 14

3 Quadratic Optimal Control of Ornstein-Uhlenbeck Processes 19
- 3.1 Dynamic Programming & Hamilton-Jacobi-Bellman equation . . . 21
- 3.2 Deterministic Optimal Control . . . 21
- 3.3 An Upwind Finite Difference scheme for HJB . . . 24
- 3.4 Ricatti Ansatz in Quadratic Optimal Control . . . 27
- 3.5 Numerical results . . . 29
  - 3.5.1 Sampling controlled SDEs . . . 29
  - 3.5.2 Pontryagin principle . . . 29
  - 3.5.3 HJB equation & Ricatti Ansatz . . . 33

4 Chance Constrained Optimal Control of SDEs 37
- 4.1 Problem statement . . . 38
- 4.2 Subgradients via KBE . . . 40
  - 4.2.1 Crank Nicolson method . . . 40
- 4.3 Relative Dual gap estimation . . . 41
- 4.4 Subgradient method . . . 46
- 4.5 Finding a suitable λ_0 . . . 47
- 4.6 Error and Work Analysis . . . 50
  - 4.6.1 Smoothing Error . . . 52
  - 4.6.2 Discretization Error . . . 54
  - 4.6.3 Computational Work Analysis . . . 55
- 4.7 Numerical results . . . 59
  - 4.7.1 Algorithm 5: Finding an initial LM λ_0 . . . 63
  - 4.7.2 Algorithm 3: Estimating the relative dual gap . . . 67
  - 4.7.3 Algorithm 4: Subgradient method . . . 69

4.7.4 Cost reduction of optimal solution 76

5 Conclusion **79**
5.1 Outlook 79

A Proofs **81**
A.1 Proof of Example 2.1.1 81
A.2 Proof of Theorem 2.2.1 83
A.3 Proof of Theorem 2.2.2 84
A.4 Proof of Theorem 3.3.1 86
A.5 Proof of Theorem 4.6.2 86
A.6 Proof of Theorem 4.6.3 87
A.7 Proof of Theorem 4.6.4 89

B Supplementary Material **91**
B.1 Convergence Analysis 91
B.1.1 Algorithm 2: UFD scheme 91
B.1.2 Crank-Nicolson method 91
B.2 Ricatti Ansatz Analysis 95

Bibliography **97**

List of Figures

1.1 Sketch visualization of controlled SDE sample paths for different Lagrange Multipliers $\lambda \geq 0$ and their corresponding chance constraint value (Relaxed version of Problem 9). 3

3.1 Results from Algorithm 1 for Problem 2 using fixed parameters and varying the regularization parameter $\mu > 0$. 31
3.2 Results from Algorithm 1 for Problem 2 using fixed parameters and varying the regularization parameter $\mu > 0$. 32
3.3 Approximate solutions to Problem 1 on the domain $[0, 1] \times [-10, 10]$ for an absolute tolerance $\text{TOL} = 0.1$. 34
3.4 Analytic solution to the HJB equation using the Ricatti Ansatz in Equation (3.26) using the representations in Equation (3.30) and (3.31). 35
3.5 Sample paths of the optimally controlled SDE in Problem 1 using the HJB approximation (red) described in Algorithm 2 for sampling as described in (3.32) and using the Ricatti Ansatz (blue) with sampling as described in (3.33). 36

4.1 Numerical verification of the asymptotical work rates of Theorem 4.2.1. The result implies that the workrate of our solver is $\Gamma = 1$ which corresponds to inverting matrices of size n with complexity $\mathcal{O}(n)$ which is possible for tridiagonal matrices. 42
4.2 A schematic comparison of the indicator function (blue) with discontinuity in $x = K$ and a smoothed out version using the sigmoid function $\text{sig}_\gamma(x - K)$. By using the sigmoid function defined in (4.2), the discontinuity is smoothened over a length of $\mathcal{O}(\gamma)$ in both directions. 51
4.3 10 sample paths of the optimally controlled SDE on a discretized set of $N = 500$ time steps. The samples were computed using the procedure described in Equation (2.24) with $N = 500$ and the parameters from Table 4.1. The resulting optimally controlled sample paths for enforcing the constraints are shown in Figure 4.11b and 4.13b. 61
4.4 Estimation of the constant c_γ and rate of the induced smoothing using a sigmoid approximation with parameter γ. 62
4.5 Example run of Algorithm 5 for Problem 8 with $\text{TOL}_\text{b} = 10^{-2}$ 4.5a, 4.5b. The violation probability is $\delta = 0.2$ and $\text{TOL} = 10^{-2}$ and $\text{TOL}_\text{r} = 10^{-1}$. The resulting penalization parameters are $\beta^* \approx 0.752$ with $\text{CC}_\text{bisec}(\beta^*) \approx 0.199$. 64
4.6 Example run of Algorithm 5 for Problem 9 with $\text{TOL}_\text{b} = 10^{-3}$ 4.6a, 4.6b. The violation probability is $\delta = 0.2$ and $\text{TOL} = 10^{-2}$ and $\text{TOL}_\text{r} = 10^{-1}$. The resulting penalization parameters are $\beta^* \approx 0.172$ with $\text{CC}_\text{bisec}(\beta^*) \approx 0.203$. 65

4.7 Runtime of the bisection method in Algorithm 5 for different tolerances $\mathrm{TOL}_\mathrm{b} > 0$. 66
4.8 Example visualization of Algorithm 3 as described in Section 4.3 to estimate the relative duality gap for Problem 8 for a violation probability of $\delta = 0.05$ and a discretization that is optimal w.r.t. Theorem 4.6.2 using $\mathrm{TOL} = 0.1$, $\mathrm{TOL}_\mathrm{r} = 0.1$. The interpolation procedure for a fixed LM $\lambda = 0.5$ with $\tilde{c} = 1$ in Equation (4.20) is visualized in Figure 4.8a. The resulting optimal controls are displayed and compared in Figure 4.8b. The relative dual gap is $d^\lambda \approx 1.512$. 68
4.9 Runtime analysis of Algorithm 3 w.r.t. to the different parameters TOL, TOL_r and δ. The results show that, practically, our solver has a work rate of $\Gamma = 1$ which aligns with the results from Section 4.2.1 and Figure 4.1. 69
4.10 An example run of Algorithm 4 for Problem 8. 72
4.11 An example run of Algorithm 4 for Problem 8. 73
4.12 An example run of Algorithm 4 for Problem 9. 74
4.13 An example run of Algorithm 4 for Problem 9. 75
4.14 A comparison of the different resulting costs in the objective for the optimally controlled dual problem solution for different violation probabilities δ. The resulting asymptotic behaviour of the costs for Problem 8 in Figure 4.14a is approximately of order $\mathcal{O}(\delta^{-0.55})$. The resulting asymptotic behaviour of the costs for Problem 9 in Figure 4.14b is approximately of order $\mathcal{O}(\delta^{-1.68})$. 77

B.1 Estimation of $\tilde{c}_1$: Problem 8 on the domain $[0,1]\times[-10,10]$ with a reference solution using $N_1 = 1500$, $M_1 = 650$ steps. Problem 9 on the domain $[0,1]\times[-2,2]$ with a reference solution using $N_1 = M_1 = 1000$ steps. . . 92
B.2 Estimation of the constant $c_2, \bar{c}_2$ and the rate of convergence of the Crank-Nicolson method for Problem 8 on the domain $[0,T]\times[-10,10]$ with a reference solution using $N_2 = M_2 = 1200$ and Lagrange Multiplier $\lambda = 50$. 93
B.3 Estimation of the constant $c_2, \bar{c}_2$ and the rate of convergence of the Crank-Nicolson method for Problem 9 on the domain $[0,T]\times[-2,2]$ with a reference solution using $N_2 = M_2 = 3000$ steps. 94

List of Tables

4.1 Standard parameters used in all experiments carried out in this section unless specified otherwise. First column is the description or interpretation of the parameter in the second column. The third and fourth column are the specific parameter values in Problem 8 and Problem 9, respectively. . 60

List of Abbreviations

PDE	**P**artial **D**ifferential **E**quation
KBE	**K**olmogorov **B**ackward **E**quation
HJB	**H**amilton **J**acobi **B**ellman
SDE	**S**tochastic **D**ifferential **E**quation
ODE	**O**rdinary **D**ifferential **E**quation
OC	**O**ptimal **C**ontrol
SOC	**S**tochastic **O**ptimal **C**ontrol
UFD	**U**pwind **F**inite **D**ifference
CN	**C**rank **N**icolson
SGM	**S**ub**g**radient **m**ethod
MC	**M**onte **C**arlo
QoI	**Q**uantity **o**f **I**nterest
LM	**L**agrange **M**ultiplier
PMP	**P**ontryagin **M**aximum **P**rinciple
CFL	**C**ourant **F**riedrichs **L**ewy
1D	one-**d**imensional

List of Parameters

α	Control
$\mu > 0$	Regularization parameter for control
A	Compact set of control values
$\theta > 0$	Drift parameter in Ornstein-Uhlenbeck process
$x_0 \in \mathbb{R}$	Initial value for SDE
$T > 0$	Terminal time of finite time horizon of OC problem
$\sigma > 0$	Diffusion parameter in SDE
u	Value function of OC problem, solution to HJB equation
v	Solution to KBE
a	Drift function in SDE
b	Diffusion function in SDE
h	Running cost function of SOC problem
g	Terminal cost function of SOC problem
G	Scalar constraint perturbation function
δ	Confidence level of chance constraint
K	Threshold for perturbed final state in chance constraint
γ	Smoothing parameter in sigmoid function for indicator approx.
λ	LM in Lagrangian relaxation technique
κ	Penalization parameter in dual gap estimation procedure
β	Penalization parameter in bisection-initialization procedure
η	Control in auxiliary problem for dual gap estimation
R	Bound for distance of initial value to optimal value in SGM assumptions
S	Bound for subgradient norm in SGM convergence assumptions
d	Upper bound for state space restriction in numerics
c	Lower bound for state space restriction in numerics
N_1	Number of time discretization steps in HJB approx.
M_1	Number of space discretization steps in HJB approx.
N_2	Number of time discretization steps in KBE approx.
M_2	Number of space discretization steps in KBE approx.
TOL	Abs. tolerance for discretization error and smoothing error in HJB approx.
TOL_r	Rel. tolerance in estimation of subgradient
TOL_SG	Tolerance in SGM
TOL_b	Tolerance in bisection-based approach
$\mathcal{W}$	Work of a procedure

Chapter 1

Introduction

This thesis deals with computational approaches to handle chance constraints in stochastic optimal control problems in which the controlled system is modelled by a SDE. SDEs arise naturally in many real-world applications where randomness plays a fundamental role. They extend ODEs by incorporating stochastic processes, typically driven by a Wiener process, to account for random fluctuations in the system's evolution. Typical application areas are finance [1] in which economic factors like interest rates, exchange rates, etc. are inherently unpredictable and require stochastic models for robust decision-making, biological systems such as neuron firing patterns [2] or population dynamics [3] involve inherent randomness and engineering systems, such as robotics and automated control, which must operate under uncertain conditions, like sensor noise and environmental disturbances [4]. Especially in the latter, when human lives are at the hand of automated systems under uncertainty, the question of safety constraints arise. Reliable systems need to satisfy safety constraints that are linked to failure probabilities of a system. These can be especially challenging to account for when the failure probabilities are minuscule. At the core of stochastic optimal control lies the HJB equation, a fundamental principle that extends Bellman's dynamic programming approach to continuous-time decision processes [1]. This equation provides a systematic way to derive optimal strategies governed by SDEs, allowing for modeling of complex real-world phenomena.

The literature for chance constrained stochastic optimal control can be divided into two categories: Continuous-time and discrete-time SOC. Especially in discrete-time formulations the literature is broad, covering a large quantity of proposed methods to tackle a lot of different forms of chance constraints [5]. Large-scale discrete-time SOC problems can be tackled by decomposing it into smaller subproblems with corresponding almost sure coupling constraints [6, 7]. There are a lot of sampling-based approaches with applications to autonomous vehicle path planning [8, 4, 9], methods to handle joint chance constraints [10], especially for multi-agent-systems with additive Gaussian noise [11], and empirical approaches to approximate the underlying distributions of the noise of the system [12]. Applications vary from mass-damper-spring systems in mechanical engineering [13, 14, 15] to chemical processes [16, 17]. There are also proposed methods to handle rare events using polynomial chaos theory and subset simulation techniques [18] and using the concept of conservative approximations [19]. This thesis considers continuous-time formulations that are related to the approach in [20] which is a discrete-time formulation. Under certain assumptions one can prove that the discrete-time formulation converges to the continuous-time formulation [21]. In the last chapter of [22], the author considers

terminal state chance constraint of the same type as we will. Afterwards, the author proposes two ways to handle expectation constraints on the terminal state using dynamic programming and Lagrangian relaxation [23]. [24] investigates different types of chance constraints and theoretical results on the applicability of dynamic programming to solve the underlying chance constrained SOC problem. Random PDE constrained stochastic optimization has also been recently worked on using sampling-based approaches [25, 26, 27]. In [26], the author shows that almost sure constraints are limits of our type of probabilistic constraints. [28] considers stochastic target constraints and names constraints in moments or probability as possible extensions of his theoretical work. There are also theoretical results on optimality conditions for terminal state constraints in SOC [29]. Applications in the continuous-time setting vary from finance, e.g. quantitative finance [24], asset-liability management [22], optimal portfolio management [30, 31, 32], to fishery management [20] and optimal power procurement for green cellular wireless networks [33]. Reliable numerical methodologies are lacking in the literature for general chance constrained optimal control of SDEs. We aim to not only give theoretical results on general problem formulations but also set the stage for a methodology to handle and solve these generalized problems numerically.

1.1 Objectives of the thesis

Let $\{X(t) \in \mathbb{R} \mid t \in [0,T]\}$ be the solution to a SDE up to the terminal time $T > 0$. We aim to incorporate chance constraints of the form

$$\mathbb{P}\left(G(X(T)) > K\right) \leq \delta,$$

in which $G : \mathbb{R} \to \mathbb{R}$ is a scalar function and $K \in \mathbb{R}$ is a threshold, to model the constraint $G(X(T)) \leq K$ that has to be uphold with a probability of $1-\delta$, $0 < \delta \ll 1$. The dynamic programming principle is not directly applicable to the chance constrained SOC formulation because the constraint controls the joint distribution of the underlying random variables, rather than each realization [34, 33]. That is why we use Lagrangian relaxation to include a penalized version of the terminal time state constraint in the objective function's terminal costs to reduce the problem to a classical SOC problem and solve it using existing optimization techniques applied to our problem setting. Chance constrained SOC presents numerous challenges, including computational complexity of solving non-linear PDEs with discontinuous boundary conditions and the difficulty of obtaining closed-form solutions for many practical applications. This thesis addresses these challenges by exploring stable solvers for the HJB equation with non-smooth boundary conditions, numerical optimization techniques to solve the dual problem arising from Lagrangian relaxation of the constraint and asymptotic optimal computational work results for the proposed procedures. By developing and analyzing efficient and robust methodologies, we seek to contribute to the broader understanding of handling chance constrained optimal control of SDEs. Through this study, we aim to deepen the theoretical and practical understanding of chance constrained SOC, addressing key challenges and paving the way for more efficient and robust decision-making frameworks in uncertain environments. Take the example of a drift-controlled Wiener process defined by the SDE $X(t) = \alpha(t, X(t))dt + 0.1 \cdot dW(t)$, $X(0) = 0.25$, $t \in [0,1]$. We aim to minimize the objective function

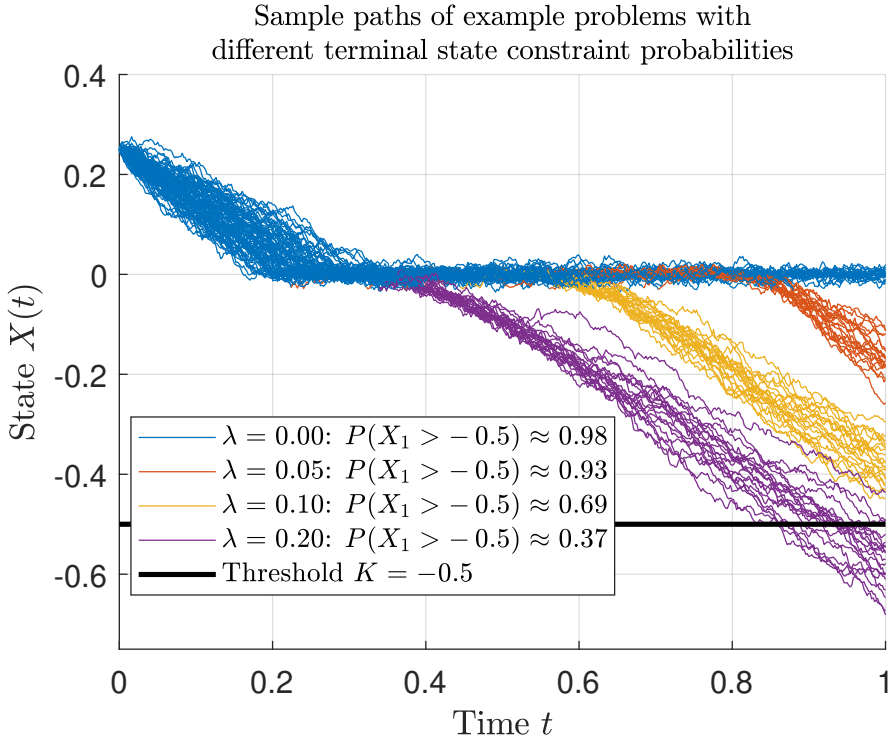


Figure 1.1: Sketch visualization of controlled SDE sample paths for different Lagrange Multipliers $\lambda \geq 0$ and their corresponding chance constraint value (Relaxed version of Problem 9).

$$\min_{\alpha \in \mathcal{A}} \mathbb{E}\left[\int_0^1 X(t)^2 dt\right], \quad \text{s.t.} \quad \mathbb{P}\left(X(1) > -0.5\right) \leq 0.2.$$

After Lagrangian relaxation of the constraint, we aim to find the optimal Lagrange Multiplier that results in a primal feasible solution to the constrained problem above. For fixed Lagrange Multipliers, we could look at the sample paths and investigate the chance constraint value heuristically. This was done before, e.g. in [35], and for our example in Figure 1.1 but is not satisfactory. We want to mathematically derive an optimal procedure to find the optimal Lagrange Multiplier that solves the dual problem numerically up to a certain tolerance $\text{TOL}_{\text{SG}} > 0$.

1.2 Structure of the thesis

The remainder of this thesis is structured as follows: Chapter 2 provides a comprehensive review of the fundamental principles of SOC and optimization, including key mathematical formulations of the underlying theory needed for the goal of this thesis. Chapter 3 reviews unconstrained optimal control of SDEs applied to the example of quadratic control of Ornstein-Uhlenbeck processes which can be seen as the stochastic counterpart of the linear quadratic regulator problem in continuous-time over a finite time horizon. Chapter 4 presents the problem statement and the methodology of solving chance constrained optimal control of SDEs. We derive the dual problem and propose a way of

solving the dual problem numerically while accounting for its reliability via an estimation of the relative dual gap. The theoretical results are tested, verified and discussed in numerical experiments. Chapter 5 concludes with a summary of findings and potential directions for future research.

Chapter 2

Mathematical Preliminaries

In this chapter, the mathematical preliminaries for the upcoming chapters are presented. We introduce the notion of SDEs, preliminaries from OC theory, constrained optimization and SGMs for numerical optimization of non-smooth objective functions. More about SDEs can be read in [1, 36, 37, 38]. [36] gives a brief overview of optimal control theory. Both [1] and [36] give an introduction to optimal control of SDEs. A more in-depth overview of constrained optimization methods and constrained optimization theory can be found in [39] and [40]. More on stochastic optimization can be found in [41]. We will introduce all the definitions and theorems with simple examples to follow along.

2.1 Stochastic Differential Equations

In this section, we will talk about SDEs driven by a Brownian motion. For the basic notion of SDEs we need the definition of a Wiener process, the Itô integral and some basic properties. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space triplet consisting of the set of all outcomes Ω, $\mathcal{F}$ the set of events and a probability measure $\mathbb{P} : \mathcal{F} \to [0, 1]$.

Definition 2.1.1 (One-dimensional Wiener Process)**.** *The one dimensional Wiener process $W : [0, \infty) \times \Omega \to \mathbb{R}$, also known as the Brownian motion, has the following properties:*

(1) With probability 1, the mapping $t \mapsto W(t)$ is continuous and $W(0) = 0$

(2) If $0 = t_0 < t_1 < \ldots < t_N = T$, then the Wiener increments

$$W(t_N) - W(t_{N-1}), \ldots, W(t_1) - W(t_0)$$

are independent

(3) For all $0 < s < t$ the increment $W(t) - W(s)$ is normally distributed, with $\mathbb{E}[W(t) - W(s)] = 0$ and $\mathbb{E}[(W(t) - W(s))^2] = t - s$, i.e. for $B \subset \mathbb{R}$, we have

$$\mathbb{P}(W(t) - W(s) \in B) = \frac{1}{\sqrt{2\pi(t-s)}} \int_B \exp\left(\frac{-y^2}{2(t-s)}\right) dy.$$

We can define the Itô integral as the limit of the forward Euler method and show that the limit is independent of the partition used in the time discretization.

Theorem 2.1.1 (Itô integral as limit of Forward Euler approximations)**.** *Let $W(t)$ be the Wiener Process and $0 = t_0 < t_1 < \ldots < t_N = T$ a discretization of the time interval $[0,T] \subset \mathbb{R}_+$. The Itô Integral with integrand $f : [0,T] \times \mathbb{R} \to \mathbb{R}$*

$$I := \int_0^T f(s, W(s))dW(s)$$

is the limit of the forward Euler (left point quadrature) approximation

$$I_{\Delta t} := \sum_{n=0}^{N-1} f(t_n, W(t_n))\Delta W_n,$$

as $\max_{n=1,\ldots,N} \Delta t_n \to 0$, where $\Delta t_n = t_{n+1} - t_n$ and $\Delta W_n = W(t_{n+1}) - W(t_n)$ for $n = 0, \ldots, N-1$. This limit (in L^2 sense) is independent of the discretization used for Lipschitz integrands, i.e. suppose there exists a positive constant $C > 0$ such that $f : [0,T] \times \mathbb{R} \to \mathbb{R}$ satisfies

$$|f(t + \Delta t, W + \Delta W) - f(t, W)| \leq C(\Delta t + |\Delta W|), \tag{2.1}$$

with any increment ΔW. Consider two different discretizations of the time interval $[0,T]$

$$\{\bar{t}_n\}_{n=0}^{\overline{N}}, \quad \bar{t}_0 = 0, \quad \bar{t}_{\overline{N}} = T, \tag{2.2}$$

$$\{\bar{\bar{t}}_n\}_{n=0}^{\bar{\overline{N}}}, \quad \bar{\bar{t}}_0 = 0, \quad \bar{\bar{t}}_{\bar{\overline{N}}} = T, \tag{2.3}$$

with corresponding forward Euler approximations

$$\overline{I} = \sum_{n=0}^{\overline{N}-1} f(\bar{t}_n, W(\bar{t}_n))(W(\bar{t}_{n+1}) - W(\bar{t}_n)), \tag{2.4}$$

$$\bar{\bar{I}} = \sum_{m=0}^{\bar{\overline{N}}-1} f(\bar{\bar{t}}_m, W(\bar{\bar{t}}_m))(W(\bar{\bar{t}}_{m+1}) - W(\bar{\bar{t}}_m)). \tag{2.5}$$

Let the maximum time step Δt_{max} be

$$\Delta t_{max} = \max\left[\max_{0 \leq n \leq \overline{N}-1} \bar{t}_{n+1} - \bar{t}_n, \max_{0 \leq m \leq \bar{\overline{N}}-1} \bar{\bar{t}}_{m+1} - \bar{\bar{t}}_m\right].$$

Then

$$\mathbb{E}[(\overline{I} - \bar{\bar{I}})^2] = \mathcal{O}(\Delta t_{max}). \tag{2.6}$$

Proof. A proof can be found in [36]. □

Furthermore, we are interested in processes that are adapted to the filtration of a Wiener process which can be interpreted as being independent of future events.

Definition 2.1.2 (Adapted process)**.** *A process $F : [0,T] \times \Omega \to \mathbb{R}$ is adapted if $F(t, \cdot)$ only depends on events which are generated by $W(s)$, $s \leq t$.*

In an adapted process, we can fully determine whether an event occurs at current time t based on the values of $W(s)$ up to time t. The upcoming processes will be adapted to the filtration of the Wiener process if not mentioned otherwise. Let us now write down basic basic properties of Itô-integrals.

Theorem 2.1.2 (Basic properties of Itô-integrals). *Suppose that $F, G : [0,T] \times \Omega \to \mathbb{R}$ are Itô-integrable and that $c_1, c_2 \in \mathbb{R}$ are constants. Then*

(i) Linearity of Itô-integrals:

$$\int_0^T (c_1 F(s,\cdot) + c_2 G(s,\cdot))dW(s) = c_1 \int_0^T F(s,\cdot)dW(s) + c_2 \int_0^T G(s,\cdot)dW(s),$$

(ii) Martingale property of Itô-integrals:

$$\mathbb{E}\left[\int_0^T F(s,\cdot)dW(s)\right] = 0,$$

(iii) Itô-isometry:

$$\mathbb{E}\left[\left(\int_0^T F(s,\cdot)dW(s)\right)\left(\int_0^T G(s,\cdot)dW(s)\right)\right] = \int_0^T \mathbb{E}[F(s,\cdot)G(s,\cdot)]ds,$$

holds.

Proof. A proof is found in [36]. □

The Ornstein-Uhlenbeck (stochastic) process is used as a running example within this thesis. It is a linear (Gaussian) SDE which can be solved explicitly. It is commonly used in finance to model interest rates [42], biology to model neuron firing rates [43], physics where it was first introduced to model the velocity of a Brownian particle under friction [44], machine learning [45] and Bayesian filtering [46]. Some of the properties using its explicit integral representation in Equation (2.7) are given in Example 2.1.1.

Example 2.1.1 (Ornstein-Uhlenbeck process). *Consider the Ornstein-Uhlenbeck process defined by*

$$X(t) = x_\infty + e^{-\theta t}(x_0 - x_\infty) + \sigma \int_0^t e^{-\theta(t-s)}dW(s), \tag{2.7}$$

where $\theta, \sigma > 0$ are given constants representing a drift parameter and the volatility, respectively. This (Gaussian) process has the following properties

1.) Expected value and variance for a fixed time $t \in [0,T]$ are given by

$$\mathbb{E}[X(t)] = x_\infty + e^{-\theta t}(x_0 - x_\infty), \quad \mathrm{Var}[X(t)] = \frac{\sigma^2}{2\theta}(1 - e^{-2\theta t}).$$

2.) The limit, as $t \to \infty$, is given by

$$\lim_{t\to\infty} \mathbb{E}[X(t)] = x_\infty, \quad \lim_{t\to\infty} \mathrm{Var}[X(t)] = \frac{\sigma^2}{2\theta}.$$

3.) The covariance between $X(t)$ and $X(t+\tau)$, $\tau > 0$, is given by

$$\mathrm{Cov}(X(t), X(t+\tau)) = \frac{\sigma^2}{2\theta}(e^{-\theta\tau} - e^{-\theta(2t+\tau)}).$$

4.) For fixed t

$$\lim_{\tau\to\infty} \mathrm{Cov}(X(t), X(t+\tau)) = 0$$

holds.

Since the limit, as $t \to \infty$, exists for both the expectation and variance, we know that a stationary distribution for the process X exists and is given by the Gaussian probability density function of the random variable

$$X_{\text{stationary}} \sim \mathcal{N}\left(x_\infty, \frac{\sigma^2}{2\theta}\right).$$

For the other limit, we can see that $X(t)$ and $X(t+\tau)$ are less correlated the more we increase their distance τ. The proofs of these properties can be found in Appendix A.1. In contrast to Brownian motion, which drifts off to infinity, the Ornstein-Uhlenbeck process reverts to a long-term mean x_∞. This models realistic scenarios where a system fluctuates around a stable equilibrium.

Using the notion of Itô integrals, which are defined as the limit of forward Euler approximations, and the proof of Theorem 2.1.1, we can prove convergence of the limit of forward Euler approximations to SDEs. First, we will extend an Euler approximation $\bar{X}(t)$ to all $t \in [0,T]$, i.e. consider a grid $\{\bar{t}_k\}_{n=0}^{\bar{N}}$, $\bar{t}_0 = 0$, $\bar{t}_{\bar{N}=T}$, and define the discrete stochastic process $\bar{X}$ by the forward Euler method

$$\bar{X}(\bar{t}_{n+1}) - \bar{X}(\bar{t}_n) = a(\bar{t}_n, \bar{X}(\bar{t}_n))(\bar{t}_{n+1} - \bar{t}_n) + b(\bar{t}_n, \bar{X}(\bar{t}_n))(W(\bar{t}_{n+1}) - W(\bar{t}_n)),$$

for $n = 0, \dots, \bar{N}-1$. Now extend $\bar{X}$ continuously to all values of $t \in [\bar{t}_n, \bar{t}_{n+1})$ for all $n = 0, \dots, \bar{N}-1$ as a continuous process $\tilde{X}$ defined by

$$\tilde{X}(t) = \bar{X}(\bar{t}_n) + \int_{\bar{t}_n}^{t} a(\bar{t}_n, \bar{X}(\bar{t}_n))ds + \int_{\bar{t}_n}^{t} b(\bar{t}_n, \bar{X}(\bar{t}_n))dW(s), \quad \bar{t}_n \le t < \bar{t}_{n+1}.$$

In other words, the process $\tilde{X} : [0,T] \times \Omega \to \mathbb{R}$ satisfies for all $n = 0, \dots, \bar{N}-1$ the SDE

$$\begin{aligned} d\tilde{X}(t) &= \bar{a}(t, \bar{X})dt + \bar{b}(t, \bar{X})dW(t), \quad \bar{t}_n \le t < \bar{t}_{n+1}, \\ \tilde{X}(\bar{t}_n) &= \bar{X}(\bar{t}_n), \end{aligned}$$

where $\bar{a}(t, \bar{X}) \equiv a(\bar{t}_n, \bar{X}(\bar{t}_n))$, $\bar{b}(t, \bar{X}) \equiv b(\bar{t}_n, \bar{X}(\bar{t}_n))$, for $\bar{t}_n \le t < \bar{t}_{n+1}$. We can prove a similar result as in Theorem 2.1.1, namely

Theorem 2.1.3. *Let $\bar{X}$ and $\bar{\bar{X}}$ be forward Euler approximations of the stochastic process $X : [0,T] \times \Omega \to \mathbb{R}$, satisfying the stochastic differential equation*

$$dX(t) = a(t, X(t))dt + b(t, X(t))dW(t), \quad 0 < t < T, \quad X(0) = x_0 \in \mathbb{R},$$

with time steps

$$\begin{aligned} &\{\bar{t}_n\}_{n=0}^{\bar{N}}, \quad \bar{t}_0 = 0, \quad \bar{t}_{\bar{N}} = T, \\ &\{\bar{\bar{t}}_n\}_{n=0}^{\bar{\bar{N}}}, \quad \bar{\bar{t}}_0 = 0, \quad \bar{\bar{t}}_{\bar{\bar{N}}} = T, \end{aligned}$$

respectively, and

$$\Delta t_{\max} = \max\left[\max_{0\le n\le \bar{N}-1} \bar{t}_{n+1} - \bar{t}_n, \max_{0\le n\le \bar{\bar{N}}-1} \bar{\bar{t}}_{n+1} - \bar{\bar{t}}_n\right].$$

Suppose that there exists a positive constant $C > 0$ such that the initial data and the drift and diffusion functions $a, b : [0,T] \times \mathbb{R} \to \mathbb{R}$ satisfy

$$\mathbb{E}\left[|\bar{X}(0)|^2 + |\bar{\bar{X}}(0)|^2\right] \le C,$$
$$\mathbb{E}\left[(\bar{X}(0) - \bar{\bar{X}}(0))^2\right] \le C \cdot \Delta t_{\max}$$

and the Lipschitz conditions

$$|a(t,x) - a(t,y)| < C|x-y|, \quad |b(t,x) - b(t,y)| < C|x-y|$$
$$|a(t,x) - a(s,x)| + |b(t,x) - b(s,x)| \le C(1+|x|)\sqrt{|t-s|}.$$

Then, there is a constant $K > 0$ such that

$$\max\left\{\mathbb{E}\left[\bar{X}^2(t,\cdot)\right], \mathbb{E}\left[\bar{\bar{X}}^2(t,\cdot)\right]\right\} \le K(T+1), \quad t < T,$$

and

$$\mathbb{E}[(\bar{X}(t,\cdot) - \bar{\bar{X}}(t,\cdot))^2] \le K\Delta t_{\max}, \quad t < T.$$

Proof. A proof is given in [36]. □

Corollary 2.1.1. *The previous theorem yields a convergence result in the L^2-norm*

$$\|X\|_2^2 = \int_0^T \mathbb{E}[X(t)^2]dt,$$

namely, the order of convergence is $1/2$, i.e.

$$\|\bar{X} - \bar{\bar{X}}\|_2 = \mathcal{O}\left(\sqrt{\Delta t_{\max}}\right).$$

This corollary shows that successive refinements of the forward Euler approximation forms a Cauchy sequence in the Hilbert space

$$V := \{F : [0,T] \times \Omega \to \mathbb{R}^l \ : \ F \text{ is adapted and } \int_0^T \mathbb{E}[|F(t,\cdot)|^2]dt < \infty\}. \tag{2.8}$$

The limit $X \in V$, of this Cauchy sequence, satisfies the stochastic equation

$$X(s) = X(0) + \int_0^s a(t, X(t))dt + \int_0^s b(t, X(t))dW(t), \quad 0 < s \le T, \tag{2.9}$$

and is unique [36]. Hence, we have constructed existence and uniqueness of solutions to Equation (2.9) by forward Euler approximations. Let X be the solution of Equation (2.9). From now on we use indistinctly the notation

$$dX(t) = a(t, X(t))dt + b(t, X(t))dW(t), \quad 0 < t \le T, \tag{2.10}$$
$$X(0) = x_0. \tag{2.11}$$

The next theorem allows us to represent functions of solutions to SDEs as SDEs themselves.

Theorem 2.1.4 (Itô's lemma). *Suppose that the assumptions in Theorem 2.1.3 hold and that X satisfies the SDE*

$$\begin{aligned} dX(s) &= a(s, X(s))ds + b(s, X(s))dW(s), \quad s > 0, \\ X(0) &= x_0, \end{aligned}$$

and let $g : (0, \infty) \times \mathbb{R} \to \mathbb{R}$ be a given bounded function in $C^2((0, \infty) \times \mathbb{R})$. Then $Y(t) \equiv g(t, X(t))$ satisfies the SDE

$$\begin{aligned} dY(t) = &\left(\partial_t g(t, X(t)) + a(t, X(t))\partial_x g(t, X(t)) + \frac{b^2(t, X(t))}{2} \partial_{xx} g(t, X(t)) \right) dt \\ &+ b(t, X(t))\partial_x g(t, X(t))dW(t). \end{aligned}$$

Proof. A proof can be found in [36]. □

Continuing the example of an Ornstein-Uhlenbeck process, we can also state the SDE whose solution is an Ornstein-Uhlenbeck process. The SDE corresponding to the process defined in Equation (2.7) is

$$dX(t) = -\theta(X(t) - x_\infty)dt + \sigma dW(t), \quad t > 0, \tag{2.12}$$

$$X(0) = x_0. \tag{2.13}$$

In future discussions, we consider $x_\infty \equiv 0$. In many applications, the QoI will be the following

$$\mathbb{E}\left[g(X(T)) \mid X(0) = x_0 \right]. \tag{2.14}$$

We have a powerful tool to get the exact value for this QoI without needing any sampling technique or estimation method. Due to the following Theorem 2.1.5, we can solve the KBE instead of, for example, using MC methods to get a statistical estimate of the QoI in Equation (2.14).

Theorem 2.1.5 (Kolmogorov Backward Equation). *Suppose that a, b and g are differentiable to any order and these derivatives are bounded. Let X be the solution of the SDE*

$$\begin{aligned} dX(t) &= a(t, X(t))dt + b(t, X(t))dW(t), \quad 0 < t < T, \\ X(0) &= x_0, \end{aligned}$$

and let

$$v(t, x) = \mathbb{E}\left[g(X(T)) \mid X(t) = x \right]. \tag{2.15}$$

Then v is the solution of the Kolmogorov Backward Equation

$$0 = \partial_t v(t, x) + a(t, x) \cdot \partial_x v(t, x) + \frac{b(t, x)^2}{2} \partial_{xx} v(t, x), \quad (t, x) \in [0, T) \times \mathbb{R}, \tag{2.16}$$

$$v(T, x) = g(x). \tag{2.17}$$

Proof. A proof is given in [36]. □

Remark 2.1.1. *All the results, for example Definition 2.1.1 and Theorem 2.1.1, can be extended to multiple dimensions. The multidimensional Wiener process W in $\mathbb{R}^\ell$ is defined by $W(t) \equiv (W^1(t), \dots, W^\ell(t))$, where W^i, $i = 1, \dots, \ell$ are independent one-dimensional Wiener processes. One can show that*

$$I_{\Delta t} \equiv \sum_{n=0}^{N-1} \sum_{i=1}^{\ell} f_i(t_n, \cdot) \Delta W_n^i,$$

where $\Delta W_n^i := W^i(t_{n+1}) - W^i(t_n)$, for a Cauchy sequence with

$$\mathbb{E}[(I_{\Delta t_1} - I_{\Delta t_2})^2] = \mathcal{O}(\Delta t_{\max}),$$

as in Theorem 2.1.1, provided $f : [0, T] \times \Omega \to \mathbb{R}^l$ is adapted and for a constant $C > 0$

$$\sqrt{\mathbb{E}\left[|F(t + \Delta t, \omega) - f(t, \omega)|^2\right]} \leq C\sqrt{\Delta t}, \quad \omega \in \Omega,$$

holds. We will stick to the 1D cases throughout the thesis although many of the results can be extended to multiple dimensions.

2.2 Optimal Control of SDEs

In this chapter, we are investigating controlled SDEs of the form

$$dX(t) = a(t, X(t), \alpha(t, X(t)))dt + b(t, X(t), \alpha(t, X(t)))dW(t), \quad 0 < t < T, \tag{2.18}$$

$$X(0) = x_0, \tag{2.19}$$

where $T > 0$ is a fixed terminal time and $x_0 \in \mathbb{R}$ is a given initial value. We are assuming that the *drift* function $a : [0, T] \times \mathbb{R} \times A \to \mathbb{R}$ and *diffusion* function $b : [0, T] \times \mathbb{R} \times A \to \mathbb{R}$ are smooth bounded functions, where A is a given compact subset of $\mathbb{R}$ of control values. The function $\alpha : [0, T] \times \mathbb{R} \to A$ is called *control* and

$$\mathcal{A} := \{\alpha : [0, T] \times \mathbb{R} \to A\}$$

the set of admissible deterministic Markov control functions. The control function

$$\alpha_t := \alpha(t, X(t)), \quad t \in [0, T],$$

uses the current time t and state $X_t := X(t)$ at time $t \in [0, T]$ to affect the dynamics of the SDE $X(t)$ by adjusting the drift and/or the diffusion functions. For $\alpha \in \mathcal{A}$, we define the cost functional

$$C_{t,x}(\alpha) = \mathbb{E}\left[\int_t^T h(s, X_s, \alpha_s)ds + g(X_T) \,\middle|\, X(t) = x\right],$$

where X solves the SDE (2.18) with control α and

$$h : [0, T] \times \mathbb{R} \times A \to \mathbb{R}, \quad g : \mathbb{R} \to \mathbb{R}$$

are given smooth bounded functions. We call h the *running cost* and g the *terminal costs*. Our goal is to find an optimal control, denoted by a star in superscript, $\alpha^* \in \mathcal{A}$,

which minimizes the expected cost defined by $C_{t,x}(\alpha)$. Assuming a minimum is attained by some $\alpha^* \in \mathcal{A}$, we define the value function as

$$u(t,x) \equiv \min_{\alpha \in \mathcal{A}} C_{t,x}(\alpha) = C_{t,x}(\alpha^*). \tag{2.20}$$

Throughout the thesis we assume that the minimum is attained. Assuming that all the assumptions above hold, we can solve a standard optimal control problem, as described above, using the dynamic programming relation.

Theorem 2.2.1 (Dynamic Programming)**.** *The (value-) function u satisfies, for all $0 < \varepsilon < T - t$, the dynamic programming relation*

$$u(t,x) = \min_{\alpha:[t,t+\varepsilon]\to A} \mathbb{E}\left[\int_t^{t+\varepsilon} h(s, X_s, \alpha_s))ds + u(t+\varepsilon, X(t+\varepsilon))\right]. \tag{2.21}$$

Proof. The proof is divided into two steps:

1. Use the optimal control to verify that

$$u(t,x) \geq \min_{\alpha:[t,t+\varepsilon]\to A} \mathbb{E}\left[\int_t^{t+\varepsilon} h(s, X_s, \alpha_s)ds + u(t+\varepsilon, X(t+\varepsilon))\right]$$

   holds.

2. Show that a suitably constructed control yields

$$u(t,x) \leq \min_{\alpha:[t,t+\varepsilon]\to A} \mathbb{E}\left[\int_t^{t+\varepsilon} h(s, X_s, \alpha_s)ds + u(t+\varepsilon, X(t+\varepsilon))\right]$$

which together imply (2.21). The full proof is given in Appendix A.2. □

Using the dynamic programming relation, we can reduce the solution of the problem to a solution to a non-linear PDE, namely the HJB equation.

Theorem 2.2.2 (Hamilton-Jacobi-Bellman equation)**.** *Assume that X solves the controlled SDE in Equation (2.18) with a Markov control function α and that the value function $u(t,x) \equiv \min_{\alpha\in\mathcal{A}} C_{t,x}(\alpha)$ is bounded and smooth. Then u satisfies the (non-linear) Hamilton-Jacobi-Bellman (partial differential) equation*

$$0 = \partial_t u(t,x) + H(t,x,\partial_x u(t,x), \partial_{xx} u(t,x)), \quad (t,x) \in [0,T) \times \mathbb{R}, \tag{2.22}$$

$$u(T,x) = g(x), \tag{2.23}$$

with the Hamiltonian function

$$H(t,x,\partial_x u(t,x), \partial_{xx} u(t,x)) = \min_{\alpha \in A}\left[a(t,x,\alpha)\partial_x u(t,x) + \frac{b(t,x,\alpha)^2}{2}\partial_{xx} u(t,x) + h(t,x,\alpha)\right].$$

Proof. The proof is given in Appendix A.3. □

Note that the Hamiltonian in Theorem 2.2.2 is a pointwise minimum over all possible control values in the compact set of control values A. This result allows us to find a pointwise representation of the optimal control for any time t and state $X(t)$. Finally, we describe the sampling process of any stochastic differential equation. This will be needed to investigate sample paths of controlled SDEs in future chapters. Let $\{t_n\}_{n=0}^N$ be an equidistant time discretization of the time domain $[0,T]$ with $\Delta t = T/N$, $t_0 = 0$ and $t_N = T$. A sample path of the SDE in Equation (2.18) on the grid $\{t_n\}_{n=0}^N$ can be obtained by computing

$$X(t_{n+1}) = X(t_n) + a(t_n, X(t_n), \alpha(t_n, X(t_n)))\Delta t + b(t_n, X(t_n), \alpha(t_n, X(t_n)))\sqrt{\Delta t}\cdot Z \tag{2.24}$$

for $n = 0, \ldots, N-1$ where $Z \sim \mathcal{N}(0,1)$ and $X(t_0) = x_0 \in \mathbb{R}$.

2.3 Optimization

This section introduces the general notation for optimization methods used in chance constrained optimal control of SDEs in Section 4. We will talk about duality, Lagrangian relaxation techniques and subgradient methods. Note that this section is about the notation and terminology. The concept of Lagrange Multipliers can be extended to optimization of functionals as well which we will be doing in future chapters. Any Calculus of Variation textbook or lecture notes should include theorems about Lagrangian Relaxation of functionals, for example [47]. The introduction and Section 2.3.1 about Lagrangian relaxation is using [39] as a guidance to introduce important notation and results. For the SGM in Section 2.3.2, [40] is used.

Let $f_i : \mathrm{dom}(f_i) \to \mathbb{R}$, $i = 0, \dots, m$ and $h_j : \mathrm{dom}(h_j) \to \mathbb{R}$, $j = 1, \dots, p$, functions on a compact domain $\mathrm{dom}(f_i), \mathrm{dom}(h_j) \subseteq \mathbb{R}^n$. Consider the following constrained optimization problem

$$\begin{cases} \min_{x \in \mathcal{D}} & f_0(x) \\ \text{s.t.} & f_i(x) \leq 0, \quad i = 1, \dots, m, \\ & h_j(x) = 0, \quad j = 1, \dots, p, \end{cases} \tag{2.25}$$

in which we want to find

$$x \in \mathcal{D} := \bigcap_{i=0}^{m} \mathrm{dom}(f_i) \cap \bigcap_{j=1}^{p} \mathrm{dom}(h_j),$$

that minimizes the function f_0 among all x that satisfy the inequality constraints $f_i(x) \leq 0$, $i = 1, \dots, m$ and equality constraints $h_j(x) = 0$, $j = 1, \dots, p$. A point $x \in \mathcal{D}$, where $\mathcal{D}$ is the domain of the optimization problem, is called *feasible* if it satisfies both the inequality and the equality constraints. $x \in \mathcal{D}$ is called the *optimization variable* and f_0 is the *objective* or *cost function*. If $m = p = 0$, the problem is called *unconstrained*. The problem itself is called *feasible* if there exists at least one feasible point, otherwise it is called *infeasible*. The set of all feasible points is called the *feasible set*. We will denote optimal values with a superscript star, i.e. the optimal value p^* of the problem is defined as

$$p^* = \min\{f_0(x) \mid f_i(x) \leq 0,\ i = 1, \dots, m,\ h_j(x) = 0,\ j = 1, \dots, p\}.$$

Accordingly, we say that x^* is an *optimal point* if x^* is feasible and $f_0(x^*) = p^*$. Finally, a feasible point x with $f_0(x) \leq p^* + \varepsilon$, $\varepsilon > 0$, is called *ε-suboptimal*.

2.3.1 Lagrangian relaxation

Consider the optimization problem in Equation (2.25). We define the *Lagrangian* $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ associated with Equation (2.25) as

$$\mathcal{L}(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{p} \nu_j h_j(x), \tag{2.26}$$

with $\mathrm{dom}(\mathcal{L}) = \mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^p$. $\lambda_i \in [0, \infty)$ and $\nu_j \in \mathbb{R}$ are called *Lagrange Multipliers* associated with the ith constraint f_i and jth constraint h_j, respectively. $\lambda = (\lambda_1, \dots, \lambda_m)$

and $\nu = (\nu_1, \ldots, \nu_p)$ are called *dual variables*. Using the Lagrangian $\mathcal{L}$, we can define the *Lagrange dual function* as

$$d(\lambda, \nu) = \inf_{x \in \mathcal{D}} \mathcal{L}(x, \lambda, \nu). \tag{2.27}$$

Due to the nature of the construction, the dual function is concave, even if the problem in Equation (2.25) is not convex. The dual function yields a lower bound for p^*. For any feasible point $\tilde{x}$ we have

$$d(\lambda, \nu) = \inf_{x \in \mathcal{D}} \mathcal{L}(x, \lambda, \nu) \leq \mathcal{L}(\tilde{x}, \lambda, \nu) \leq f_0(\tilde{x}).$$

Since $\tilde{x}$ is an arbitrary feasible point

$$d(\lambda, \nu) \leq p^*.$$

This is a non-trivial bound only if $d(\lambda, \nu) \neq -\infty$, e.g. when $\lambda_i \geq 0$, $i = 1, \ldots, m$ and $(\lambda, \nu) \in \mathrm{dom}(d)$. Such a pair (λ, ν) is called *dual feasible*. For any dual feasible pair (λ, ν), the difference

$$p^* - d(\lambda, \nu) \geq 0 \tag{2.28}$$

is called the *duality gap* corresponding to the dual feasible pair (λ, ν). Finally, we call the optimization problem

$$\max_{\lambda, \nu \in \mathbb{R}^n_+} d(\lambda, \nu),$$

the *dual problem*. The optimization problem (2.29) is a concave, possibly non-differentiable, optimization problem. We can reformulate this optimization problem to be a convex minimization problem by

$$\max_{\lambda, \nu \in \mathbb{R}^n_+} d(\lambda, \nu) \iff \min_{\lambda, \nu \in \mathbb{R}^n_+} -d(\lambda, \nu). \tag{2.29}$$

2.3.2 Subgradient methods

In many applications the functions f_i for some $i = 0, \ldots, m$ or h_j for some $j = 1, \ldots, p$ in Equation (2.25) might be non-differentiable. Thus, the necessity for numerical optimization techniques that can work around the non-differentiability of the resulting Lagrangian or dual function arise. First, we define the subgradient of a non-differentiable function.

Definition 2.3.1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex C^0-function. A vector $g \in \mathbb{R}^n$ is called subgradient of f in x_0 if for all $x \in \mathbb{R}^n$*

$$f(x) \geq f(x_0) + \langle g, x - x_0 \rangle \tag{2.30}$$

holds. Here $\langle \cdot, \cdot \rangle$ denotes the scalar product. The subdifferential $\partial f(x_0)$ is the set of all subgradients of f in x_0.

We will use the subdifferential to write down a basic subgradient method for the unconstrained convex non-differentiable dual problem in (2.29). The subgradient method for convex optimization problems uses the simple iteration

$$x^{(k+1)} = x^{(k)} - \alpha_k g^{(k)}, \tag{2.31}$$

where $x^{(k)}$ is the kth iterate, $g^{(k)}$ is any subgradient of f at $x^{(k)}$, i.e. $g^{(k)} \in \partial f(x^{(k)})$ and $\alpha_k > 0$ is the kth step size. It might happen that $-g^{(k)}$ is not a descent direction such that we have $f(x^{(k+1)}) > f(x^{(k)})$. This can be verified by a simple example.

Example 2.3.1 ($-g^{(k)}$ is not a descent direction)**.** *Consider the absolute value function $f(x) = |x|$, for $x \in \mathbb{R}$. f is not differentiable in $x = 0$ with subdifferential $\partial f(0) = [-1, 1]$. If we choose the subgradient $g^{(0)} = 1$ in the first iteration starting at $x_0 = 0$ it is clear that for any step size $\alpha_0 \neq 0$*

$$0 = f(0) = f(x_0) < f(x_0 - \alpha_0 \cdot 1) = f(x_1).$$

Thus, it is common to keep track of the best iterate found so far. At each step, set

$$f_{\text{best}}^{(k)} = \min\{f_{\text{best}}^{(k-1)}, f(x^{(k)})\}.$$

We can write the best objective value found in k iterations as

$$f_{\text{best}}^{(k)} = \min\{f(x^{(1)}), \ldots, f(x^{(k)})\}.$$

The sequence $f_{\text{best}}^{(k)}$ is decreasing and has a limit (possibly $-\infty$). Standard gradient descent methods rely on the function values in each iteration to decrease whereas for subgradient methods this is not a guarantee. The convergence results depend on the choice of the step size rule and the distance from the optimal set itself. Namely, suppose x^* is an arbitrary point that minimizes the objective function f. It holds

$$\begin{aligned} \|x^{(k+1)} - x^*\|_2^2 &= \|x^{(k)} - \alpha_k g^{(k)} - x^*\|_2^2 \\ &= \|x^{(k)} - x^*\|_2^2 - 2\alpha_k g^{(k)T}(x^{(k)} - x^*) + \alpha_k^2 \|g^{(k)}\|_2^2 \\ &\leq \|x^{(k)} - x^*\|_2^2 - 2\alpha_k (f(x^{(k)}) - f(x^*)) + \alpha_k^2 \|g^{(k)}\|_2^2. \end{aligned} \tag{2.32}$$

The latter is due to the definition of the subgradient in Equation (2.30), i.e.

$$f(x^*) \geq f(x^{(k)}) + g^{(k)T}(x^* - x^{(k)}).$$

The convergence analysis relies on two main assumptions:

1. An upper bound R for the distance of the initial point to the optimal points is known, i.e. there exists a constant $R > 0$ such that

$$\|x^{(1)} - x^*\|_2 \leq R. \tag{2.33}$$

2. The norm of the subgradients is bounded, i.e. for all iterations k there exists a constant $S > 0$ such that

$$\|g^{(k)}\|_2 \leq S. \tag{2.34}$$

We can repeat the inequality derivation in (2.32) recursively to get

$$0 \leq \|x^{(k+1)} - x^*\|_2^2 \leq \|x^{(1)} - x^*\|_2^2 - 2\sum_{i=1}^{k} \alpha_i (f(x^{(i)}) - f(x^*)) + \sum_{i=1}^{k} \alpha_i^2 \|g^{(i)}\|_2^2. \tag{2.35}$$

Rearranging (2.35) and using Assumption 1, Equation (2.33), we have

$$2\sum_{i=1}^{k} \alpha_i (f(x^{(i)}) - f(x^*)) \leq R^2 + \sum_{i=1}^{k} \alpha_i^2 \|g^{(i)}\|_2^2. \tag{2.36}$$

Combining (2.36) with

$$\sum_{i=1}^{k} \alpha_i (f(x^{(i)}) - f(x^*)) \geq \left(\sum_{i=1}^{k} \alpha_i\right) \min_{i=1,\dots,k} \left(f(x^{(i)}) - f(x^*)\right) = \left(\sum_{i=1}^{k} \alpha_i\right) \left(f_{\text{best}}^{(k)} - f(x^*)\right)$$

yields

$$f_{\text{best}}^{(k)} - f(x^*) = \min_{i=1,\dots,k} f(x^{(i)}) - f(x^*) \leq \frac{R^2 + \sum_{i=1}^{k} \alpha_i^2 \|g^{(i)}\|_2^2}{2 \sum_{i=1}^{k} \alpha_i}.$$

Finally, using Assumption 2, Equation (2.34), we obtain the main inequality for the upcoming convergence results

$$f_{\text{best}}^{(k)} - f(x^*) \leq \frac{R^2 + S^2 \sum_{i=1}^{k} \alpha_i^2}{2 \sum_{i=1}^{k} \alpha_i}. \tag{2.37}$$

We can now apply this inequality to different step size rules to get convergence results depending on the step size rule.

Corollary 2.3.1 (Step size rules and their convergence)**.** *The following step size rules together with their different convergence result using (2.37) may be used in the subgradient method.*

1. *Constant step sizes, i.e. $\alpha_k = \alpha$, where α is a positive constant independent on the current iteration k. Inequality (2.37) reduces to*

$$f_{\text{best}}^{(k)} - f(x^*) \leq \frac{R^2 + S^2\alpha^2 k}{2\alpha k} \xrightarrow{k\to\infty} \frac{S^2\alpha}{2},$$

   i.e. $f_{\text{best}}^{(k)}$ converges to within $S^2\alpha/2$-optimal within at most $R^2/(S^2\alpha^2)$ steps.

2. *Constant step length, i.e. $\alpha_i = \gamma/\|g^{(i)}\|_2$, where $\gamma > 0$. Using $\alpha_i \geq \gamma/S$, inequality (2.37) reduces to*

$$f_{\text{best}}^{(k)} - f(x^*) \leq \frac{R^2 + \gamma^2 k}{2 \sum_{i=1}^{k} \alpha_i} \leq \frac{R^2 + \gamma^2 k}{2\gamma k/S} \xrightarrow{k\to\infty} \frac{S\gamma}{2}$$

   i.e. $f_{\text{best}}^{(k)}$ converges to within $S\gamma/2$-optimal.

3. *Square summable but not summable, i.e. a step size α_i such that*

$$\alpha_i \geq 0, \quad \sum_{i=1}^{\infty} \alpha_i^2 < \infty, \quad \sum_{i=1}^{\infty} \alpha_i = \infty.$$

   In this case, inequality (2.37) reduces to

$$f_{\text{best}}^{(k)} - f(x^*) \leq \frac{R^2 + S^2\|\alpha\|_2^2}{2 \sum_{i=1}^{k} \alpha_i} \xrightarrow{k\to\infty} 0,$$

   i.e. the subgradient method converges to the optimum.

More examples can be found in [40].

Additionally, we may be interested in the finite sequence of $K \in \mathbb{N}$ step sizes $\{\alpha_i\}_{i=1}^K$ that minimizes the right hand side of inequality (2.37). Since the function

$$B(\alpha_1, \ldots, \alpha_K) := \frac{R^2 + S^2 \sum_{i=1}^K \alpha_i^2}{2 \sum_{i=1}^K \alpha_i}$$

is convex and symmetric in $\alpha_1, \ldots, \alpha_K$, we can use standard calculus to deduce that the optimum is attained when all α_i's are equal. This reduces, for all $\alpha_i \equiv \alpha$, the sub-optimality bound in (2.37) to

$$\frac{R^2 + S^2 K \alpha^2}{2K\alpha},$$

which is minimized by $\alpha = (R/S)/\sqrt{K}$ yielding

$$f_{\text{best}}^{(K)} - f(x^*) \leq RS/\sqrt{K}. \tag{2.38}$$

This means, if we want to achieve a guaranteed accuracy of $\text{TOL}_{\text{SG}} > 0$, the number of steps K^* needed in the subgradient method has to satisfy

$$\left(\frac{RS}{\text{TOL}_{\text{SG}}}\right)^2 \leq K^*. \tag{2.39}$$

This tells us, that the subgradient method is very slow if the constants in Assumption 1 and 2 are too large, i.e. we are also interested in finding good initial values for the subgradient method.

Chapter 3

Quadratic Optimal Control of Ornstein-Uhlenbeck Processes

This chapter will build the groundwork for handling chance constraints in optimal control of SDEs in Chapter 4. For this cause, we will use an optimal control example in which we control Ornstein-Uhlenbeck type dynamics introduced in Equation (2.12). This was briefly motivated in Section 2.1 prior to introducing some basic properties of Ornstein-Uhlenbeck processes in Example 2.1.1. Independent of the wide range of applications, e.g. finance [42], biology [43] or physics [44], it is also very appealing to investigate as a toy problem from a mathematical view. The Ornstein-Uhlenbeck process is a mean-reverting, linear, stationary and Gaussian process. There exists an explicit representation of the solution to the SDE (2.12) with initial condition (2.13), namely Equation (2.7), and thus you can compute mean, variance, covariances and the full underlying probability distribution at all times $t \in [0, T]$. Additionally, the Ornstein-Uhlenbeck process reverts to a long term mean x_∞. In contrast, Brownian motion for example drifts off to infinity. This enables modelling of realistic scenarios in which a system fluctuates around a stable equilibrium. Gaussian processes are well studied and widely used in machine learning and filtering techniques [45, 46]. The mean of the Ornstein-Uhlenbeck process $X(t)$ defined in (2.12) with initial condition (2.13) and $x_\infty = 0$ can be written as

$$m(t) := E[X(t)] = x_0 \cdot e^{-\theta t}, \quad t \in [0, T].$$

Thus, the mean can be described by the ODE

$$\frac{dm(t)}{dt} = -\theta m(t), \quad m(0) = x_0. \tag{3.1}$$

In this chapter we are considering a controlled version of the Ornstein-Uhlenbeck process

$$dX(t) = (-\theta X(t) + \alpha(t, X(t)))dt + \sigma dW(t), \quad 0 < t < T, \tag{3.2}$$
$$X(0) = x_0 \in \mathbb{R}, \tag{3.3}$$

where the diffusion $\sigma > 0$ is constant and the drift is additively controlled. The objective will be a quadratic function in state $X(t)$ and control $\alpha(t, X(t))$, i.e.

$$C_{t,x}(\alpha) = \mathbb{E}\left[\int_t^T X(t)^2 + \mu \cdot \alpha(t, X(t))^2 dt \;\middle|\; X(t) = x\right],$$

where $\mu > 0$ is a regularization parameter for the controls, meaning that larger values for μ makes controlling the SDE more expensive in terms of running costs. The quadratic objective will result in an analytic representation of the optimal control, making the problem as appealing, as an example, as the Ornstein-Uhlenbeck process for uncontrolled SDEs. The problem can be formulated as follows.

Problem 1. *Consider the following optimization problem,*

$$\min_{\alpha \in \mathcal{A}} \mathbb{E}\left[\int_0^T X(t)^2 + \mu \cdot \alpha(t, X(t))^2 \, dt\right]$$

over the set of admissible deterministic Markovian controls $\mathcal{A}$, which are adapted to the filtration $\mathcal{F}_t$ of the process $X(t) \in \mathbb{R}$ for all $t \in [0, T]$ which follows the controlled dynamics

$$\begin{cases} dX(t) &= (-\theta X(t) + \alpha(t, X(t)))\, dt + \sigma \cdot dW(t), \quad 0 < t < T, \\ X(0) &= x_0, \end{cases}$$

for given parameters $\theta > 0, \mu > 0, \sigma, x_0 \in \mathbb{R}$ and finite terminal time $T > 0$.

We can compare the mean behaviour of this SOC problem by investigating the deterministic OC problem in which the underlying dynamics is described by the ODE in (3.1) describing the mean of the Ornstein-Uhlenbeck process. In the deterministic framework, the controls will not be state-dependent due to the fact that the state is uniquely determined by the ODE. The problem will reduce to the following.

Problem 2. *Consider the following optimization problem,*

$$\min_{\alpha \in \mathcal{A}_d} \left[\int_0^T Y(t)^2 + \mu \cdot \alpha(t)^2 \, dt\right]$$

over the set of control functions $\mathcal{A}_d := \{\alpha : [0, T] \to A\}$ and subject to the controlled deterministic ODE

$$\begin{aligned} \frac{dY(t)}{dt} &= (-\theta Y(t) + \alpha(t)), \quad 0 < t < T, \\ Y(0) &= x_0, \end{aligned}$$

for given parameters $\theta > 0$, $\mu > 0$, $x_0 \in \mathbb{R}$ and $T > 0$. Here A is a compact subset of $\mathbb{R}$.

We differentiate between Problem 1 and Problem 2 by using $X(t)$ for the solution to the controlled SDE and $Y(t)$ for the solution to the controlled ODE, respectively. Comparing both resulting optimal solutions will not only give insights into the theoretical differences between SOC and OC but on resulting similarities of the problems at hand and their interpretation. The main objectives of this chapter are

3.1 State the dynamic programming relation and derive the HJB equation for Problem 1

3.2 Derive the Pontryagin principle and HJB for Problem 2 and propose a numerical discretization

3.3 Construct a stable numerical scheme to solve the corresponding HJB PDE

3.4 Use the Ricatti Ansatz from quadratic optimal control to reduce Problem 1 to the solution of Ricatti ODEs

3.1 Dynamic Programming & Hamilton-Jacobi-Bellman equation

The dynamic programming relation from Equation (2.21) for Problem 1 is straight forward. Using Theorem 2.2.1, we know that for all $\varepsilon > 0$, the following dynamic programming relation holds

$$u(t,x) = \min_{\alpha:[t,t+\varepsilon]\to A} \mathbb{E}\left[\int_t^{t+\varepsilon} X(s)^2 + \mu\cdot\alpha(s,X(s))^2 ds + u(t+\varepsilon, X(t+\varepsilon)) \;\middle|\; X(t)=x\right].$$

Furthermore using Theorem 2.2.2, the HJB equation for Problem 1 is given by

$$0 = \partial_t u(t,x) + \min_{\alpha\in A}\left[(-\theta x+\alpha)\partial_x u(t,x) + \frac{\sigma^2}{2}\partial_{xx}u(t,x) + x^2 + \mu\cdot\alpha^2\right], \tag{3.4}$$

$$u(T,x) = 0, \tag{3.5}$$

for $(t,x)\in[0,T)\times\mathbb{R}$. We can derive an analytic representation of the optimal control by differentiating the Hamiltonian function H w.r.t. α

$$\bar{H}(\alpha) = (-\theta x+\alpha)\partial_x u + \frac{\sigma^2}{2}\partial_{xx}u + x^2 + \mu\cdot\alpha^2$$

and using the first order optimality condition $d\bar{H}/d\alpha = 0$ yielding

$$\alpha^*(t,x) = -\frac{\partial_x u(t,x)}{2\mu}, \quad (t,x)\in[0,T]\times\mathbb{R}. \tag{3.6}$$

Substituting α^* into Equation (3.4) yields the non-linear PDE

$$0 = \partial_t u - \theta x\cdot\partial_x u - \frac{1}{4\mu}(\partial_x u)^2 + \frac{\sigma^2}{2}\partial_{xx}u + x^2, \quad (t,x)\in[0,T)\times\mathbb{R} \tag{3.7}$$

$$u(T,x) = 0. \tag{3.8}$$

In Equation (3.7) you can see the non-linearity $(\partial_x u)^2$ of the HJB equation. When building a PDE solver for HJB equations, see Section 3.3, we will stick with the representation involving taking a minimum over controls as in Equation (3.4) or in more general Equation (2.22). This is done because we would like to have a PDE solver that can solve other SOC problems as well in which we might not have the luxury of an analytic representation of the optimal control as in Equation (3.6).

3.2 Deterministic Optimal Control

In deterministic optimal control theory it is possible to compute optimal controls and optimal solution paths using characteristics of the HJB equation. In general, Problem 2, can be formulated as the following optimization problem:

$$\min_{\alpha\in\mathcal{A}_d}\left[\int_0^T h(t,Y(t),\alpha(t))dt + g(Y(T))\right],$$

subject to the controlled ODE

$$\begin{aligned} Y'(t) &= a(t, Y(t), \alpha(t)), \quad 0 < t \leq T, \\ Y(0) &= x_0. \end{aligned}$$

The HJB equation in this case is given by

$$0 = \partial_t u + \min_{\alpha \in A} H(t, y, \partial_y u), \tag{3.9}$$

$$u(T, y) = g(y), \tag{3.10}$$

with Hamiltonian H taking on the form

$$H(t, y, \alpha, \partial_y u) = a(t, y, \alpha)\partial_y u(t, y) + h(t, y, \alpha).$$

For problems with higher dimensional underlying dynamics $1 \ll d$, the HJB is numerically intractable because PDE solver's computational complexity increase exponentially w.r.t. the dimension of the problem. For deterministic problems the PMP can be used to formulate a solution methodology independent of solving a PDE. The PMP provides necessary conditions for optimality in OC problems. To apply the PMP, we define the Hamiltonian H using an adjoint variable $\lambda(t)$, sometimes called costate,

$$H(t, y, \alpha, \lambda(t)) = a(t, y, \alpha)\lambda(t) + h(t, y, \alpha).$$

The adjoint variable $\lambda(t)$ satisfies the ODE, running backwards in time,

$$\frac{d\lambda(t)}{dt} = -\partial_y H(t, y, \alpha, \lambda(t)), \quad \lambda(T) = \partial_y g(Y(T)). \tag{3.11}$$

Equation (3.11) can be derived using Lagrangian relaxation from Calculus of Variations for the problem statement above. A derivation can be read in Section 8.1.2 of [36]. At each time $t \in [0, T]$, the optimal control $\alpha^*(t)$ minimizes the Hamiltonian

$$\alpha^*(t) = \operatorname*{argmin}_{\alpha \in A} H(t, y, \alpha, \lambda(t)).$$

Summarized, the PMP states that the optimal control $\alpha^*(t)$, corresponding state $Y^*(t)$ and costate $\lambda^*(t)$ must satisfy

(i) State dynamics for $t \in [0, T]$:

$$\frac{dY^*(t)}{dt} = a(t, Y^*(t), \alpha^*(t)), \quad Y^*(0) = x_0.$$

(ii) Adjoint dynamics for $t \in [0, T]$:

$$\begin{aligned} \frac{d\lambda^*(t)}{dt} &= -\partial_y h(t, Y^*(t), \lambda^*(t)) - \lambda^*(t) \cdot \partial_y a(t, Y^*(t), \alpha^*(t)), \\ \lambda^*(T) &= \partial_y g(Y^*(T)). \end{aligned}$$

(iii) Optimal control:

$$\alpha^*(t) = \operatorname*{argmin}_{\alpha \in A} H(t, y, \alpha, \lambda(t)).$$

For our explicit Problem 2, this reduces to solving the following Hamiltonian system

$$Y'(t) = -\theta Y(t) - \frac{\lambda(t)}{2\mu}, \quad Y(0) = x_0, \tag{3.12}$$

$$\lambda'(t) = \theta\lambda(t) - 2Y(t), \quad \lambda(T) = 0. \tag{3.13}$$

Solving the system in (3.12) and (3.13) numerically using a standard explicit Euler method is not possible due to the fact that the second equation (3.13) is backwards in time. We can discretize this Hamiltonian system using a Symplectic Euler method which essentially is a combination of explicit and implicit Euler. Using an equidistant timestepping scheme

$$\{t_k\}_{k=0}^N, \quad \text{with} \quad \Delta t = t_1 - t_0,$$

and the notation $Y_n := Y(t_n)$, for $n = 0, \dots, N$, we will use an explicit Euler step for Equation (3.12) and an implicit Euler step for Equation (3.13), i.e.

$$\begin{aligned} &\text{[Explicit]} \quad Y_{n+1} = Y_n + \Delta t \cdot \partial_\lambda H(\lambda_{n+1}, Y_n), \qquad X_0 = x_0, \\ &\text{[Implicit]} \quad \lambda_n = \lambda_{n+1} + \Delta t \cdot \partial_y H(\lambda_{n+1}, Y_n), \quad \lambda_N = \partial_y g(Y_N). \end{aligned}$$

For Problem 2 this reduces to a coupled forward-backward system

$$\begin{aligned} Y_{n+1} &= Y_n + \Delta t \left(-\theta Y_n - \frac{\lambda_{n+1}}{2\mu} \right), \qquad Y_0 = X_0, \\ \lambda_n &= \lambda_{n+1} - \Delta t \left(-\theta\lambda_{n+1} + 2Y_n \right), \quad \lambda_N = 0. \end{aligned}$$

Given the initial value x_0 and a guess initial value $\lambda_0 = \ell_0 \in \mathbb{R}$ we can rewrite the forward-backward system into a forward-forward system by

$$\begin{aligned} \lambda_{n+1} &= \lambda_n \frac{1}{1-\theta\Delta t} - \frac{2\Delta t}{1-\theta\Delta t} Y_n, \quad \lambda_0 = \ell_0, \\ Y_{n+1} &= Y_n(1-\theta\Delta t) - \frac{\Delta t}{2\mu}\lambda_{n+1}, \quad Y_0 = x_0. \end{aligned}$$

This allows us to write the n-th value as a function of the previously computed values or inputs (x_0, ℓ_0), i.e. for $n = 1, \dots, N$

$$\lambda_n = \ell_0(1-\theta\Delta t)^{-n} - \frac{2\Delta t}{(1-\theta\Delta t)} \left(\sum_{k=0}^{n-1} Y_k (1-\theta\Delta t)^{k-n+1} \right), \tag{3.14}$$

$$Y_n = x_0(1-\theta\Delta t)^n - \frac{\Delta t}{2\mu} \left(\sum_{k=1}^{n} \lambda_k (1-\theta\Delta t)^{n-k} \right). \tag{3.15}$$

We will use a shooting method to solve this system of equations which uses a root-finding algorithm to find the value $\ell_0 \in \mathbb{R}$ such that the iteration above returns $\lambda_N = 0$. In more detail: For a given discretization t_k, $k = 0, \dots, N$, we define a function

$$\Lambda : \mathbb{R} \to \mathbb{R}, \ell_0 \mapsto \lambda_N(\ell_0)$$

that maps the input guess $\ell_0 \in \mathbb{R}$ to the terminal costate λ_N given the iteration rules in Equation (3.14) and Equation (3.15). For this function, we are looking for a value $\ell_0^* \in \mathbb{R}$ such that

$$\Lambda(\ell_0^*) = 0,$$

which can be found by any root-finding procedure like a bisection or Newton method. We can obtain the optimal control using the third part of the PMP (iii), i.e.

$$\alpha^*(t) = \underset{\alpha \in A}{\operatorname{argmin}}\, H(t, y, \alpha, \lambda(t)) = \underset{\alpha \in A}{\operatorname{argmin}}\, \left[(-\theta y + \alpha)\lambda(t) + y^2 + \mu \cdot \alpha^2\right] = -\frac{\lambda(t)}{2\mu}.$$

Thus, we can approximate the optimal control and the value function by

$$\alpha_n^* = \frac{-\lambda_n}{2\mu}, \quad \text{and} \quad u_n = \sum_{k=n}^{N} \left(Y_k^2 + \mu(\alpha_k^*)^2\right) \Delta t, \tag{3.16}$$

where $\lambda_n = \lambda(t_n)$ and $\alpha_n = \alpha(t_n)$ for $n = 0, \ldots, N$. A summarizing algorithm is written down below in Algorithm 1. Numerical results are presented in Section 3.5.2.

Algorithm 1 Algorithm for solving OC using PMP

Require: Step size choice $\{t_k\}_{k=0}^N$, initial guess ℓ_0.
Solve $\Lambda(\ell_0^*) = 0$ using root-finding algorithm.
Compute state Y_n and costate λ_n arrays using (3.14) and (3.15) for $n = 0, \ldots, N$.
Compute optimal control α_n^* and value function u_n using (3.16).
return Optimal control α_n^* and value function u_n for $n = 0, \ldots, N$.

3.3 An Upwind Finite Difference scheme for HJB

The PMP is not applicable in the SOC setting because of the state-dependency of the control. Thus, we are interested in solving the HJB equation numerically. The main goal of this section is to contruct a PDE solver for the HJB equation in Equation (2.22) with terminal time boundary condition (2.23). We will use the proposed UFD scheme by Sun et al. in [48] and modify it to be able to solve the HJB equation resulting from the optimal control of SDEs. We choose the UFD scheme because of its effectiveness and well-adaptiveness. It has been studied thoroughly and it was shown that the solution converges to the solution of the corresponding OC problem [48]. Additionally, the UFD scheme is explicit in time such that we can compute the next step by taking the previous steps' point-wise optimal control. Other possible ways to solve the HJB can be read in SOC literature like [49, 50]. Recall the HJB equation from Equation (2.22) with terminal time boundary condition (2.23) for $(t, x) \in [0, T) \times \mathbb{R}$

$$\begin{aligned} 0 &= \partial_t u(t,x) + \min_{\alpha \in A} \left[a(t,x,\alpha)\partial_x u(t,x) + \frac{\sigma^2}{2}\partial_{xx} u(t,x) + h(t,x,\alpha) \right], \\ u(T,x) &= g(x). \end{aligned}$$

For implementation purposes, we have to restrict the state space to a compact subset $[c, d] \subset \mathbb{R}$, as well as impose boundary conditions on the boundary of the restriction $\{c, d\}$ to ensure uniqueness of the numerical solutions. We will impose the boundary condition

$$\frac{\partial^2 u}{\partial x^2} = 0 \quad \text{on} \quad \{c, d\}, \quad \text{for all } t \in [0, T], \tag{3.17}$$

which can be interpreted as a linear interpolation at the boundaries. This is a natural choice when no further information about the problem at hand is given. Given positive

integers N_1, M_1 corresponding to the time and space discretization respectively, set

$$t_n = n \cdot \Delta t,\ n = 0, \dots, N_1, \qquad \Delta t = \frac{T}{N_1},$$
$$x_m = c + m \cdot \Delta x,\ m = 0, \dots, M_1,\ \Delta x = \frac{|d-c|}{M_1}$$

and define, using $u_m^n := u(t_n, x_m)$,

$$a_m^{n^+}(\alpha) = \max\{a(t_n, x_m, \alpha), 0\}, \qquad D^+ u_m^n = \frac{(u_{m+1}^n - u_m^n)}{\Delta x},$$
$$a_m^{n^-}(\alpha) = \min\{a(t_n, x_m, \alpha), 0\}, \qquad D^- u_m^n = \frac{(u_m^n - u_{m-1}^n)}{\Delta x}.$$

Deviating from the HJB equation in [48], we have an additional second order partial derivative in the Hamiltonian which we will approximate using a central second order finite difference, i.e.

$$D^2 u_m^n = \frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{(\Delta x)^2}, \quad m = 1, \dots, M_1 - 1.$$

We will use the following Upwind Finite Difference (UFD) scheme

$$\frac{u_m^{n-1} - u_m^n}{\Delta t} = a_m^{n^+}(\alpha_m^n) \cdot D^+ u_m^n + a_m^{n^-}(\alpha_m^n) \cdot D^- u_m^n + \frac{\sigma^2}{2} D^2 u_m^n + h(t_n, x_m, \alpha_m^n), \tag{3.18}$$

where

$$\alpha_m^n = \underset{\alpha \in A}{\operatorname{argmin}}\ H_m^n(\alpha), \tag{3.19}$$
$$H_m^n(\alpha) = \left[a_m^{n^+}(\alpha) \cdot D^+ u_m^n + a_m^{n^-}(\alpha) \cdot D^- u_m^n + \frac{\sigma^2}{2} D^2 u_m^n + h(t_n, x_m, \alpha) \right]. \tag{3.20}$$

In general, the minimization in Equation (3.19) can be solved using a numerical minimization procedure. For Problem 1, we can solve the argmin-problem analytically by differentiating between the following two cases arising from the nature of the UFD scheme. In the first case, the drift is positive and we consider foward finite-difference approximations. In the second case the drift is negative and we consider backwards finite-difference approximations, i.e.

Case 1: $\alpha_m^{n+1} \geq \theta x_m$

$$\alpha_m^n = \operatorname{argmin}_{\alpha \in \mathbb{R}} \left[(-\theta x_m + \alpha) D^+ u_m^n + x^2 + \mu \alpha^2 \right] = -\frac{1}{2\mu} D^+ u_m^n.$$

Case 2: $\alpha_m^{n+1} < \theta x_m$

$$\alpha_m^n = \operatorname{argmin}_{\alpha \in \mathbb{R}} \left[(-\theta x_m + \alpha) D^- u_m^n + x^2 + \mu \alpha^2 \right] = -\frac{1}{2\mu} D^- u_m^n.$$

In addition, we impose the boundary condition in Equation (3.17) using a central finite difference at the boundary $\{c, d\}$, i.e.

$$0 \overset{!}{=} \frac{\partial^2 u}{\partial x^2} \approx \frac{u_m^n - 2u_{m-1}^n + u_{m-2}^n}{(\Delta x)^2}, \quad \text{for } m = 2, M_1$$
$$\implies u_{M_1}^n = 2u_{M_1-1}^n - u_{M_1-2}^n \quad \text{and} \quad u_0^n = 2u_1^n - u_2^n$$

Finally, we can write down a compact algorithm for solving the HJB equation (3.4) with arbitrary terminal time boundary condition $u(T,x) = g(x)$ for some function g in x. Note that we are only considering an optimal control setting without a controlled diffusion. For this case, we can simply approximate the additional term with plain central finite differences. Additionally, we are considering arbitrary terminal time boundary conditions in preparation for the next chapter. The algorithm can be described as follows.

Algorithm 2 UFD scheme for Problem 1 with arbitrary terminal costs

Require: $u_m^{N_1} = g(x_m)$ and $\alpha_m^{N_1} = 0$ for all $m = 0, \dots, M_1$
  for $n = N_1, \dots, 1$ **do**
    $u_1^n \leftarrow 2u_2^n - u_3^n$ ▷ Imposed BC
    $u_{M_1}^n \leftarrow 2u_{M_1-1}^n - u_{M_1-2}^n$
    for $m = 1, \dots, M_1 - 1$ **do**
      if $\alpha_m^n \geq \theta x_m$ **then** ▷ Cases for optimal control
        $\alpha_m^{n-1} \leftarrow -1/(2\mu) \cdot D^+ u_m^n$
      else if $\alpha_m^n < \theta x_m$ **then**
        $\alpha_m^{n-1} \leftarrow -1/(2\mu) \cdot D^- u_m^n$
      end if
      $u_m^{n-1} \leftarrow u_m^n + \Delta t \cdot H_m^n(\alpha_m^n)$ ▷ UFD scheme
    end for
  end for
  $u_1^0 \leftarrow 2u_2^0 - u_3^0$
  $u_{M_1}^0 \leftarrow 2u_{M_1-1}^0 - u_{M_1-2}^0$

For the stability of Algorithm 2, we are assuming the Courant-Friedrichs-Lewy (CFL) condition

$$\frac{\Delta t}{\Delta x}\left(\|a\|_\infty + \frac{\sigma^2}{\Delta x}\right) \leq 1 \tag{3.21}$$

$$\implies T \cdot \frac{M_1}{|d-c|}\left(\|a\|_\infty + \frac{\sigma^2 \cdot M_1}{|d-c|}\right) \leq N_1, \tag{3.22}$$

where

$$\|a\|_\infty := \max_{(t,x,\alpha)\in[0,T]\times[c,d]\times A} |a(t,x,\alpha)|. \tag{3.23}$$

Algorithm 2 describes a first order finite difference scheme, i.e. the absolute error can be bounded by

$$|u(0,x_0) - u^{N_1,M_1}(0,x_0)| \leq \frac{c_1}{N_1} + \frac{\bar{c}_1}{M_1}, \quad c_1, \bar{c}_1 > 0, \tag{3.24}$$

where u is the analytic solution and u^{N_1,M_1} is the approximation using the proposed UFD scheme. Due to the CFL condition, estimating the constant $c_1 > 0$ is impractical. Thus, we will further bound the error in (3.24) using (3.22), i.e.

$$|u(0,x_0) - u^{N_1,M_1}(0,x_0)| \leq \frac{c_1}{N_1} + \frac{\bar{c}_1}{M_1} \overset{(3.22)}{\leq} \tilde{c}_1 \cdot M_1^{-1}. \tag{3.25}$$

The computational work, using Algorithm 2, to compute an approximation is

$$\mathcal{W} = \mathcal{O}(N_1 \cdot M_1).$$

Here $\mathcal{O}$ is the big-O-notation meaning, that the work is at most a positive constant multiple, say $C > 0$, of the absolute value of $N_1 \cdot M_1$, i.e.

$$\mathcal{W} \leq C \cdot N_1 \cdot M_1.$$

A result on the optimal work is given in the following Theorem 3.3.1. We aim to satisfy a prescribed absolute tolerance $\mathrm{TOL} > 0$ on (3.25) as well as the CFL condition (3.22).

Theorem 3.3.1 (Optimal Work UFD scheme)**.** *Consider the following optimization problem*

$$\begin{aligned} \min_{N,M} \ \mathcal{W} &= \min_{N,M} \ N_1 \cdot M_1, \\ s.t. \quad \tilde{c}_1 \cdot M_1^{-1} &\approx \mathrm{TOL}, \\ T \cdot \frac{M_1}{|d-c|} \left(\|a\|_\infty + \frac{\sigma^2 \cdot M_1}{|d-c|} \right) &\leq N_1. \end{aligned}$$

Then

$$\mathcal{W}^* = \mathcal{O}(\mathrm{TOL}^{-3})$$

and

$$N_1^* = \frac{T\|a\|_\infty \tilde{c}_1}{|d-c|} \cdot \mathrm{TOL}^{-1} + \frac{\sigma^2 T \tilde{c}_1^2}{|d-c|^2} \cdot \mathrm{TOL}^{-2}, \quad M_1^* = \tilde{c}_1 \cdot \mathrm{TOL}^{-1}.$$

Proof. The proof is found in Appendix A.4. □

3.4 Ricatti Ansatz in Quadratic Optimal Control

The Ricatti Ansatz assumes a quadratic value function in the state x for the solution to the HJB equation (3.7), i.e. the value function u is of the following structure

$$u(t,x) = f(t) \cdot x^2 + \ell(t), \quad (t,x) \in [0,T] \times \mathbb{R} \tag{3.26}$$

with functions in time $f, \ell : [0,T] \to \mathbb{R}$. We can reduce the problem of solving the HJB (3.7) into solving a system of Ricatti ODEs. This can be constructed as follows. For the partial derivatives in assumption (3.26), we have

$$\begin{aligned} \partial_t u(t,x) &= f'(t)x^2 + \ell'(t), & \partial_x u(t,x) &= 2f(t)x, \\ \partial_{xx} u(t,x) &= 2f(t), & (\partial_x u(t,x))^2 &= 4f^2(t)x^2. \end{aligned}$$

Substituting this Ansatz's partial derivatives into the HJB equation (3.7) with terminal time boundary condition (3.8) yields

$$\begin{aligned} 0 &= f'(t)x^2 + \ell'(t) - \theta x \cdot 2f(t)x + x^2 - \frac{1}{4\mu}(2f(t)x)^2 + \frac{\sigma^2}{2} 2f(t), \\ 0 &= f(T)x^2 + \ell(T). \end{aligned}$$

Factoring out x^2 reduces the problem to the solution of a system of ODEs with a Ricatti ODE

$$0 = f' - 2\theta f - \frac{f^2}{\mu} + 1, \quad f(T) = 0, \quad \ell(t) = -\sigma^2 \cdot \int_t^T f(s)\, ds. \tag{3.27}$$

The optimal control can be obtained by

$$\alpha^*(t,x) = -\frac{\partial_x u(t,x)}{2\mu} = -\frac{f(t)x}{\mu} \tag{3.28}$$

and the optimally controlled SDE can thus be written as

$$dX(t) = \left(-\theta - \frac{f(t)}{\mu}\right) X(t)\, dt + \sigma \cdot dW(t), \quad X(0) = x_0. \tag{3.29}$$

First, we will analyse the Ricatti ODE in Equation (3.27). We can write the non-linear first order Ricatti ODE in Equation (3.27) as a linear second order ODE

$$w'' - 2\theta w' - \frac{1}{\mu} w = 0, \quad w(T) = 0,$$

with two fundamental solutions

$$w_1(t) = \exp(\lambda_1 t), \quad w_2(t) = \exp(\lambda_2 t)$$

where $\lambda_1 = \theta + \sqrt{\theta^2 + \frac{1}{\mu}}$, $\lambda_2 = \theta - \sqrt{\theta^2 + \frac{1}{\mu}}$. The analytic solution to the Ricatti ODE is then given by

$$f(t) = -\mu \frac{d}{dt} \left(\log(c_1 w_1(t) + c_2 w_2(t))\right) = -\mu \frac{c_1 \lambda_1 w_1(t) + c_2 \lambda_2 w_2(t)}{c_1 w_1(t) + c_2 w_2(t)}.$$

The final time condition $f(T) = 0$ implies

$$f(T) = -\mu \frac{c_1 \lambda_1 w_1(T) + c_2 \lambda_2 w_2(T)}{c_1 w_1(T) + c_2 w_2(T)} \overset{!}{=} 0 \implies c_2 = -\frac{c_1 \lambda_1 w_1(T)}{\lambda_2 w_2(T)}$$

and substituting back yields

$$f(t) = -\mu \cdot \frac{\lambda_1 \cdot \left(w_1(t) - e^{T(\lambda_1 - \lambda_2)} w_2(t)\right)}{w_1(t) - \frac{\lambda_1}{\lambda_2} e^{T(\lambda_1 - \lambda_2)} w_2(t)}. \tag{3.30}$$

Finally, using the notation

$$q(t) := w_1(t) - \frac{\lambda_1}{\lambda_2} e^{T(\lambda_1 - \lambda_2)} w_2(t)$$

we can compute $\ell(t)$ for any $t \in [0, T]$ using

$$\ell(t) = -\sigma^2 \cdot \int_t^T f(s) ds = -\sigma^2 \mu \cdot \int_t^T \frac{q'(s)}{q(s)} ds = \sigma^2 \mu \cdot \log \left| \frac{q(T)}{q(t)} \right|. \tag{3.31}$$

Further analysis of the Ricatti Ansatz and the analytic representation of the optimally controlled SDE in Equation (3.29) can be found in the Supplementary Material B.2.

3.5 Numerical results

This section presents all the numerical experiments that were conducted on the quadratic optimal control of Ornstein-Uhlenbeck processes as well as a section dedicated to the sampling process of the solution.

3.5.1 Sampling controlled SDEs

When sampling optimally controlled paths of the underlying Ornstein-Uhlenbeck dynamics, we use an approximation of the x-gradient of the value function resulting from the UFD approximation, i.e.

$$\partial_x u_m^n = \frac{u_{m+1}^n - u_{m-1}^n}{2\Delta x}.$$

To obtain optimal controlled sample paths on a grid $\{t_n\}_{n=0}^N$, we interpolate the x-gradient of the value function and evaluate the gradient at the current processes state, updating the control by

$$\alpha^*(t_n, X(t_n)) = -\frac{\partial_x u(t_n, X(t_n))}{2\mu}, \quad n = 0, \dots, N$$

and receiving the next optimally controlled sample path value by

$$X(t_{n+1}) = X(t_n) + (-\theta X(t_n) + \alpha^*(t_n, X(t_n)))\Delta t + \sigma \cdot \sqrt{\Delta t} \cdot Z, \quad n = 0, \dots, N-1 \tag{3.32}$$

where $Z \sim \mathcal{N}(0,1)$. Using the Ricatti Ansatz (3.28) in Equation (3.32), we can reduce the sampling to

$$X(t_{n+1}) = X(t_n) + \left(-\theta - \frac{f(t_n)}{\mu}\right) \cdot X(t_n) \cdot \Delta t + \sigma \cdot \sqrt{\Delta t} \cdot Z \tag{3.33}$$

where $f(\cdot)$ is given by Equation (3.30) or a linear interpolation of a highly accurate solution $\tilde{f}$ to the Ricatti ODE in Equation (3.27) on a grid $\{\tilde{t}_n\}_{n=0}^{\tilde{N}}$ with $N \ll \tilde{N}$, i.e.

$$f(t) := \tilde{f}(\tilde{t}_n) + \frac{t - \tilde{t}_n}{\tilde{t}_{n+1} - \tilde{t}_n}\left(\tilde{f}(\tilde{t}_{n+1}) - \tilde{f}(\tilde{t}_n)\right), \quad t \in [\tilde{t}_n, \tilde{t}_{n+1}], \quad n = 0, \dots, \tilde{N}-1. \tag{3.34}$$

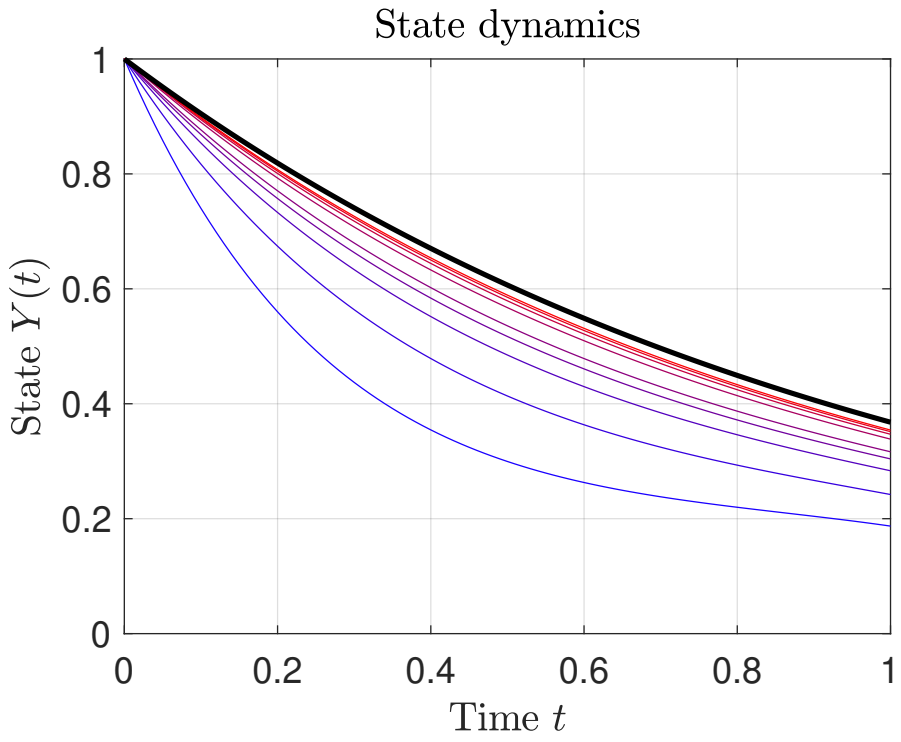
3.5.2 Pontryagin principle

To investigate the solution to the deterministic OC Problem 2, we will note down the solution to the uncontrolled process

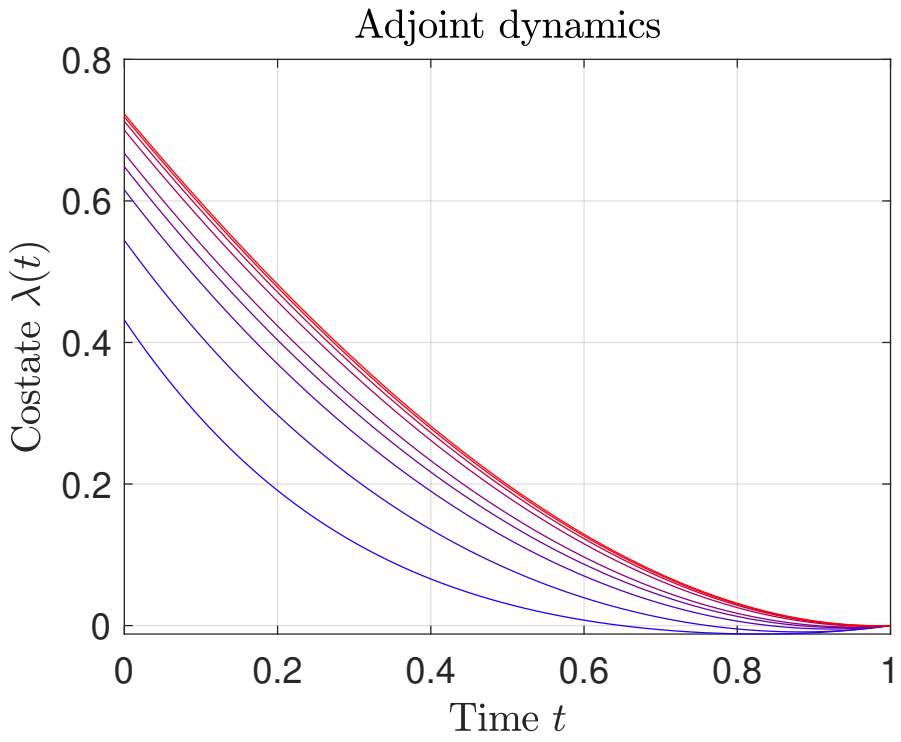
$$\frac{dY(t)}{dt} = -\theta Y(t), \quad Y(0) = x_0, \tag{3.35}$$

i.e. for $\alpha \equiv 0$. In Problem 2, the optimal control will try to *push* down the state trajectories $Y(t)$ to zero to minimize the quadratic running costs. This is also motivated by the zero terminal state costs and the decaying nature of the underlying dynamics. We will investigate the impact of the regularization parameter $\mu > 0$ on the mean-trajectories of the SOC Problem 1. For that, we will solve Problem 2 using

Algorithm 1 for fixed parameters and vary the parameter $\mu > 0$. We will investigate Problem 2 using terminal time $T = 1$, initial condition $x_0 = 1$, drift-parameter $\theta = 1$ and an equidistant time stepping scheme using $N = 100$ steps. We used $\mu \in \{0.1, 0.25, 0.5, 0.75, 1, 2, 3, 4, 5\}$ resulting in 9 runs of Algorithm 1. The parameter $\mu > 0$ corresponds to the costs of controlling the dynamics, i.e. we expect the optimally controlled state path $Y(t)$ for lower values of μ to deviate more from the path of the uncontrolled dynamics and for greater values of μ we expect the optimally controlled dynamics path $Y(t)$ to be close to the uncontrolled path. In Figure 3.1 and 3.2, we displayed the resulting trajectories of the state dynamics $Y(t)$ on the grid $t_k = 1/k$, $k = 0, \ldots, 100$, Figure 3.1a, the corresponding costate $\lambda(t)$ of the adjoint dynamics, 3.1b, the corresponding optimal controls $\alpha^*(t)$, Figure 3.2a, and their corresponding value functions $u(t)$, Figure 3.2b. We choose to plot the results in the same plot with a color shift from red to blue corresponding to the value of μ. For small values of μ we use blue coloring and for greater values of μ we use red coloring. In Figure 3.1a, we indicate the uncontrolled state dynamics solution using a thicker, black plot. We see the expected drift away from the trajectory of the uncontrolled dynamics (in black) when decreasing the value of μ to near zero (color shift from red to blue). In Figure 3.2b, we can see that smaller values for μ result in lower values for the value function. This is supported by the fact that controlling the dynamics is cheaper and thus the impact of the state $Y(t)$ on the value function is greater resulting in lower values for the running costs. In other words, the running costs shift from $Y(t)^2$ to $\mu \cdot \alpha(t)^2$ where the costs are damped because μ is small.

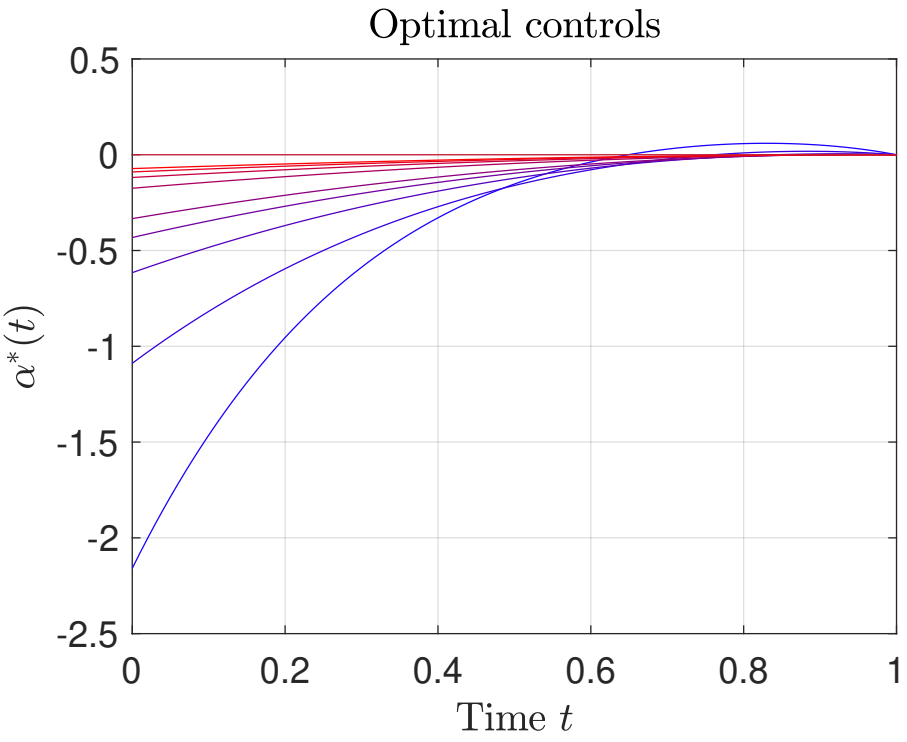


(a) State dynamic trajectories $Y(t)$

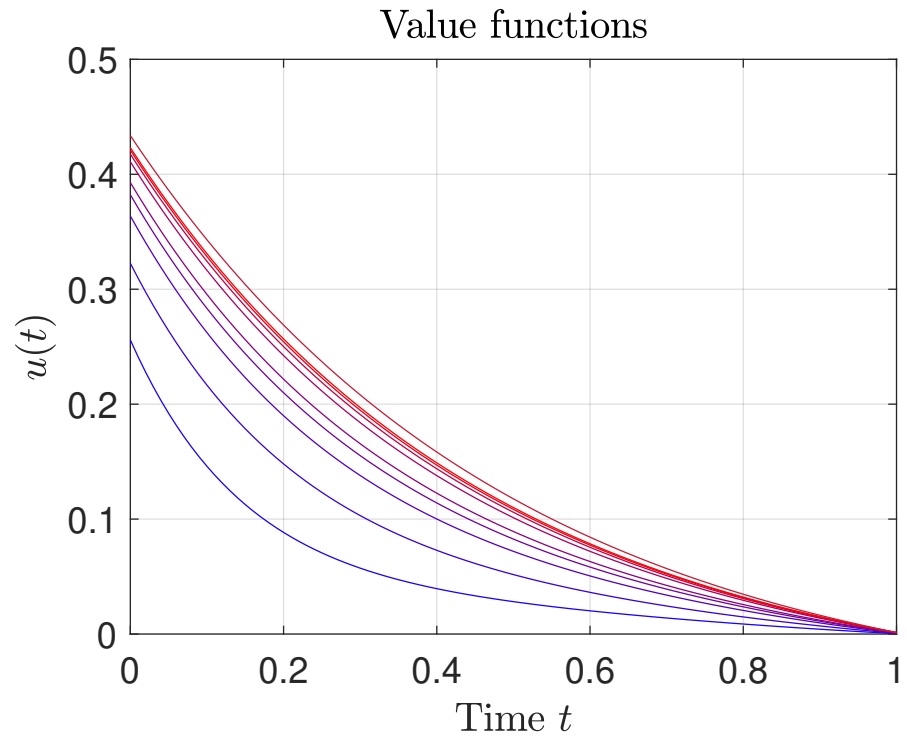


(b) Adjoint state dynamics $\lambda(t)$

Figure 3.1: Results from Algorithm 1 for Problem 2 using fixed parameters and varying the regularization parameter $\mu > 0$.



(a) Optimal controls $\alpha^*(t)$



(b) Value functions $u(t)$

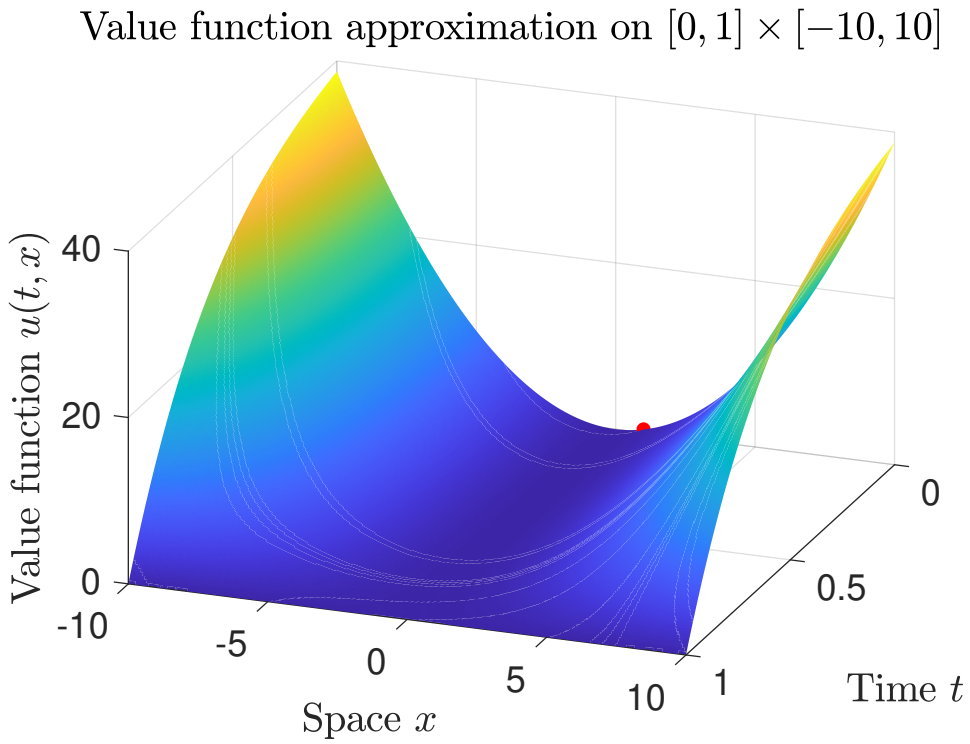
Figure 3.2: Results from Algorithm 1 for Problem 2 using fixed parameters and varying the regularization parameter $\mu > 0$.

3.5.3 HJB equation & Ricatti Ansatz

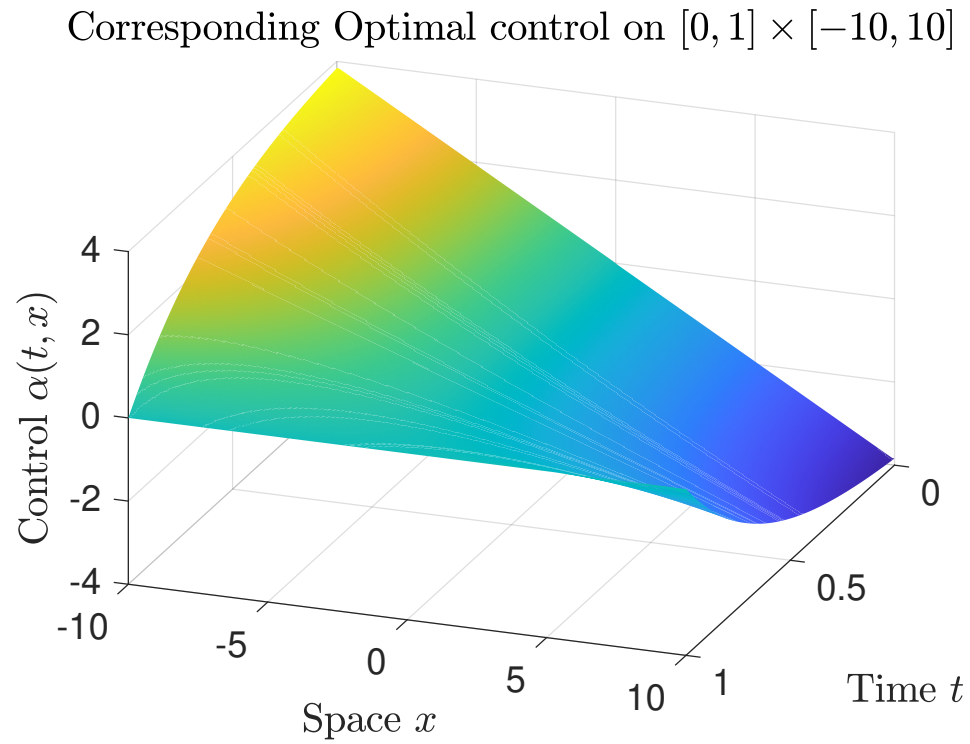
In this section, we will investigate the results for Problem 1 using Algorithm 2 to solve the corresponding HJB equation (3.4) with terminal time boundary condition (3.5). We will adapt the parameters used in the previous Section 3.5.2, i.e. $x_0 = 1$, $\theta = 1$, $T = 1$, and further use $\sigma = 0.1$ and fix $\mu = 1$. Choosing the space restriction is crucial and needs to ensure that the state $X(t)$ for our problem setting will not get out of bound. Since we are using a constant diffusion $\sigma = 0.1$, it is very unlikely for the state $X(t)$ with an initial state $X(0) = 1$ to get out of bound when choosing $[c, d] = [-10, 10]$, i.e. $X(t) \notin [-10, 10]$ for all $t \in [0, 1]$. We will use the estimation of the constant $\tilde{c}_1 \approx 41.87$, see Supplementary Material B.1.1, in combination with the results from Theorem 3.3.1 to approximate the HJB equation on the domain $[0, 1] \times [-10, 10]$. For a prescribed absolute tolerance $\text{TOL} = 10^{-1}$, the resulting time and space discretization steps are $N_1^* = 298$ and $M_1^* = 419$. The resulting plots of the solution to the HJB equation on this domain is displayed in Figure 3.3a. On the right hand side, the corresponding optimal control for all possible time-state pairs $(t, X(t)) \in [0, 1] \times [-10, 10]$ is displayed. The small red dot in the value function approximation indicates the time-state pair at which we can evaluate the value function to get the total costs of the SOC problem at hand, i.e. $u(0, x_0) \approx 0.3999$ in our case. The corresponding total costs of Problem 2 using $\mu = 1$, see Figure 3.2b, is $u(0) \approx 0.3928$. When considering the relation between the value function u and the optimal control α^*, i.e.

$$\alpha^*(t, x) = -\frac{\partial_x u(t, x)}{2}, \quad (t, x) \in [0, 1] \times [-10, 10],$$

we can deduce the quadratic behaviour of the value function since the optimal control is linear in the state, see Figure 3.3b. The next step is to compare the HJB approximation in Figure 3.3 with the value function approximation using the Ricatti Ansatz in Section 3.4. For that, we compute the function evaluations of f in Equation (3.30) and ℓ in Equation (3.31) on the grid discretized by $N_1 = 81$ and $M_1 = 81$. These values are chosen such that the resulting costs of the SOC problem is given by the red dot indicated in Figure 3.4a and we do not need to interpolate the value and thus directly get the value function evaluation is $u(0, x_0) = 0.3832$. Finally, we will compare some resulting sample paths of the optimally controlled SDE using the HJB approximation, see Equation (3.32), and the Ricatti Ansatz using (3.33). We will sample three paths on an equidistant grid t_k, $k = 0, \ldots, 500$ using a fixed seed such that we can directly compare the paths that are controlled by the HJB approximation approach and the Ricatti Ansatz. The samples are displayed in Figure 3.5 with a subplot zooming in at terminal time to show the similarities of the paths. This leads to the conclusion that for the example of quadratic optimal control of Ornstein-Uhlenbeck processes the usage of the Ricatti Ansatz is justified and preferred due to its lower computational complexity, i.e. we do not need to solve a non-linear PDE but can simply evaluate the analytic representation (3.26) using (3.30) and (3.31).

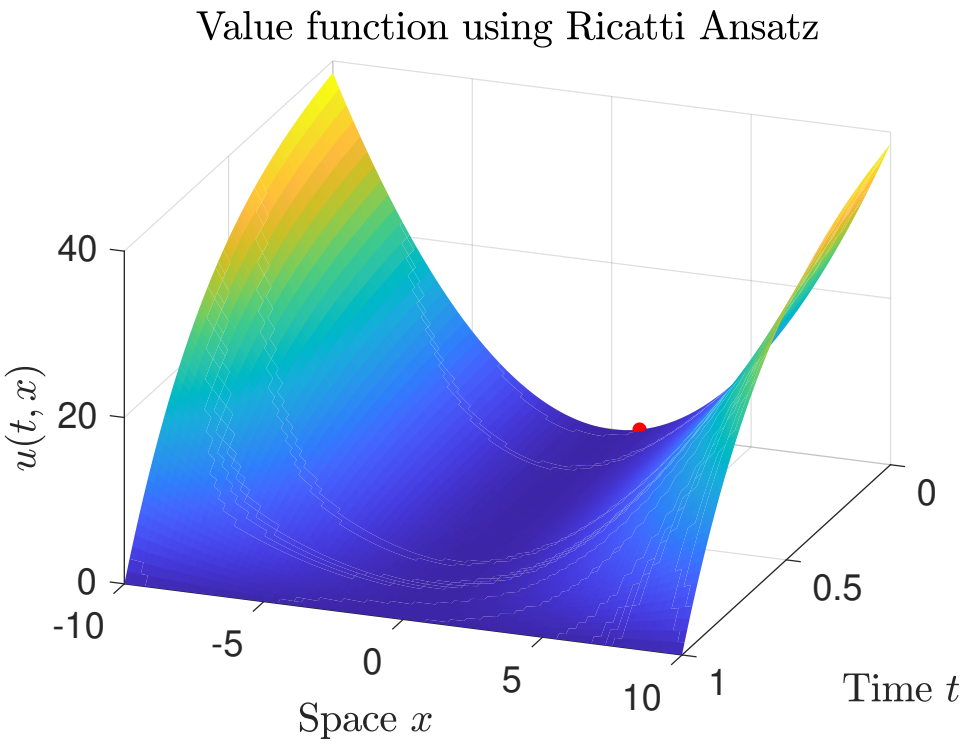


(a) HJB approximation

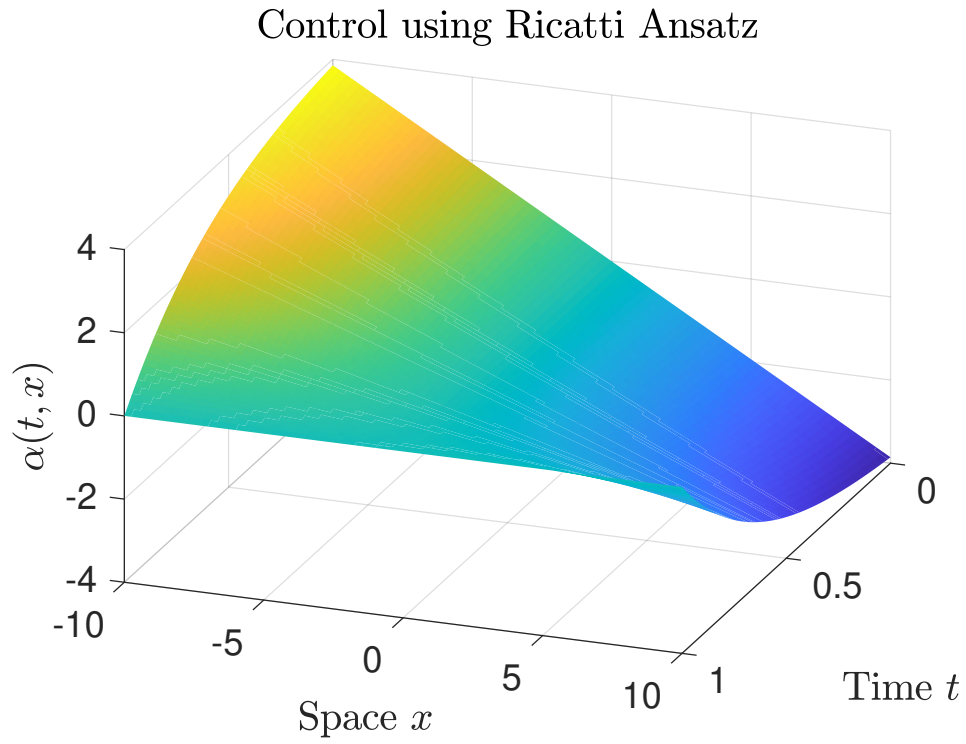


(b) Optimal control approximation

Figure 3.3: Approximate solutions to Problem 1 on the domain $[0, 1] \times [-10, 10]$ for an absolute tolerance TOL $= 0.1$.



(a) Value function



(b) Optimal control approximation

Figure 3.4: Analytic solution to the HJB equation using the Ricatti Ansatz in Equation (3.26) using the representations in Equation (3.30) and (3.31).

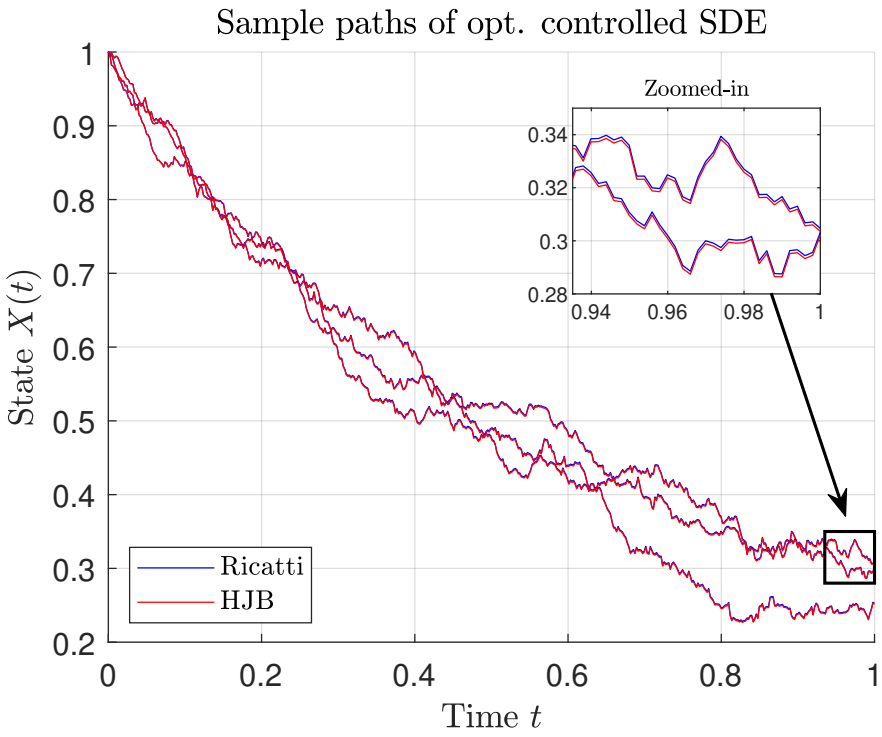


Figure 3.5: Sample paths of the optimally controlled SDE in Problem 1 using the HJB approximation (red) described in Algorithm 2 for sampling as described in (3.32) and using the Ricatti Ansatz (blue) with sampling as described in (3.33).

Chapter 4

Chance Constrained Optimal Control of SDEs

The goal of this chapter is to give a roadmap to solving a class of chance constrained SOC problems theoretically and numerically. In particular, we are solving chance constrained SOC problems with a probability constraint on the terminal state. Previous work has explored the possibility of solving these kind of problems using relaxation techniques, namely Lagrangian relaxation [22, 23, 35]. Applying the dynamic programming relation directly on the chance constrained problem is not possible because the constraint controls the joint distribution of the random variable, rather than each realization [34, 33]. [35] and [23] lay the theoretical groundwork for solving chance constrained SOC problems using dual theory but lack numerical optimization procedures on finding the optimal Lagrange Mutlipliers in the dual formulation. In [35], the numerical experiments are only carried out for fixed penalizations of the constraint, only investigating wether or not the chance constraint is satisfied, disregarding the optimality of the penalization. We aim to close this gap by proposing a numerical optimization method to solve the dual problem. This is done by proposing a subgradient based numerical optimization algorithm in Section 4.4 with a solution methodology and numerical solver to compute a subgradient in each iteration of the SGM in Section 4.2 and 4.2.1. In Section 4.3, we propose a novel heuristic approach to estimate the relative dual gap for chance constrained SOC problems which involves solving an auxiliary chance constrained SOC problem. Additionally, in Section 4.5, we propose a novel approach to finding a suitable initial value for the underlying numerical optimization procedure of Section 4.4. Everything together leads to an algorithm that computes the optimal control of the constrained SOC problem, with a probability constraint in the terminal state, up to a prescribed tolerance $\text{TOL}_{\text{SG}} > 0$. The procedures are tested using two numerical examples. A chance constrained version of Problem 1 discussed in detail in Chapter 3 and an additional example in which the optimal control is not smooth. Finally, an investigation in the cost reduction of using chance constraints as relaxations of almost sure constraints is performed. In the following sections, we will use the shorthand notation

$$X_t \equiv X(t), \quad \alpha_t \equiv \alpha(t, X(t)), \quad W_t \equiv W(t), \quad t \in [0, T].$$

4.1 Problem statement

We are interested in solving the following problem with a single terminal state chance constraint.

Problem 3 (Primal problem)**.** *Consider the following optimization problem*

$$\min_{\alpha\in\mathcal{A}} \mathbb{E}\left[\int_0^T h(t, X_t, \alpha_t)\, dt + g(X_T)\right], \quad s.t. \quad \mathbb{P}\left(G(X_T) > K\right) \leq \delta$$

over the set of admissible deterministic Markovian controls $\mathcal{A}$, that are adapted to the filtration $\mathcal{F}_t$ of the process $X_t \in \mathbb{R}$ for all $t \in [0, T]$ which follows the controlled dynamics

$$\begin{cases} dX_t &= a(t, X_t, \alpha_t)dt + b(t, X_t, \alpha_t)dW_t, \quad 0 < t \leq T, \\ X_0 &= x_0 \in \mathbb{R}, \end{cases}$$

for a given terminal time $T > 0$, violation probability $0 < \delta \ll 1$, threshold $K \in \mathbb{R}$, $G : \mathbb{R} \to \mathbb{R}$, drift $a : [0, T] \times \mathbb{R} \times \mathcal{A} \to \mathbb{R}$, diffusion $b : [0, T] \times \mathbb{R} \times \mathcal{A} \to \mathbb{R}$ and independent 1-dimensional Wiener process W_t.

We call Problem 3 the *primal* problem. The chance constraint prevents direct application of the dynamic programming principle [34]. The constraint controls the underlying joint probability distribution of the random variable, rather than each realization [33]. Thus, we will use Lagrangian Relaxation to relax the constraint and solve the corresponding dual problem by applying a numerical optimization technique to find the optimal LM. The relaxed sub-problem can be solved using the dynamic programming relation and thus by solving a corresponding HJB equation. Relaxing the terminal state probabilitic constraint will result in a SOC problem with pertubed terminal costs in the objective. The probabilistic contraint can be written as an expectation constraint with a discontinuous indicator function, i.e.

$$\mathbb{P}(G(X_T) > K) = \mathbb{E}[\mathbb{1}_{\{G(X_T)>K\}}]. \tag{4.1}$$

This will result in a discontinuous terminal time boundary condition in the HJB equation. This motivates the use of a smoothened version, namely a sigmoid function, of the indicator function in the probability

$$\mathrm{sig}_\gamma : \mathbb{R} \to [0, 1], \quad x \mapsto \frac{1}{1 + \exp\left(\gamma^{-1} \cdot x\right)}. \tag{4.2}$$

Later, we will see that the error induced by using a sigmoid approximation with a smoothening width of $\mathcal{O}(\gamma)$ can be controlled. The *relaxed* problem associated with Problem 3 can be written as follows.

Problem 4 (Relaxed problem)**.** *For a fixed $\lambda \geq 0$, consider the following optimization problem*

$$\min_{\alpha\in\mathcal{A}_{\mathrm{rel}}} \mathbb{E}\left[\int_0^T h(t, X_t^\lambda, \alpha_t)\, dt + g(X_T^\lambda) + \lambda \cdot \Big(\mathrm{sig}_\gamma(G(X_T^\lambda) - K) - \delta\Big)\right]$$

over the set of admissible deterministic Markovian controls $\mathcal{A}_{\text{rel}}$*, that are adapted to the filtration* $\mathcal{F}_t$ *of the process* $X_t^\lambda \in \mathbb{R}$ *for all* $t \in [0,T]$ *which follows the controlled dynamics*

$$\begin{cases} dX_t^\lambda &= a(t, X_t^\lambda, \alpha_t)dt + b(t, X_t^\lambda, \alpha_t)dW_t, \quad 0 < t \leq T, \\ X_0^\lambda &= x_0 \in \mathbb{R}, \end{cases}$$

for a given terminal time $T > 0$*, violation probability* $0 < \delta \ll 1$*, threshold* $K \in \mathbb{R}$*,* $G : \mathbb{R} \to \mathbb{R}$*, drift* $a : [0,T] \times \mathbb{R} \times \mathcal{A}_{\text{rel}} \to \mathbb{R}$*, diffusion* $b : [0,T] \times \mathbb{R} \times \mathcal{A}_{\text{rel}} \to \mathbb{R}$ *and independent* 1*-dimensional Wiener process* W_t*.*

Let us denote the dual function by $d : [0,\infty) \to \mathbb{R}$,

$$\lambda \mapsto \min_{\alpha \in \mathcal{A}_{\text{rel}}} \mathbb{E}\left[\int_0^T h(t, X_t^\lambda, \alpha_t)\, dt + g(X_T^\lambda) + \lambda \cdot \left(\text{sig}_\gamma(G(X_T^\lambda) - K) - \delta \right) \,\middle|\, X_0^\lambda = x_0 \right], \tag{4.3}$$

subject to the controlled dynamics X_t^λ from Problem 4, emphasizing the dependence on the fixed Lagrange Multiplier $\lambda \geq 0$. We can evaluate the dual function for a fixed $\lambda \geq 0$ by solving the HJB equation from Theorem 2.2.2, i.e. we define the value function u^λ for the relaxed Problem 4 as

$$u^\lambda(t,x) := \min_{\alpha \in \mathcal{A}_{\text{rel}}} \mathbb{E}\left[\int_t^T h(s, X_s^\lambda, \alpha_s)\, ds + g(X_T^\lambda) + \lambda \cdot \left(\text{sig}_\gamma(G(X_T^\lambda) - K) - \delta \right) \,\middle|\, X_t^\lambda = x \right],$$

where superscript λ indicates the dependency on the fixed LM $\lambda \geq 0$. u^λ solves the HJB equation for $(t,x) \in [0,T) \times \mathbb{R}$

$$0 = \partial_t u^\lambda + \min_{\alpha \in A} \left[a(t,x,\alpha)\partial_x u^\lambda + \frac{b(t,x,\alpha)^2}{2} \partial_{xx} u^\lambda + h(t,x,\alpha) \right], \tag{4.4}$$

$$u^\lambda(T,x) = g(x) + \lambda \cdot \left(\text{sig}_\gamma(G(x) - K) - \delta \right), \tag{4.5}$$

with optimal control at each time-space pair $(t,x) \in [0,T] \times \mathbb{R}$ denoted by

$$\alpha^\lambda(t,x) := \operatorname*{argmin}_{\alpha \in A} \left[a(t,x,\alpha)\partial_x u^\lambda + \frac{b(t,x,\alpha)^2}{2} \partial_{xx} u^\lambda + h(t,x,\alpha) \right]. \tag{4.6}$$

The *dual* problem associated with Problem 3 is formulated as follows.

Problem 5 (Dual problem)**.** *Find*

$$\lambda^* = \operatorname*{argmax}_{\lambda \geq 0} d(\lambda),$$

where the dual function is given by

$$d(\lambda) = u^\lambda(0, x_0),$$

and u^λ *solves the HJB equation (4.4) with terminal time boundary conditions (4.5).*

We will solve the dual problem 5 numerically by applying numerical optimization methods in order to find the optimal Lagrange Multiplier λ^* up to a prescribed tolerance TOL_{SG}.

4.2 Subgradients via KBE

We will solve the dual Problem 5 numerically using a subgradient method since the objective, namely the dual function, might not be differentiable for all LM $\lambda \geq 0$. This paves the possibility to substitute the standard subgradient method, presented in Section 2.3.2, with a more sophisticated numerical optimization method. Possible class of methods include bundle methods where the update direction is found as the least-norm convex combination of some *bundle* of previous subgradients [51, 40]. The basics for subgradient methods are described in Section 2.3.2. The goal of this section is to derive a representation of a subgradient of the objective in the relaxed problem 4 using the KBE from Theorem 2.1.5. According to [40], the violation in constraint is a subgradient in the fixed LM $\lambda \geq 0$, i.e.

$$\mathbb{E}\left[\operatorname{sig}_\gamma(G(X_T^\lambda) - K) - \delta \;\middle|\; X_0^\lambda = x_0\right] \in \partial d(\lambda).$$

We define the QoI

$$v^\lambda(t,x) := \mathbb{E}\left[\operatorname{sig}_\gamma(G(X_T^\lambda) - K) - \delta \;\middle|\; X_t^\lambda = x\right], \tag{4.7}$$

such that a subgradient in $\lambda \geq 0$ can be computed by evaluating the QoI (4.7) in the initial conditions $(0, x_0)$ of the controlled SDE

$$v^\lambda(0,x_0) \in \partial d(\lambda).$$

We can solve for the QoI using Theorem 2.1.5, i.e. by solving the KBE

$$0 = \partial_t v^\lambda + a(t,x,\alpha^\lambda)\partial_x v^\lambda + \frac{b(t,x,\alpha^\lambda)^2}{2}\partial_{xx}^2 v^\lambda, \quad (t,x) \in [0,T) \times \mathbb{R} \tag{4.8}$$

$$v^\lambda(T,x) = \operatorname{sig}_\gamma(G(x) - K) - \delta. \tag{4.9}$$

4.2.1 Crank Nicolson method

The next step is to provide a numerical scheme to solving the second order linear parabolic PDE (4.8) with terminal time boundary condition (4.9). We choose to use the Crank-Nicolson method. The CN method is a combination of forward and backward Euler approximations. For linear PDEs of the form

$$\partial_t v(t,x) = F\left(t,x,v,\partial_x v,\partial_{xx} v\right), \quad (t,x) \in [0,T] \times [c,d], \tag{4.10}$$

with a linear operator F and arbitrary smooth boundary conditions on the boundary of the domain $[0,T] \times [c,d]$, the CN method uses, for positive integers N_2, M_2, the approximation

$$\frac{v_m^{n+1} - v_m^n}{\Delta t} = \frac{1}{2}\left(F_m^{n+1} + F_m^n\right),$$

where $\Delta t = T/N_2$, $\Delta x = |d-c|/M_2$ and for $n = 0, \ldots, N_2$, $m = 1, \ldots, M_2 - 1$,

$$v_m^n := v(n\Delta t, m\Delta x), \quad F_m^n := F\left(n\Delta t, m\Delta x, v_m^n, \frac{v_{m+1}^n - v_{m-1}^n}{2\Delta x}, \frac{v_{m+1}^n - 2v_m^n + v_{m-1}^n}{(\Delta x)^2}\right).$$

This method results in a linear system of equations with an underlying tridiagonal matrix. This method is unconditionally stable for linear parabolic PDEs, meaning that independent of the time and space discretization chosen, the numerical solution will not blow up [52]. This is in contrast to the UFD scheme proposed in Section 3.3 where the time and space discretization have to satisfy the CFL condition (3.21). If v is the exact solution to the PDE (4.10) and $\tilde{v}$ is the numerical approximation using the CN method, the theoretical rate of convergence is given by

$$\Delta v := \tilde{v}(0, x_0) - v(0, x_0) = \mathcal{O}\left((\Delta t)^2 + (\Delta x)^2\right).$$

This can be looked up in e.g. [53] or in Supplementary Material B.1.2 where this is verified numerically. For the second order partial derivative approximation, we use a central finite difference scheme resulting in a quadratic space accuracy. For the time discretization we are using a centered finite difference but the spatial terms are evaluated at the midpoint in time, which improves the accuracy resulting in a second order accurate method in the time discretization. The approximation of the KBE using the CN method has a computational work of $\mathcal{W} = \mathcal{O}(N_2 \cdot M_2^\Gamma)$ where Γ is the workrate of the solver used when solving the linear system of equations. For tridiagonal matrices $\Gamma = 1$ [53]. The optimal work $\mathcal{W}^*$ to approximate the quantity of interest $v(0, x_0)$ for a given control α^λ using the KBE with a relative tolerance $\text{TOL}_\text{r} > 0$ is described in the following Theorem 4.2.1. The resulting work rates are numerically verified and displayed in Figure 4.1. We aim to approximate values that are approximately around a prescribed violation probability of $0 < \delta \ll 1$.

Theorem 4.2.1 (Optimal Work Crank Nicolson)**.** *Consider the following optimization problem*

$$\begin{aligned} &\min_{N_2, M_2} \ \mathcal{W} = \min_{N_2, M_2} \ N_2 \cdot M_2^\Gamma, \\ s.t. \quad &\frac{c_2}{N_2} + \frac{\bar{c}_2}{M_2} \approx \text{TOL}_\text{r} \cdot \delta. \end{aligned}$$

Then,

$$\mathcal{W}^* = \mathcal{O}\left((\text{TOL}_\text{r} \cdot \delta)^{-\frac{1+\Gamma}{2}}\right)$$

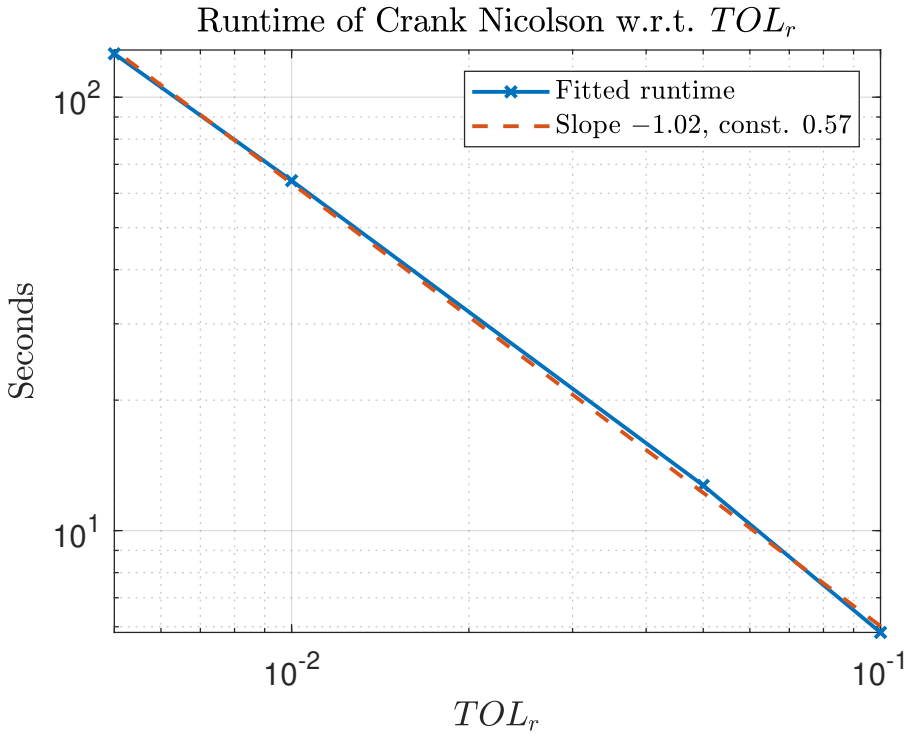
and

$$N_2^* = \sqrt{c_2 \cdot (\Gamma + 1)}\,(\text{TOL}_\text{r} \cdot \delta)^{-\frac{1}{2}}, \quad M_2^* = \sqrt{\bar{c}_2 \cdot (1 + \Gamma^{-1})}\,(\text{TOL}_\text{r} \cdot \delta)^{-\frac{1}{2}}.$$

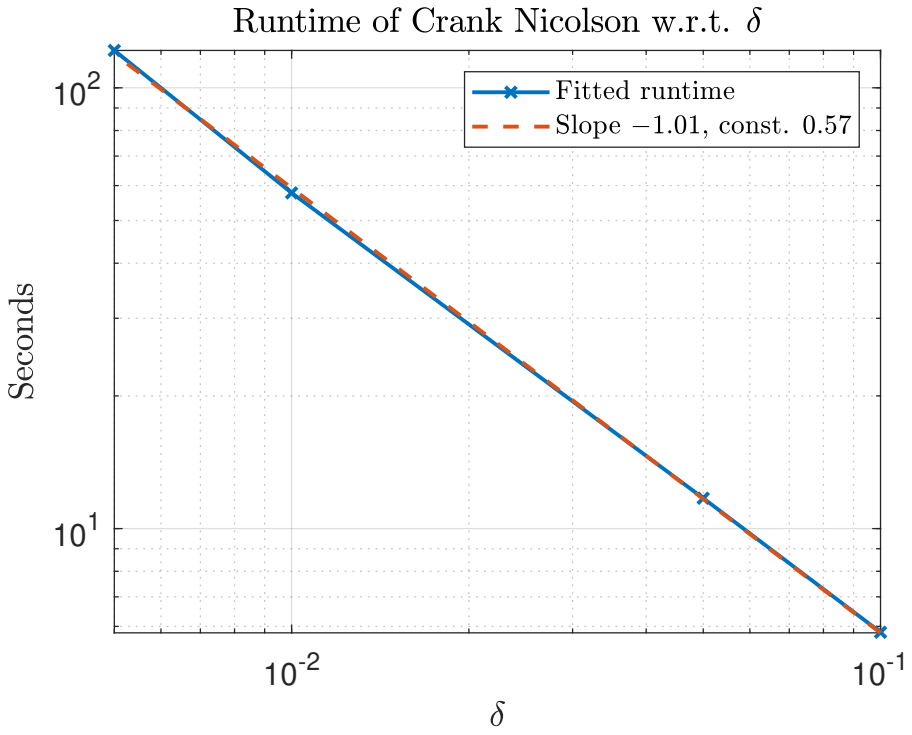
Proof. The proof is done similar to Proof A.4 using Lagrangian relaxation. □

4.3 Relative Dual gap estimation

The convergence of the SGM is dependent on the two the assumptions (2.33) and (2.34) introduced in the theoretical analysis of the SGM in Section 2.3.2. In practice, an estimation of the constants $R, S > 0$ in (2.33) and (2.34), respectively, is not trivial and needs sufficient information on the problem at hand. That is why we are interested in a reliable stopping criterion in each iteration independent of these constants. In computational optimization it is common to approximate the dual gap in each iteration



(a) Runtime w.r.t. $\mathrm{TOL_r}$



(b) Runtime w.r.t. δ

Figure 4.1: Numerical verification of the asymptotical work rates of Theorem 4.2.1. The result implies that the workrate of our solver is $\Gamma = 1$ which corresponds to inverting matrices of size n with complexity $\mathcal{O}(n)$ which is possible for tridiagonal matrices.

for a fixed arbitrary LM $\lambda \geq 0$ using a feasible but suboptimal solution to the primal problem. This is also motivated by the fact that an estimation of the exact primal solution p^*, as described in Equation (2.28), is in general not possible. We propose an auxiliary problem in which we construct a suboptimal primal feasible solution η^λ to Problem 3 and its resulting primal costs p^λ corresponding to the current dual solution α^λ and its dual costs $u^\lambda(0, x_0)$ of Problem 4. We aim to find a primal feasible sub-optimal control η^κ such that

$$\mathbb{E}\left[\operatorname{sig}_\gamma(G(X_T^\kappa) - K) \,\middle|\, X_0^\kappa = x_0\right] \approx \delta,$$

where X_t^κ is the optimally controlled SDE governed by the following SOC problem.

Problem 6. *Given a control α^λ, consider the following OC problem*

$$\min_{\eta \in \mathcal{A}_{\text{rel}}} \mathbb{E}\left[\int_0^T \left(\alpha^\lambda(s, X_s^\kappa) - \eta(s, X_s^\kappa)\right)^2 ds\right] + \kappa \cdot \mathbb{E}\left[\operatorname{sig}_\gamma(G(X_T^\kappa) - K)\right]$$

over the set of admissible deterministic Markovian controls $\mathcal{A}_{\text{rel}}$, that are adapted to the filtration $\mathcal{F}_t^\kappa$ of the process $X_t^\kappa \in \mathbb{R}$ for all $t \in [0, T]$ which follow the controlled dynamics

$$\begin{cases} dX_t^\kappa &= a(t, X_t^\kappa, \eta_t)dt + b(t, X_t^\kappa, \eta_t)dW_t, \quad 0 < t \leq T, \\ X_0^\kappa &= x_0 \in \mathbb{R} \end{cases}$$

for a given terminal time $T > 0$, threshold $K \in \mathbb{R}$, penalization parameter $\kappa \geq 0$, $G : \mathbb{R} \to \mathbb{R}$, drift $a : [0, T] \times \mathbb{R} \times \mathcal{A}_{\text{rel}} \to \mathbb{R}$, diffusion $b : [0, T] \times \mathbb{R} \times \mathcal{A}_{\text{rel}} \to \mathbb{R}$ and independent 1-dimensional Wiener process W_t.

This problem can be interpreted as a regularization problem in which we would like to find the closest control η to the current LM's optimal control α^λ such that η is primal feasible, i.e. the chance constraint is satisfied. We can solve Problem 6 using previous results. Namely, we define the value function for an arbitrary initial condition $(t, x) \in [0, T] \times \mathbb{R}$ of the controlled SDE corresponding to Problem 6 as

$$u^\kappa(t, x) := \min_{\eta \in \mathcal{A}_{\text{rel}}} \mathbb{E}\left[\int_t^T \left(\alpha^\lambda(s, X_s^\kappa) - \eta(s, X_s^\kappa)\right)^2 ds + \kappa \cdot \operatorname{sig}_\gamma(G(X_T^\kappa) - K) \,\middle|\, X_t^\kappa = x\right].$$

u^κ solves the following HJB equation

$$0 = \partial_t u^\kappa + \min_{\eta \in A}\left[a(t, x, \eta)\partial_x u^\kappa + \frac{b(t, x, \eta)^2}{2}\partial_{xx} u^\kappa + \left(\alpha^\lambda(t, x) - \eta\right)^2\right], \tag{4.11}$$

$$u^\kappa(T, x) = \kappa \cdot \operatorname{sig}_\gamma(G(x) - K), \tag{4.12}$$

for all $(t, x) \in [0, T) \times \mathbb{R}$. For a given solution

$$\eta^\kappa(t, x) = \operatorname*{argmin}_{\eta \in \mathcal{A}_{\text{rel}}}\left[a(t, x, \eta(t, x))\partial_x v^\kappa + \frac{b(t, x, \eta(t, x))^2}{2}\partial_{xx} v^\kappa + \left(\alpha^\lambda(t, x) - \eta(t, x)\right)^2\right], \tag{4.13}$$

for $(t, x) \in [0, T) \times \mathbb{R}$, of Problem 6, we define

$$v^\kappa(t, x) := \mathbb{E}\left[\operatorname{sig}_\gamma(G(X_T^\kappa) - K) \,\middle|\, X_t^\kappa = x\right]$$

such that the QoI of this section is given by the evaluation at the initial conditions $(0, x_0)$, i.e.

$$v^\kappa(0, x_0) = \mathbb{E}\left[\operatorname{sig}_\gamma(G(X_T^\kappa) - K) \;\middle|\; X_0^\kappa = x_0\right].$$

We can find $v^\kappa(t, x)$ for an arbitrary initial condition $(t, x) \in [0, T] \times \mathbb{R}$ of the controlled SDE corresponding to Problem 6 using the KBE

$$0 = \partial_t v^\kappa + a(t, x, \eta^\kappa)\partial_x v^\kappa + \frac{b(t, x, \eta^\kappa)^2}{2}\partial_{xx} v^\kappa, \tag{4.14}$$

$$v^\kappa(T, x) = \operatorname{sig}_\gamma(G(x) - K), \tag{4.15}$$

for $(t, x) \in [0, T) \times \mathbb{R}$. We are now able to define the chance constraint function as

$$\mathrm{CC} : [0, \infty) \to [0, 1], \quad \kappa \mapsto \mathbb{E}\left[\operatorname{sig}_\gamma(G(X_T^\kappa) - K) \mid X_0^\kappa = x_0\right] = v^\kappa(0, x_0). \tag{4.16}$$

The chance constraint function $\mathrm{CC}(\kappa)$ is a monotonically decreasing function in the penalization parameter $\kappa \in [0, \infty)$. The construction of the primal feasible sub-optimal control η^κ is motivated by the following key idea: We seek a minimal terminal cost in the objective of Problem 6. The terminal costs is bounded by a constant, more specific

$$\kappa \cdot \mathrm{CC}(\kappa) = \mathcal{O}(1)$$

with respect to the violation probability $0 < \delta \ll 1$. We aim the chance constraint to be approximately around the violation probability, i.e. the chance constraint function is of order

$$\mathrm{CC}(\kappa) = \mathcal{O}(\delta)$$

such that we can deduce that the penalizatin parameter is of order

$$\kappa = \mathcal{O}(\delta^{-1}).$$

Due to the fact that estimating the relative dual gap is a subroutine for the upcoming subgradient method, the goal is to approximate the dual gap without much computational overload. This *order splitting* enables us to find suitable parameters κ_i, $i = 1, 2, 3$, to use a quadratic regression model to approximate the chance constraint function (4.16) around the point of interest κ^* for which $\mathrm{CC}(\kappa^*) \approx \delta$. We chose a quadratic interpolation because of its balance between simplicity and accuracy. In return, solving Problem 6 using the parameter κ^* yields an optimal control η^{κ^*} which we can use to construct the desired primal costs approximation p^λ and thus the estimated numerical relative dual gap

$$d^\lambda := \frac{p^\lambda - d(\lambda)}{d(\lambda)}, \tag{4.17}$$

where d is the dual function defined in (4.3). In detail, the novel procedure can be described as follows: For a given LM $\lambda \geq 0$ and corresponding optimal control α^λ given by Problem 4, we want to solve Problem 6 for three well chosen points κ_i, $i = 1, 2, 3$, to get a unique quadratic interpolation polynomial. The first parameter κ_1 should yield an optimal control η^{κ_1} such that the optimally controlled process $X_t^{\kappa_1}$ satisfies

$$\mathbb{E}\left[\operatorname{sig}_\gamma(G(X_T^{\kappa_1}) - K) \mid X_0^{\kappa_1} = x_0\right] > \delta \tag{4.18}$$

and the third parameter κ_3 should yield

$$\mathbb{E}\left[\mathrm{sig}_\gamma(G(X_T^{\kappa_3}) - K) \mid X_0^{\kappa_3} = x_0\right] < \delta. \tag{4.19}$$

The first parameter can be chosen as $\kappa_1 = \lambda$. For the third point κ_3, we use the *order splitting* described above, i.e.

$$\kappa_3 = \kappa_1 + \tilde{c} \cdot \delta^{-1} \tag{4.20}$$

for some constant $\tilde{c} > 0$ such that

$$\mathbb{E}\left[\mathrm{sig}_\gamma(G(X_T^{\kappa_3}) - K) \mid X_0^{\kappa_3} = x_0\right] < \delta \tag{4.21}$$

holds and η^{κ_3} is primal feasible. The constant $\tilde{c} > 0$ is application dependent. In case an estimation of the constant $\tilde{c}$ is too imprecise and does not yield a primal feasible control η^{κ_3}, i.e. not satisfying inequality (4.21), we have to choose a larger constant. The second parameter κ_2 is chosen as the intermediate point

$$\kappa_2 = \frac{1}{2}\left(\kappa_1 + \kappa_3\right).$$

When quadratically interpolating the three solution points

$$(\kappa_1, \mathrm{CC}(\kappa_1)), (\kappa_2, \mathrm{CC}(\kappa_2)), (\kappa_3, \mathrm{CC}(\kappa_3)) \tag{4.22}$$

with a quadratic (Lagrange) polynomial

$$p_2(\kappa) := \sum_{i=1}^{3} \left(\prod_{\substack{1 \le j \le 3 \\ j \neq i}} \frac{\kappa - \kappa_j}{\kappa_i - \kappa_j} \right) \mathrm{CC}(\kappa_i), \tag{4.23}$$

we can solve the root-finding problem

$$p_2(\kappa^*) - \delta = 0 \tag{4.24}$$

for an approximation of a penalization parameter κ^* such that

$$\mathrm{CC}(\kappa^*) \approx \delta.$$

The resulting control η^{κ^*} to Problem 6, using κ^*, can be used to estimate the relative dual gap defined in (4.17). We can compute the primal costs p^λ using the primal feasible control η^{κ^*}. This is done by solving for the function w defined as

$$w(t,x) := \mathbb{E}\left[\int_t^T h(t, X_s^{\kappa^*}, \eta^{\kappa^*}) ds + g(X_T^{\kappa^*}) \;\middle|\; X_t^{\kappa^*} = x \right]$$

subject to the controlled dynamics of the primal problem 3 with arbitrary initial conditions $(t,x) \in [0,T] \times \mathbb{R}$

$$\begin{aligned} dX_s^{\kappa^*} &= a(s, X_s^{\kappa^*}, \eta_s^{\kappa^*}) ds + b(s, X_s^{\kappa^*}, \eta_s^{\kappa^*}) dW_s, \quad t < s \le T, \\ X_t^{\kappa^*} &= x, \end{aligned}$$

using the PDE

$$0 = \partial_t w + a(t,x,\eta^{\kappa^*})\partial_x w + \frac{b(t,x,\eta^{\kappa^*})^2}{2}\partial_{xx} w + h(t,x,\eta^{\kappa^*}), \tag{4.25}$$

$$w(T,x) = g(x), \tag{4.26}$$

defined on the domain $(t,x) \in [0,T) \times \mathbb{R}$, such that $p^\lambda := w(0,x_0)$. We can then estimate the relative dual gap (4.17) for a fixed LM $\lambda \geq 0$ by computing

$$d^\lambda = \frac{p^\lambda - d(\lambda)}{d(\lambda)} = \frac{w(0,x_0) - u^\lambda(0,x_0)}{u^\lambda(0,x_0)}.$$

The procedure is summarized in Algorithm 3.

Algorithm 3 Relative Dual Gap Estimation

Require: Lagrange Multiplier $\lambda \geq 0$, optimal control α^λ, dual costs $d(\lambda) = u^\lambda(0,x_0)$ and application-dependent constant $\tilde{c} > 0$.
Set $\kappa_1 = \lambda$, $\kappa_3 = \kappa_1 + \tilde{c} \cdot \delta^{-1}$, $\kappa_2 = 0.5 \cdot (\kappa_1 + \kappa_3)$.
for $i = 1,2,3$ **do**
Solve Problem 6 for fixed $\kappa = \kappa_i$ using (4.11), (4.12).
Compute $\mathrm{CC}(\kappa_i)$ using (4.14), (4.15) with η^{κ_i} defined in (4.13).
end for
Construct the quadratic (Lagrange) polynomial p_2 defined in (4.23).
Solve the root-finding problem $p_2(\kappa^*) - \delta = 0$.
Solve Problem 6 using $\kappa = \kappa^*$ to get the primal feasible η^{κ^*}.
Compute $p^\lambda = w(0,x_0)$ by solving (4.25), (4.26).
Compute the relative dual gap d^λ defined in (4.17).
return d^λ.

4.4 Subgradient method

We can use the results from previous sections to write down a SGM for chance constrained optimal control of SDEs. For given step sizes $\{t_k\}$, $k = 0,1,\ldots$, and an initial LM $\lambda_0 \geq 0$, we can write the SGM's update rule, Equation (2.31), for the dual *maximization*, Problem 5, as

$$\lambda_{k+1} = \lambda_k + t_k \cdot v^{\lambda_k}(0,x_0), \quad k = 0,1,2\ldots$$

Due to the inequality in (2.37) in the convergence analysis of the SGM from Section 2.3.2, we are interested in a suitable choice of step-sizes as well as a suitable initial value $\lambda_0 \geq 0$ for the SGM to improve the convergence of the method. For a single (scalar) chance constraint, we propose a bisection-based method of finding a suitable initial value $\lambda_0 \geq 0$ in Section 4.5. Section 4.3 introduced a stopping criterion for the SGM based on an approximation of the relative dual gap in a fixed LM $\lambda \geq 0$. The fact that the subgradient method is based on a subgradient in the current iteration prevents us from solely taking the norm of the subgradient as a stopping criterion. A procedure using the results of the previous sections is written down in Algorithm 4. In this algorithm, we aim to solve the dual problem up to a prescribed relative tolerance $\mathrm{TOL}_{\mathrm{SG}} > 0$ for the relative dual gap esimation introduced in Section 4.3. Note that this is due to the fact that estimating the constants R and S in the assumptions for the convergence of

subgradient methods is, in general, not always possible. For the relative dual gap, defined in (4.17), we will use the shorthand notation

$$d_k := d^{\lambda_k} = \frac{w(0,x_0) - u^{\lambda_k}(0,x_0)}{u^{\lambda_k}}$$

in the current iteration $k \in \mathbb{N}$ with corresponding LM $\lambda_k \geq 0$.

Algorithm 4 SGM for Problem 5

Require: Tolerance $\text{TOL}_{\text{SG}} > 0$, step size choice t_k, initial LM λ_0.
Set $k = 0$, $d_0 > \text{TOL}_{\text{SG}}$.
while $d_k > \text{TOL}_{\text{SG}}$ **do**
Compute u^{λ_k} using (4.4), (4.5) and corresponding control α^{λ_k} (4.6). ▷ UFD scheme (3.18)
Compute the relative dual gap d_k (4.17). ▷ Algorithm 3
Compute a subgradient $v^{\lambda_k}(0,x_0)$ using (4.8), (4.9). ▷ Crank-Nicolson, Section 4.2.1
Update $\lambda_{k+1} = \lambda_k + t_k \cdot v^{\lambda_k}(0,x_0)$. ▷ Subgradient Update
Set $k = k + 1$.
end while
Set $\lambda^* = \lambda_{k+1}$ and compute α^* using (4.4), (4.5) and (4.6).
return Optimal control α^* solving the *dual* Problem 5 up to a relative dual gap of $\text{TOL}_{\text{SG}} > 0$.

4.5 Finding a suitable λ_0

The convergence results of the subgradient method in Section 2.3.2 is the outcome of two main assumptions:

1. The distance of the initial point λ_0 to the optimal point λ^* is bounded by a constant $R > 0$, i.e.

$$\|\lambda_0 - \lambda^*\|_2 \leq R.$$

2. For all iterations k, the subgradients are bounded by some constant $S > 0$, i.e.

$$\|v^{\lambda}(0,x_0)\|_2 \leq S.$$

In our 1 dimensional problem setting and due to the concave nature of the dual function, the second assumption is satisfied and the bound S is given by the subgradient norm of a subgradient in the initial Lagrange Multiplier λ_0, i.e. for all iterations $k = 1, 2, \ldots$

$$\|v^{\lambda_k}(0,x_0)\|_2 \leq \|v^{\lambda_0}(0,x_0)\|_2 =: S.$$

These assumptions result in the main inequality (2.37) for all convergence results from Corollary 2.3.1. In this section, we will propose a bisection-based method to find a suitable initial value $\lambda_0 \geq 0$ for single (scalar) chance constraints as in Problem 3. The bisection method is a reliable and robust method commonly used to find initial values for methods that are not as robust [53]. This method will allow us to control the constant R in the first assumption. Consider the following problem.

Problem 7. *Consider the following optimization problem*

$$\min_{\alpha \in \mathcal{A}_{\text{rel}}} \mathbb{E}\left[\int_0^T h(t, X_t^\beta, \alpha_t)\, dt + g(X_T^\beta) + \beta \cdot \text{sig}_\gamma(G(X_T^\beta) - K)\right]$$

over the set of admissible deterministic Markovian controls $\mathcal{A}_{\text{rel}}$*, that are adapted to the filtration* $\mathcal{F}_t^\beta$ *of the process* $X_t^\beta \in \mathbb{R}$ *for all* $t \in [0,T]$ *which follows the controlled dynamics*

$$\begin{cases} dX_t^\beta &= a(t, X_t^\beta, \alpha_t)ds + b(t, X_t^\beta, \alpha_t)dW_t, \quad 0 < t \leq T, \\ X_0^\beta &= x_0 \in \mathbb{R}, \end{cases}$$

for a given terminal time $T > 0$*, threshold* $K \in \mathbb{R}$*,* $G : \mathbb{R} \to \mathbb{R}$*, drift* $a : [0,T] \times \mathbb{R} \times \mathcal{A}_{\text{rel}} \to \mathbb{R}$*, diffusion* $b : [0,T] \times \mathbb{R} \times \mathcal{A}_{\text{rel}} \to \mathbb{R}$ *and independent* 1*-dimensional Wiener process* W_t*.*

Similar to solving the relaxed Problem 4, we can solve Problem 7 by defining the value function

$$u^\beta(t,x) := \min_{\alpha \in \mathcal{A}_{\text{rel}}} \left[\int_t^T h(s, X_s^\beta, \alpha_s)ds + g(X_T^\beta) + \beta \cdot \text{sig}_\gamma(G(X_T^\beta) - K) \,\middle|\, X_t^\beta = x \right],$$

which is the solution to the HJB equation with pertubed terminal time boundary condition

$$\begin{aligned} 0 &= \partial_t u^\beta + \min_{\alpha \in A} \left[a(t,x,\alpha)\partial_x u^\beta + \frac{b(t,x,\alpha)^2}{2} \partial_{xx} u^\beta + h(t,x,\alpha) \right], \\ u^\beta(T,x) &= g(x) + \beta \cdot \text{sig}_\gamma(G(x) - K), \end{aligned}$$

for all $(t,x) \in [0,T) \times \mathbb{R}$. We denote the optimal control by

$$\alpha^\beta = \operatorname*{argmin}_{\alpha \in \mathcal{A}_{\text{rel}}} \left[a(t,x,\alpha)\partial_x u^\beta + \frac{b(t,x,\alpha)^2}{2} \partial_{xx} u^\beta + h(t,x,\alpha) \right].$$

We can further define

$$v^\beta(t,x) := \mathbb{E}\left[\text{sig}_\gamma(G(X_T^\beta) - K) \,\middle|\, X_t^\beta = x \right], \quad (t,x) \in [0,T] \times \mathbb{R}$$

and know that v^β is the solution to the KBE

$$0 = \partial_t v^\beta + a(t,x,\alpha^\beta)\partial_x v^\beta + \frac{b(t,x,\alpha^\beta)^2}{2} \partial_{xx} v^\beta, \tag{4.27}$$

$$v^\beta(T,x) = \text{sig}_\gamma(G(x) - K), \tag{4.28}$$

on the domain $(t,x) \in [0,T) \times \mathbb{R}$. We are interested in the QoI

$$v^\beta(0,x_0) = \mathbb{E}\left[\text{sig}_\gamma(G(X_T^\beta) - K) \,\middle|\, X_0^\beta = x_0 \right].$$

Similar to the derivation in Section 4.3, we define a chance constraint function

$$\text{CC}_{\text{bisec}} : [0,\infty) \to [0,1], \beta \mapsto v^\beta(0,x_0). \tag{4.29}$$

The proposal is to use the bisection method on the function $\text{CC}_{\text{bisec}}(\cdot) - \delta$. In each iteration $k \in \mathbb{N}$, the bisection method consideres a compact subset $[\beta_k, \bar{\beta}_k] \subseteq [0,\infty)$ of the domain of the function $\text{CC}_{\text{bisec}}(\cdot) - \delta$ with different function signs, i.e.

$$(\text{CC}_{\text{bisec}}(\beta_k) - \delta) \cdot (\text{CC}_{\text{bisec}}(\bar{\beta}_k) - \delta) < 0,$$

and computes its intermediate value

$$\tilde{\beta}_k = \frac{1}{2}\left(\beta_k + \bar{\beta}_k\right).$$

Depending on the function evaluation at the intermediate point $\tilde{\beta}_k$, the next iteration considers the compact interval $[\beta_k, \tilde{\beta}_k]$ or $[\tilde{\beta}_k, \bar{\beta}_k]$. More precisely, the bisection method starts with a pair of values $(\beta_0, \bar{\beta}_0) \in [0,\infty) \times [0,\infty)$ such that

$$\mathrm{CC}_{\mathrm{bisec}}(\beta_0) - \delta > 0 > \mathrm{CC}_{\mathrm{bisec}}(\bar{\beta}_0) - \delta.$$

For iterations $k = 1, \ldots,$ calculate

1. the midpoint $\tilde{\beta}_k = (\beta_{k-1} + \bar{\beta}_{k-1})/2$
2. the function evaluation at the midpoint $\mathrm{CC}_{\mathrm{bisec}}(\tilde{\beta}_k) - \delta$
3. check if $|\tilde{\beta}_k - \beta_{k-1}|$ is sufficiently small, return $\tilde{\beta}_k$ and stop the method or
4. check the sign of $\mathrm{CC}_{\mathrm{bisec}}(\tilde{\beta}_k) - \delta$: If
   - $\mathrm{CC}_{\mathrm{bisec}}(\tilde{\beta}_k) - \delta < 0$: Set $\beta_k = \beta_{k-1}, \bar{\beta}_k = \tilde{\beta}_k$
   - $\mathrm{CC}_{\mathrm{bisec}}(\tilde{\beta}_k) - \delta > 0$: Set $\beta_k = \tilde{\beta}_k, \bar{\beta}_k = \beta_{k-1}$

   and continue with $(\beta_k, \bar{\beta}_k)$.

The goal is to find a $\tilde{\beta}_k \in [0,\infty)$, $k \in \mathbb{N}$, such that we can bound the distance between optimal penalization parameter β^* and penalization parameter $\tilde{\beta}_k$ after $k \in \mathbb{N}$ iterations by an absolute tolerance $\mathrm{TOL}_{\mathrm{b}} > 0$, i.e.

$$|\beta^* - \tilde{\beta}_k| \overset{!}{<} \mathrm{TOL}_{\mathrm{b}}. \tag{4.30}$$

If the function $\mathrm{CC}_{\mathrm{bisec}}(\cdot) - \delta$ is continuous, the method converges linearly and the absolute error is halved in each step [53]. This results in an error bound and thus a result on the iterations $k \in \mathbb{N}$ needed to satisfy (4.30), i.e.

$$|\beta^* - \tilde{\beta}_k| \leq \frac{|\bar{\beta}_0 - \beta_0|}{2^k} \overset{!}{<} \mathrm{TOL}_{\mathrm{b}} \implies k > \frac{1}{\log(2)} \cdot \log\left(\frac{|\beta_0 - \bar{\beta}_0|}{\mathrm{TOL}_{\mathrm{b}}}\right). \tag{4.31}$$

Thus, the bound in (4.30) is satisfied after at most

$$k^* := \left\lceil \frac{1}{\log(2)} \cdot \log\left(\frac{|\beta_0 - \bar{\beta}_0|}{\mathrm{TOL}_{\mathrm{b}}}\right) \right\rceil \tag{4.32}$$

iterations. When using the resulting parameter as an initial value $\lambda_0 = \tilde{\beta}_{k^*}$ for the SGM in Algorithm 4, we can reduce the value of the constant R in the first assumption (2.33) and thus decrease the runtime of the SGM. The choice of initial interval

$$[\beta_0, \bar{\beta}_0] \subseteq [0,\infty)$$

is also crucial. For the left bound of the interval, we can choose $\beta_0 = 0$, which corresponds to the unconstrained problem with $\mathrm{CC}_{\mathrm{bisec}}(0) - \delta > 0$, otherwise the unconstrained problem already satisfies the chance constraint and we do not need to relax the problem at

all. For the second value $\bar{\beta}_0$, we can use the *order splitting* introduced in Section 4.3. In particular,

$$\beta = \mathcal{O}(\delta^{-1})$$

implies that choosing $\bar{\beta}_0 = \mathcal{O}(\delta^{-2})$ results in $\mathrm{CC}_{\mathrm{bisec}}(\bar{\beta}_0) - \delta < 0$.

Algorithm 5 Bisection method for chance constrained SOC

Require: Tolerance $\mathrm{TOL}_{\mathrm{b}} > 0$, initial values $\beta_0 = 0$, $\bar{\beta}_0 = \mathcal{O}(\delta^{-2})$ with $\mathrm{CC}_{\mathrm{bisec}}(0) > \delta$ and $\mathrm{CC}_{\mathrm{bisec}}(\bar{\beta}_0) < \delta$.
  Compute k^* using (4.32).
  for $k = 1, \dots, k^*$ **do**
    Compute $\tilde{\beta}_k = (\beta_{k-1} + \bar{\beta}_{k-1})/2$, $\mathrm{CC}_{\mathrm{bisec}}(\tilde{\beta}_k) = v^{\beta_k}(0, x_0)$ using (4.27), (4.28).
    if $\mathrm{CC}_{\mathrm{bisec}}(\tilde{\beta}_k) - \delta < 0$ **then**
      Set $\beta_k = \beta_{k-1}, \bar{\beta}_k = \tilde{\beta}_k$.
    else
      Set $\beta_k = \tilde{\beta}_k, \bar{\beta}_k = \beta_{k-1}$.
    end if
return $\lambda_0 = \beta_{k^*}$.

4.6 Error and Work Analysis

In the following sections, we aim to deepen the numerical understanding of solving chance constrained SOC problems using the procedures presented. Previous work, e.g. [22] or [35], lacks rigorous numerical analysis of the procedures presented. In that regard, we will discuss the different error sources arising in the numerical implementation of the theoretical results. We will have a special focus on the smoothing error induced by using a sigmoid approximation for the indicator function in the chance constraint in Section 4.6.1. Furthermore, in Section 4.6.3, we will investigate the asymptotic computational work of the algorithms presented. This enables us to solve the problems at hand for prescribed tolerances with minimal computational work. For the reference solution, we need some notation. For a fixed LM $\lambda \geq 0$, consider the value function

$$u(t,x) := \min_{\alpha \in \mathcal{A}_{\mathrm{rel}}} \mathbb{E}\left[\int_t^T h(s, X_s, \alpha_s) ds + g(X_T) + \lambda \left(\mathbb{1}_{\{G(X_T) > K\}} - \delta \right) \,\middle|\, X_t = x \right]$$

subject to the controlled dynamics

$$\begin{aligned} dX_s &= a(s, X_s, \alpha_s) ds + b(s, X_s, \alpha_s) dW_s, \quad t < s \leq T, \\ X_t &= x. \end{aligned}$$

u is the solution to the HJB equation

$$0 = \partial_t u + \min_{\alpha \in A} \left[a(t,x,\alpha) \partial_x u + \frac{b(t,x,\alpha)^2}{2} \partial_{xx} u + h(t,x,\alpha) \right], \tag{4.33}$$

$$u(T,x) = g(x) + \lambda \cdot (\mathbb{1}_{\{G(x) > K\}} - \delta), \tag{4.34}$$

on the domain $(t,x) \in [0,T) \times \mathbb{R}$, which is directly related to the HJB equation (4.4) of the relaxed Problem 4 with perturbed terminal time boundary condition (4.5).

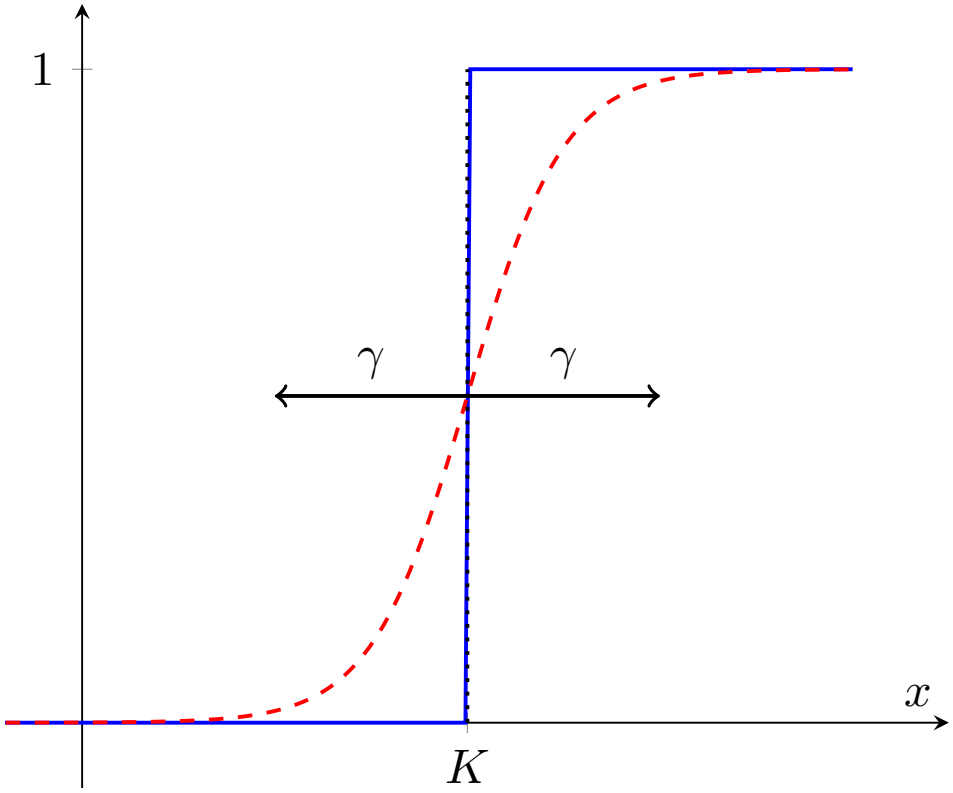


Figure 4.2: A schematic comparison of the indicator function (blue) with discontinuity in $x = K$ and a smoothed out version using the sigmoid function $\mathrm{sig}_\gamma(x - K)$. By using the sigmoid function defined in (4.2), the discontinuity is smoothened over a length of $\mathcal{O}(\gamma)$ in both directions.

The only difference is the use of an indicator function, representing the actual probability (4.1). Instead of using the notation u^λ for the solution of Equation (4.4) with terminal time boundary condition (4.5), we will denote the solution as u^γ, emphasizing the dependency on the smoothing parameter γ. A numerical approximation will be indicated by the discretization parameters Δt and Δx in superscript, i.e. $u^{\gamma,\Delta t,\Delta x}$ for a numerical approximation using $\Delta t > 0$ as the uniform step step and $\Delta x >$ as the uniform mesh in the spatial grid, see Equation (4.41). We can write the total error $\varepsilon_{\text{total}}$ of solving Problem 5, i.e. evaluating the dual function $d(\lambda)$, as the sum of the smoothing error ε_{s} and discretization error ε_{d}:

$$\begin{aligned}
\varepsilon_{\text{total}} &= |u(0,x_0) - u^{\gamma,\Delta t,\Delta x}(0,x_0)| && (4.35)\\
&= |u(0,x_0) - u^{\gamma}(0,x_0) + u^{\gamma}(0,x_0) - u^{\gamma,\Delta t,\Delta x}(0,x_0)| && (4.36)\\
&\leq \underbrace{|u(0,x_0) - u^{\gamma}(0,x_0)|}_{=:\varepsilon_{\text{s}}} + \underbrace{|u^{\gamma}(0,x_0) - u^{\gamma,\Delta t,\Delta x}(0,x_0)|}_{=:\varepsilon_{\text{d}}}. && (4.37)
\end{aligned}$$

4.6.1 Smoothing Error

For the analysis of the smoothing error

$$\varepsilon_{\mathrm{s}} = |u(0,x_0) - u^\gamma(0,x_0)| \tag{4.38}$$

we will state a Lemma and use its result in combination with an analysis carried out in [54] to bound the smoothing error asymptotically with respect to the smoothing parameter $\gamma > 0$. Without loss of generality, we assume that the terminal costs function g in the objective is zero, i.e. $g \equiv 0$.

Lemma 4.6.1. *Assume the optimal controls α and α^γ corresponding to u and u^γ respectively will be attained. For fixed $\lambda \geq 0$, it holds*

$$|u(0,x_0) - u^\gamma(0,x_0)| \leq \lambda \cdot \mathbb{E}\left[\mathbb{1}_{\{G(X_T)>K\}} - \mathrm{sig}_\gamma(G(X_T)-K)\right]. \tag{4.39}$$

Proof. Let us denote the optimally controlled dynamics of the problem corresponding to u by X and corresponding to u^γ by X^γ. By definition

$$\begin{aligned}
u^\gamma(0,x_0) - u(0,x_0) &\overset{(1)}{=} \mathbb{E}\left[\int_0^T h(t,X_t^\gamma,\alpha_t^\gamma)dt + \lambda\cdot\left(\mathrm{sig}_\gamma(G(X_T^\gamma)-K)-\delta\right)\right] - u(0,x_0)\\
&\quad + \mathbb{E}\left[\lambda\cdot\left(\mathbb{1}_{\{G(X_T^\gamma)>K\}}-\delta\right)\right] - \mathbb{E}\left[\lambda\cdot\left(\mathbb{1}_{\{G(X_T^\gamma)>K\}}-\delta\right)\right]\\
&\overset{(2)}{=} \lambda\cdot\mathbb{E}\left[\mathrm{sig}_\gamma(G(X_T^\gamma)-K) - \mathbb{1}_{\{G(X_T^\gamma)>K\}}\right] + \mathbb{E}\left[\int_0^T h(t,X_t^\gamma,\alpha_t^\gamma)dt\right]\\
&\quad + \mathbb{E}\left[u(T,X_T^\gamma)\right] - u(0,x_0)\\
&\overset{(3)}{=} \lambda\cdot\mathbb{E}\left[\mathrm{sig}_\gamma(G(X_T^\gamma)-K) - \mathbb{1}_{\{G(X_T^\gamma)>K\}}\right] + \mathbb{E}\left[\int_0^T h(t,X_t^\gamma,\alpha_t^\gamma)dt\right]\\
&\quad + \mathbb{E}\left[\int_0^T du(t,X_t^\gamma)dt\right]\\
&\overset{(4)}{=} \lambda\cdot\mathbb{E}\left[\mathrm{sig}_\gamma(G(X_T^\gamma)-K) - \mathbb{1}_{\{G(X_T^\gamma)>K\}}\right]\\
&\quad + \mathbb{E}\left[\int_0^T \left(\partial_t u + a\cdot\partial_x u + \frac{b^2}{2}\partial_{xx}u\right)(t,X_t^\gamma,\alpha_t^\gamma) + h(t,X_t^\gamma,\alpha_t^\gamma)dt\right]\\
&\overset{(5)}{\geq} \lambda\cdot\mathbb{E}\left[\mathrm{sig}_\gamma(G(X_T^\gamma)-K) - \mathbb{1}_{\{G(X_T^\gamma)>K\}}\right]\\
&\quad + \mathbb{E}\left[\int_0^T (-H+H)(t,X_t^\gamma,\partial_x u(t,X_t^\gamma),\partial_{xx}u(t,X_t^\gamma))dt\right]\\
&= \lambda\cdot\mathbb{E}\left[\mathrm{sig}_\gamma(G(X_T^\gamma)-K) - \mathbb{1}_{\{G(X_T^\gamma)>K\}}\right],
\end{aligned}$$

where in (1) we add and subtract the terminal time costs along the path $(X_t^\gamma, \alpha_t^\gamma)$. In (2), we rearranged the terms and used

$$\mathbb{E}\left[\lambda\cdot\left(\mathbb{1}_{\{G(X_T^\gamma)>K\}}-\delta\right)\right] = u(T,X_T^\gamma).$$

(3) is using the assumption that the value function or the solution to the Hamilton-Jacobi-Bellman equation from Theorem 2.2.2 is smooth and thus, without issues, differentiable or integrable. (4) is using Itô's formula from Theorem 2.1.4 on the latter integral and

rearranging terms to put the running costs at the end. Inequality (5) can be verified by Theorem 2.2.2, i.e. the fact that

$$\begin{aligned}
&a(t, X_t^\gamma, \alpha_t^\gamma)\partial_x u(t, X_t^\gamma) + \frac{b(t, X_t^\gamma, \alpha_t^\gamma)^2}{2}\partial_{xx} u(t, X_t^\gamma) + h(t, X_t^\gamma, \alpha_t^\gamma) \\
&\geq \min_{\alpha \in A}\left[a(t, X_t^\gamma, \alpha)\partial_x u(t, X_t^\gamma) + \frac{b(t, X_t^\gamma, \alpha)^2}{2}\partial_{xx} u(t, X_t^\gamma) + h(t, X_t^\gamma, \alpha)\right] \\
&= H(t, X_t^\gamma, \partial_x u(t, X_t^\gamma), \partial_{xx} u(t, X_t^\gamma))
\end{aligned}$$

and

$$\partial_t u(t, X_t^\gamma) = -H(t, X_t^\gamma, \partial_x u, \partial_{xx} u).$$

Similarly but along the path (X_t, α_t), we can derive

$$\begin{aligned}
u(0,x_0) - u^\gamma(0,x_0) &= \mathbb{E}\left[\int_0^T h(t, X_t, \alpha_t)dt + \lambda \cdot \left(\mathbb{1}_{\{G(X_T)>K\}} - \delta\right)\right] - u^\gamma(0,x_0) \\
&\quad + \mathbb{E}\left[\lambda \cdot \left(\mathrm{sig}_\gamma(G(X_T) - K) - \delta\right)\right] - \mathbb{E}\left[\lambda \cdot \left(\mathrm{sig}_\gamma(G(X_T) - K) - \delta\right)\right] \\
&= \lambda \cdot \mathbb{E}\left[\mathbb{1}_{\{G(X_T)>K\}} - \mathrm{sig}_\gamma(G(X_T) - K)\right] + \mathbb{E}\left[\int_0^T h(t, X_t, \alpha_t)dt\right] \\
&\quad + \mathbb{E}\left[u^\gamma(T, X_T)\right] - u^\gamma(0,x_0) \\
&= \lambda \cdot \mathbb{E}\left[\mathbb{1}_{\{G(X_T)>K\}} - \mathrm{sig}_\gamma(G(X_T) - K)\right] + \mathbb{E}\left[\int_0^T h(t, X_t, \alpha_t)dt\right] \\
&\quad + \mathbb{E}\left[\int_0^T du^\gamma(t, X_t)dt\right] \\
&= \lambda \cdot \mathbb{E}\left[\mathbb{1}_{\{G(X_T)>K\}} - \mathrm{sig}_\gamma(G(X_T) - K)\right] \\
&\quad + \mathbb{E}\left[\int_0^T \left(\partial_t u^\gamma + a\partial_x u^\gamma + \frac{b^2}{2}\partial_{xx} u^\gamma\right)(t, X_t) + h(t, X_t, \alpha_t)dt\right] \\
&\geq \lambda \cdot \mathbb{E}\left[\mathbb{1}_{\{G(X_T)>K\}} - \mathrm{sig}_\gamma(G(X_T) - K)\right] \\
&\quad + \mathbb{E}\left[\int_0^T \left(-H^\gamma + H^\gamma\right)(t, X_t, \partial_x u^\gamma(t, X_t), \partial_{xx} u^\gamma(t, X_t))dt\right] \\
&= \lambda \cdot \mathbb{E}\left[\mathbb{1}_{\{G(X_T)>K\}} - \mathrm{sig}_\gamma(G(X_T) - K)\right],
\end{aligned}$$

where

$$-\partial_t u^\gamma(t, X_t) = H^\gamma(t, X_t, \partial_x u^\gamma, \partial_{xx} u^\gamma).$$

Both results together imply inequality (4.39), i.e.

$$|u(0,x_0) - u^\gamma(0,x_0)| \leq \lambda \cdot \mathbb{E}\left[\mathbb{1}_{\{G(X_T)>K\}} - \mathrm{sig}_\gamma(G(X_T) - K)\right].$$

□

The result from Lemma 4.6.1 can now be used to analyse the behaviour of taking a smoothened version of the indicator function (4.2) in the probability (4.1). We can get

the asymptotic behaviour of the quantity from Lemma 4.6.1 using a Taylor-expansion of the density function ρ_T of the random variable $Y_T := G(X_T)$, i.e.

$$\begin{aligned}\mathbb{E}\left[\mathbb{1}_{\{G(X_T)>K\}} - \operatorname{sig}_\gamma(G(X_T)-K)\right] &= \int_{-\infty}^{\infty}\left(\mathbb{1}_{\{y>K\}} - \operatorname{sig}_\gamma(y-K)\right)\rho_T(y)dy \\ &\overset{(*)}{=} \gamma\cdot\int_{-\infty}^{\infty}\left(\mathbb{1}_{\{x>0\}} - \frac{1}{1+\exp(-x)}\right)\rho_T(\gamma x+K)dx \\ &= a_1\cdot\rho_T(K)\gamma + a_2\cdot\rho_T'(K)\gamma^2 + \mathcal{O}(\gamma^3),\end{aligned}$$

where in $(*)$ we transformed the integral using $y = \gamma x + K$ and

$$\begin{aligned}a_1 &= \int_{-\infty}^{\infty}\left(\mathbb{1}_{\{x>0\}} - \frac{1}{1+\exp(-x)}\right)dx = 0, \\ a_2 &= \int_{-\infty}^{\infty} x\cdot\left(\mathbb{1}_{\{x>0\}} - \frac{1}{1+\exp(-x)}\right)dx = \frac{\pi^2}{6}.\end{aligned}$$

The asymptotic behaviour of the smoothing error w.r.t. $\gamma > 0$ is thus of order

$$\varepsilon_{\mathrm{s}} \overset{(4.39)}{\leq} \lambda\cdot\mathbb{E}\left[\mathbb{1}_{\{G(X_T)>K\}} - \operatorname{sig}_\gamma(G(X_T)-K)\right] = \lambda\cdot\left(\frac{\pi^2\rho_T(K)}{6}\cdot\gamma^2 + \mathcal{O}(\gamma^3)\right) = \mathcal{O}\left(\gamma^2\right).$$

The fixed Lagrange Multiplier $\lambda \geq 0$ acts as a constant. We will continue using the constant $c_\gamma > 0$ as the constant in this asymptotic behaviour, i.e.

$$\varepsilon_{\mathrm{s}} = |u(0,x_0) - u^\gamma(0,x_0)| \leq c_\gamma\cdot\gamma^2. \tag{4.40}$$

4.6.2 Discretization Error

Recall the HJB equation in (4.4) with terminal time boundary condition (4.5) and the KBE in Equation (4.8) with terminal time boundary condition (4.9). The state space of the controlled process is $\mathbb{R}$, thus both the HJB equation and KBE are defined on the domain $[0,T]\times\mathbb{R}$. Numerically, we can not implement an infinite state space domain. Thus, we will limit the state space by a compact set $[c,d] \subset \mathbb{R}$, $c < d$, such that we can approximate the solutions to both equations on a grid discretizing the bounded domain $[0,T]\times[c,d]$ with artificial boundary condition

$$\frac{\partial^2 u}{\partial x^2} = 0 \quad \text{on} \quad \{c,d\}, \quad \text{for all } t\in[0,T],$$

as described in Section 3.3. Let $N_i, M_i \in \mathbb{N}$, $i = 1,2$ be positive integers corresponding to the discretization of $[0,T]\times[c,d]$ for the HJB equation ($i=1$) and KBE ($i=2$), i.e.

$$\Delta t = \frac{T}{N_i}, \quad \Delta x = \frac{|d-c|}{M_i}, \quad i = 1,2. \tag{4.41}$$

The analysis of the discretization error

$$\varepsilon_{\mathrm{d}} = |u^\gamma(0,x_0) - u^{\gamma,\Delta t,\Delta x}(0,x_0)|$$

from (4.35) can be adapted from Sun et al. in [48]. The proposed UFD scheme to solve the HJB equation in Equation (4.4) with terminal time boundary condition (4.5), see

Section 3.3, is a first order finite difference scheme. This implies an error bound for the discretization error of the form

$$\varepsilon_{\mathrm{d}} = |u^{\gamma}(0,x_0) - u^{\gamma,\Delta t,\Delta x}(0,x_0)| \leq c_1 N_1^{-1} + \bar{c}_1 M_1^{-1}$$

for constants $c_1, \bar{c}_1 > 0$. Additionally, due to the fact that the discretization parameters N_1, M_1 have to satisfy the CFL-condition in Equation (3.21), i.e.

$$\begin{aligned} &\frac{\Delta t}{\Delta x}\left(\|a\|_\infty + \frac{\|b\|_\infty}{\Delta x}\right) \leq 1 \\ \implies\ & T \cdot \frac{M_1}{|d-c|}\left(\|a\|_\infty + \frac{\|b\|_\infty \cdot M_1}{|d-c|}\right) \leq N_1, \end{aligned}$$

we can asymptotically bound the UFD scheme error by

$$\varepsilon_{\mathrm{d}} = |u^{\gamma}(0,x_0) - u^{\gamma,\Delta t,\Delta x}(0,x_0)| \leq c_1 N_1^{-1} + \bar{c}_1 M_1^{-1} \leq \tilde{c}_1 \cdot M_1^{-1}. \tag{4.42}$$

For the next section, we also need an error estimate on the subgradient evaluation, i.e. solving the KBE in Equation (4.8) with terminal time boundary condition (4.9) on $[0,T]\times[c,d]$. We use a Crank-Nicolson method, see Section 4.2.1, which is a second order finite difference scheme, i.e. the error can be bounded by

$$\varepsilon_{\mathrm{CN}} := |v^{\lambda}(0,x_0) - v^{\lambda,\Delta t,\Delta x}(0,x_0)| \leq c_2 N_2^{-2} + \bar{c}_2 M_2^{-2} \tag{4.43}$$

for some constants $c_2, \bar{c}_2 > 0$. We are aiming for an approximation that is relative to the violation probability $0 < \delta \ll 1$. Thus, if we impose a relative tolerance $\mathrm{TOL}_{\mathrm{r}} > 0$ on the absolute error representation (4.43), we have a relative error representation for the CN method

$$\varepsilon_{\mathrm{CN}}^{\mathrm{rel}} := \varepsilon_{\mathrm{CN}} \cdot \delta^{-1} \overset{!}{<} \mathrm{TOL}_{\mathrm{r}} \implies \varepsilon_{\mathrm{CN}} < \mathrm{TOL}_{\mathrm{r}} \cdot \delta. \tag{4.44}$$

4.6.3 Computational Work Analysis

In this section, we will investigate the computational work of the algorithms proposed in this chapter. Recall the positive integers N_i, M_i, $i = 1,2$ for the time and space discretization of the domain $[0,T]\times[c,d]$ of the HJB equation and KBE with relation (4.41). We can write the costs of solving the HJB equation using the UFD scheme as

$$\mathcal{W}_{\mathrm{UFD}} = \mathcal{O}(N_1 \cdot M_1) \tag{4.45}$$

and solving the KBE using the Crank-Nicolson method with a solver of work rate $\Gamma \in [1,3]$ as

$$\mathcal{W}_{\mathrm{CN}} = \mathcal{O}(N_2 \cdot M_2^{\Gamma}). \tag{4.46}$$

The costs of Algorithm 3 is given by

$$\mathcal{W}_{\mathrm{DG}} = \mathcal{O}\left(N_1 \cdot M_1 + N_2 \cdot M_2^{\Gamma}\right). \tag{4.47}$$

This is justified by the fact that for each interpolation point in (4.22), we need to solve one HJB equation and one KBE equation. Similarly, in each iteration of the subgradient

method in Algorithm 4, we need to solve multiple HJB equations and KBEs, resulting in the work

$$\mathcal{W}_{\text{SG}} = \mathcal{O}\left(K\left(N_1 \cdot M_1 + N_2 \cdot M_2^{\Gamma}\right)\right) \tag{4.48}$$

after $K \in \mathbb{N}$ iterations. Generally, the following Theorems 4.6.3, 4.6.2 and 4.6.4 are structured as follows: We aim to minimize the computational operations needed of solving a specific problem while accounting for an error with a prescribed tolerance. Different errors will have different imposed tolerances. For the UFD scheme approximation of the HJB equation and the corresponding error (4.42) as well as the smoothing error (4.40), we will denote the prescribed absolute tolerance as $\text{TOL} > 0$, resulting in a constraint in the optimization (minimization of the costs) given by

$$\tilde{c}_1 \cdot M_1^{-1} \approx \text{TOL}, \tag{4.49}$$
$$c_\gamma \cdot \gamma^2 \approx \text{TOL}, \tag{4.50}$$

respectively. Additionally, we have to account for the stability of the UFD scheme, i.e. the CFL condition (3.21), yielding an additional constraint

$$T \cdot \frac{M_1}{|d-c|}\left(\|a\|_\infty + \frac{\|b\|_\infty \cdot M_1}{|d-c|}\right) \approx N_1$$

whenever we approximate a HJB equation. For the KBE approximation using the CN method proposed in Section 4.2.1 and corresponding absolute error representation (4.43), we impose a relative tolerance $\text{TOL}_\text{r} > 0$. The constraint included in the minimization of the costs is a combination of the results in (4.43) and (4.44), namely

$$c_2 N_2^{-2} + \bar{c}_2 M_2^{-2} \approx \text{TOL}_\text{r} \cdot \delta. \tag{4.51}$$

These three constraints will arise in all the following Theorems. The first result is on the asymptotic optimal computational work of Algorithm 3, i.e. computing the relative dual gap for a fixed LM $\lambda \geq 0$. Note that the costs of solving the root-finding problem in (4.24) can be neglected with respect to the costs of solving the HJB equation and KBE.

Theorem 4.6.2 (Optimal Work Algorithm 3)**.** *Let* $\text{TOL} > 0$ *be the prescribed absolute tolerance in (4.49), (4.50) and* $\text{TOL}_\text{r} > 0$ *be the prescribed relative tolerance in (4.51). Consider the following optimization problem*

$$\begin{aligned}
\min_{N_1,M_1,N_2,M_2} \mathcal{W}_{\text{DG}} = \min_{N_1,M_1,N_2,M_2} N_1 \cdot M_1 + N_2 \cdot M_2^{\Gamma} \\
s.t. \qquad \tilde{c}_1 \cdot M_1^{-1} \approx \text{TOL}, \\
T \cdot \frac{M_1}{|d-c|}\left(\|a\|_\infty + \frac{\|b\|_\infty \cdot M_1}{|d-c|}\right) \approx N_1, \\
c_2 \cdot N_2^{-2} + \bar{c}_2 \cdot M_2^{-2} \approx \text{TOL}_\text{r} \cdot \delta, \\
c_\gamma \cdot \gamma^2 \approx \text{TOL}.
\end{aligned}$$

Then, the computational work of Algorithm 3 that minimizes the computational work of all the subroutines involved, satisfying the constraints, is

$$\mathcal{W}^*_{\text{DG}} = \mathcal{O}\left(\text{TOL}^{-3} + (\text{TOL}_\text{r} \cdot \delta)^{-\frac{1+\Gamma}{2}}\right)$$

with optimal parameters

$$N_1^* = \frac{c_3\tilde{c}_1}{\mathrm{TOL}} + \frac{\bar{c}_3\tilde{c}_1^2}{\mathrm{TOL}^2}, \quad N_2^* = \sqrt{c_2(1+\Gamma)}\,(\mathrm{TOL}_\mathrm{r}\cdot\delta)^{-\frac{1}{2}}$$
$$M_1^* = \frac{\tilde{c}_1}{\mathrm{TOL}}, \qquad\qquad M_2^* = \sqrt{\bar{c}_2\cdot\frac{\Gamma+1}{\Gamma}}\cdot(\mathrm{TOL}_\mathrm{r}\cdot\delta)^{-\frac{1}{2}},$$

where

$$c_3 = \frac{T\cdot\|a\|_\infty}{|d-c|}, \quad \bar{c}_3 = \frac{T\cdot\|b\|_\infty}{|d-c|^2}$$

and

$$\gamma^* = c_\gamma^{-\frac{1}{2}}\cdot\mathrm{TOL}^{\frac{1}{2}}.$$

Proof. The proof is given in Appendix A.5. □

Theorem 4.6.2 is essentially the combined results of Theorem 3.3.1 and Theorem 4.2.1. To perform the same analysis on Algorithm 5, we recall the result in (4.31), in which an absolute tolerance $\mathrm{TOL}_\mathrm{b} > 0$ on the bisection methods error on the distance after the kth iteration and the optimal penalization parameter β^* is imposed. Additionally to Theorem 4.6.2 constraints, the next Theorem 4.6.3 includes (4.31) as a constraint and gives a result on the number of iterations $k \in \mathbb{N}$ needed to satisfy the tolerance $\mathrm{TOL}_\mathrm{b} > 0$ while accounting for a minimal asymptotic computational work. The total computational work $\mathcal{W}_\mathrm{bisec}$ of Algorithm 5 is

$$\mathcal{W}_\mathrm{bisec} = \mathcal{O}\left(k\cdot\left(N_1\cdot M_1 + N_2\cdot M_2^\Gamma\right)\right).$$

Theorem 4.6.3 (Optimal Work Algorithm 5)**.** *Let* $\mathrm{TOL} > 0$ *be the prescribed absolute tolerance in (4.49), (4.50),* $\mathrm{TOL}_\mathrm{r} > 0$ *be the prescribed relative tolerance in (4.51) and* $\mathrm{TOL}_\mathrm{b} > 0$ *be the prescribed absolute tolerance in (4.31). Consider the following optimization problem*

$$\min_{N_1,M_1,N_2,M_2,k} \mathcal{W}_\mathrm{bisec} = \min_{N_1,M_1,N_2,M_2,k} k\left(N_1\cdot M_1 + N_2\cdot M_2^\Gamma\right)$$
$$\begin{aligned}
s.t. \qquad \tilde{c}_1\cdot M_1^{-1} &\approx \mathrm{TOL},\\
T\cdot\frac{M_1}{|d-c|}\left(\|a\|_\infty + \frac{\|b\|_\infty\cdot M_1}{|d-c|}\right) &\approx N_1,\\
c_2\cdot N_2^{-2} + \bar{c}_2\cdot M_2^{-2} &\approx \mathrm{TOL}_\mathrm{r}\cdot\delta,\\
c_\gamma\cdot\gamma^2 &\approx \mathrm{TOL},\\
\frac{|\beta_0-\bar{\beta}_0|}{2^k} &\approx \mathrm{TOL}_\mathrm{b}.
\end{aligned}$$

Then, the computational work of Algorithm 5 that minimizes the computational work of all the subroutines involved, satisfying the constraints, is

$$\mathcal{W}_\mathrm{bisec}^* = \mathcal{O}\left(\log(\mathrm{TOL}_\mathrm{b}^{-1})\left(\mathrm{TOL}^{-3} + (\mathrm{TOL}_\mathrm{r}\cdot\delta)^{-\frac{1+\Gamma}{2}}\right)\right)$$

with optimal parameters

$$N_1^* = \frac{c_3\tilde{c}_1}{\mathrm{TOL}} + \frac{\bar{c}_3\tilde{c}_1^2}{\mathrm{TOL}^2}, \quad N_2^* = \sqrt{c_2(1+\Gamma)}\,(\mathrm{TOL}_\mathrm{r}\cdot\delta)^{-\frac{1}{2}}$$
$$M_1^* = \frac{\tilde{c}_1}{\mathrm{TOL}}, \qquad\qquad M_2^* = \sqrt{\bar{c}_2\cdot\frac{\Gamma+1}{\Gamma}}\cdot(\mathrm{TOL}_\mathrm{r}\cdot\delta)^{-\frac{1}{2}},$$

where

$$c_3 = \frac{T\cdot\|a\|_\infty}{|d-c|}, \quad \bar{c}_3 = \frac{T\cdot\|b\|_\infty}{|d-c|^2}$$

and

$$\gamma^* = c_\gamma^{-\frac{1}{2}}\cdot\mathrm{TOL}^{\frac{1}{2}}, \quad k^* = \left\lceil \frac{1}{\log(2)}\cdot\log\left(\frac{|\beta_0-\bar{\beta}_0|}{\mathrm{TOL}_b}\right)\right\rceil.$$

Proof. The proof is given in Appendix A.6. □

Finally, we would like to investigate the subgradient method from Algorithm 4. The relative dual gap estimation proposed in Section 4.3 does not return a theoretical bound on the number of iterations needed to satisfy a prescribed tolerance $\mathrm{TOL}_{\mathrm{SG}} > 0$ on the distance of the optimal solution to the current best iterations solution. Hence, we will use the sub-optimality bound in (2.37) to give a result on the maximum number of iterations $K \in \mathbb{N}$ needed to reach an absolute tolerance of $\mathrm{TOL}_{\mathrm{SG}} > 0$. We recall the positive constants R and S from the convergence assumptions of the subgradient method in Section 4.4 as well as the fact that the sub-optimality bound in (2.37) is minimized by a sequence of step sizes

$$t_k = \frac{R}{S\sqrt{K}}, \quad k = 1,\dots,K.$$

In the following Theorem 4.6.4, K is the number of steps in the subgradient method to converge within some absolute tolerance $\mathrm{TOL}_{\mathrm{SG}} > 0$. The additional constraint in Theorem 4.6.4 comes from the result in (2.38).

Theorem 4.6.4 (Optimal Work Algorithm 4)**.** *Let* $\mathrm{TOL} > 0$ *be the prescribed absolute tolerance in (4.49), (4.50),* $\mathrm{TOL}_\mathrm{r} > 0$ *be the prescribed relative tolerance in (4.51) and* $\mathrm{TOL}_{\mathrm{SG}} > 0$ *be the prescribed absolute tolerance in (2.38). Consider the following optimization problem*

$$\begin{aligned}
\min_{N_1,M_1,N_2,M_2,K} \mathcal{W}_{\mathrm{SG}} &= \min_{N_1,M_1,N_2,M_2,K} K\left(N_1\cdot M_1 + N_2\cdot M_2^{\Gamma}\right)\\
s.t. \qquad \tilde{c}_1\cdot M_1^{-1} &\approx \mathrm{TOL},\\
T\cdot\frac{M_1}{|d-c|}\left(\|a\|_\infty + \frac{\|b\|_\infty\cdot M_1}{|d-c|}\right) &\approx N_1\\
c_2\cdot N_2^{-2} + \bar{c}_2\cdot M_2^{-2} &\approx \mathrm{TOL}_\mathrm{r}\cdot\delta,\\
c_\gamma\cdot\gamma^2 &\approx \mathrm{TOL},\\
\frac{R\cdot S}{\sqrt{K}} &\approx \mathrm{TOL}_{\mathrm{SG}}.
\end{aligned}$$

Then

$$\mathcal{W}^*_{\mathrm{SG}} = \mathcal{O}\left(\mathrm{TOL}_{\mathrm{SG}}^{-2}\left(\mathrm{TOL}^{-3} + (\mathrm{TOL}_{\mathrm{r}}\cdot\delta)^{-\frac{1+\Gamma}{2}}\right)\right)$$

with

$$N_1^* = \frac{c_3\tilde{c}_1}{\mathrm{TOL}} + \frac{\bar{c}_3\tilde{c}_1^2}{\mathrm{TOL}^2}, \quad N_2^* = \sqrt{c_2(1+\Gamma)}\,(\mathrm{TOL}_{\mathrm{r}}\cdot\delta)^{-\frac{1}{2}}$$
$$M_1^* = \frac{\tilde{c}_1}{\mathrm{TOL}}, \qquad\qquad M_2^* = \sqrt{\bar{c}_2\cdot\frac{\Gamma+1}{\Gamma}}\cdot(\mathrm{TOL}_{\mathrm{r}}\cdot\delta)^{-\frac{1}{2}},$$

where

$$c_3 = \frac{T\cdot\|a\|_\infty}{|d-c|}, \quad \bar{c}_3 = \frac{T\cdot\|b\|_\infty}{|d-c|^2}$$

and

$$\gamma^* = c_\gamma^{-\frac{1}{2}}\cdot\mathrm{TOL}^{\frac{1}{2}} \quad \text{and} \quad K^* = \left(\frac{RS}{\mathrm{TOL}_{\mathrm{SG}}}\right)^2.$$

Proof. The proof is given in Appendix A.7. □

Theorem 4.6.4 can be used to get an upper bound on the runtime using the constant sequence of step sizes $t_k = R/(S\sqrt{K})$ for a prescribed maximum number of iterations $K \in \mathbb{N}$.

4.7 Numerical results

In this section, we will visualize the proposed solution methodology on simple examples and check the convergence results of Theorem 4.6.2 and 4.6.3 numerically. We will investigate the runtime of the algorithms and show that taking a chance constraint with a violation probability $0 < \delta \ll 1$ instead of enforcing an almost sure constraint can reduce the running costs of the resulting objective. First, we will recap the example from Chapter 3 in which we investigated quadratic optimal control of Ornstein-Uhlenbeck processes without chance constraints. We will now use Problem 1 and add a chance constraint. The problem can be written as follows.

Problem 8 (Chance constrained QSOC)**.** *Consider the following optimization problem,*

$$\min_{\alpha\in\mathcal{A}} \mathbb{E}\left[\int_0^T X(t)^2 + \mu\cdot\alpha(t,X(t))^2\,dt\right], \quad s.t. \quad \mathbb{P}\left(X_T > K\right) \leq \delta$$

over the set of admissible deterministic Markovian controls $\mathcal{A} = \{\alpha : [0,T]\times\mathbb{R}\to A\}$*, which are adapted to the filtration* $\mathcal{F}_t$ *of the process* $X(t)\in\mathbb{R}$ *for all* $t\in[0,T]$ *which follows the controlled dynamics*

$$\begin{cases} dX(t) &= (-\theta X(t) + \alpha(t,X(t)))\,dt + \sigma\cdot dW(t), \quad 0 < t < T, \\ X(0) &= x_0, \end{cases}$$

for given parameters $\theta > 0, \mu > 0, \sigma > 0, x_0 \in \mathbb{R}$*, finite terminal time* $T > 0$ *and set of control values* $A \subset \mathbb{R}$*.*

Additionally, we will have a second (toy-) problem, namely a simple drift-controlled Wiener process, on which we will also apply our theory for numerical results. The problem can be written as follows.

Problem 9 (Chance constrained Bang-Bang)**.** *Consider the following optimization problem,*

$$\min_{\alpha\in\mathcal{A}} \mathbb{E}\left[\int_0^T X(t)^2\, dt\right], \quad s.t. \quad \mathbb{P}\left(X_T > K\right) \leq \delta$$

over the set of admissible deterministic Markovian controls $\mathcal{A} = \{\alpha : [0,T] \times \mathbb{R} \to A\}$*, which are adapted to the filtration* $\mathcal{F}_t$ *of the process* $X(t) \in \mathbb{R}$ *for all* $t \in [0,T]$ *which follows the controlled dynamics*

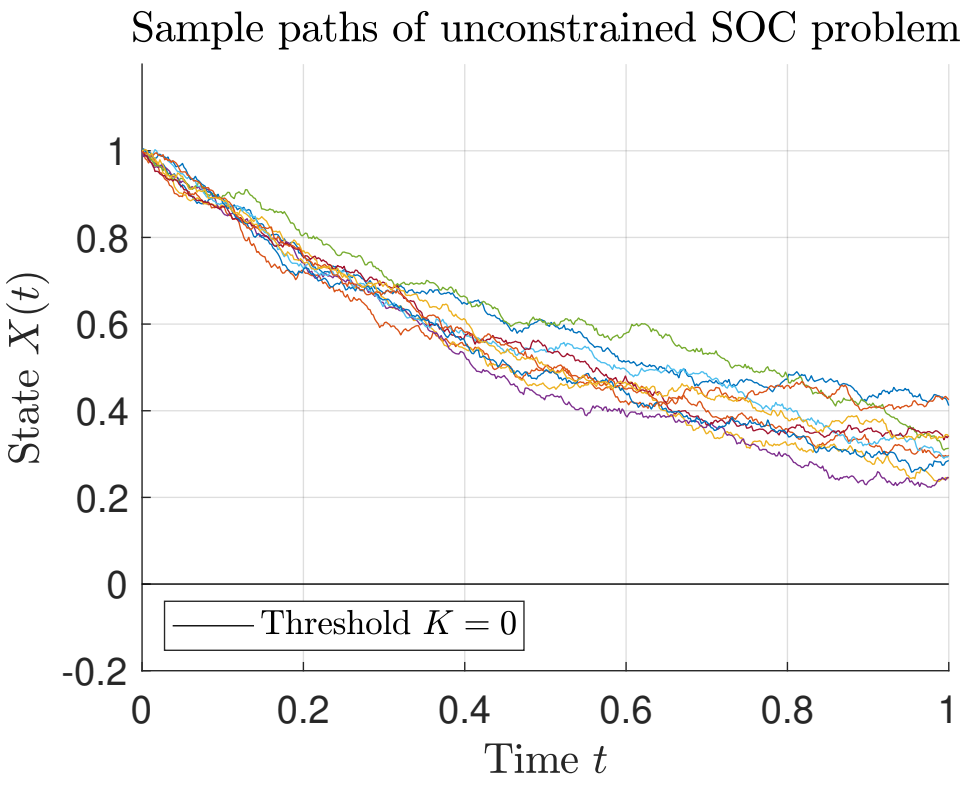
$$\begin{cases} dX(t) &= \alpha(t, X(t))\, dt + \sigma \cdot dW(t), \quad 0 < t < T, \\ X(0) &= x_0, \end{cases}$$

for given parameters $\sigma > 0, x_0 \in \mathbb{R}$*, finite terminal time* $T > 0$ *and compact set of control values* $A \subset \mathbb{R}$*.*

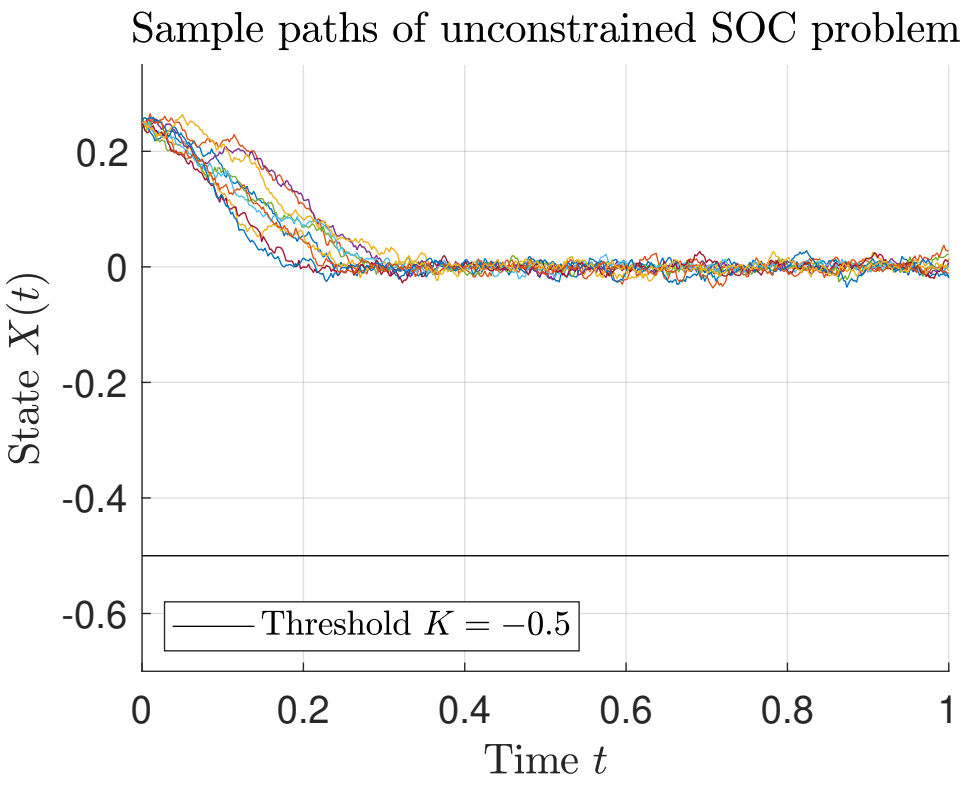
The main difference between Problem 8 and Problem 9 is the fact that in the first problem we have a regularized control in form of a regularization parameter $\mu > 0$ and its appearance in the objective itself. In Problem 8, the optimal control is dependent on the partial derivative w.r.t. the state space direction x of the value function, see Equation (3.6). This will result in a singularity of the optimal control due to the indicator function in the terminal time boundary condition (4.34) of the non-smoothened relaxed problem. This also supports the necessity of using a smoothened version of the indicator function in the numerical solver. In contrast, Problem 9 yields an optimal control that only takes the extreme values in the compact set $A \subset \mathbb{R}$ of control values. This phenomenon is called a bang-bang control. The set $A \subseteq \mathbb{R}$ in Problem 8 does not need to be compact and can be $A = \mathbb{R}$. Due to numerical issues, namely oscillations arising with solving the KBE using the proposed CN method and the optimal controls of the relaxed Problem 4, we impose a compact subset $A \subset \mathbb{R}$ for Problem 8 as well. The following experiments were conducted using the values from Table 4.1. Some sample paths of the unconstrained optimal control problem for both examples are displayed in Figure 4.3 using Theorem 3.3.1 to get the optimal discretization parameters for a prescribed absolute tolerance of $\text{TOL} = 10^{-1}$ for the

Description	Parameter	Problem 8	Problem 9
Terminal time	T	1	1
State space boundaries	$[c,d]$	$[-10,10]$	$[-2,2]$
Initial value of controlled SDE	x_0	1	0.25
Drift of controlled SDE	θ	1	-
Diffusion of controlled SDE	σ	0.1	0.1
Threshold of chance constraint	K	0	-0.5
Control regularization	μ	1	-
Compact set of control values	A	$[-5,5]$	$[-1,1]$

Table 4.1: Standard parameters used in all experiments carried out in this section unless specified otherwise. First column is the description or interpretation of the parameter in the second column. The third and fourth column are the specific parameter values in Problem 8 and Problem 9, respectively.

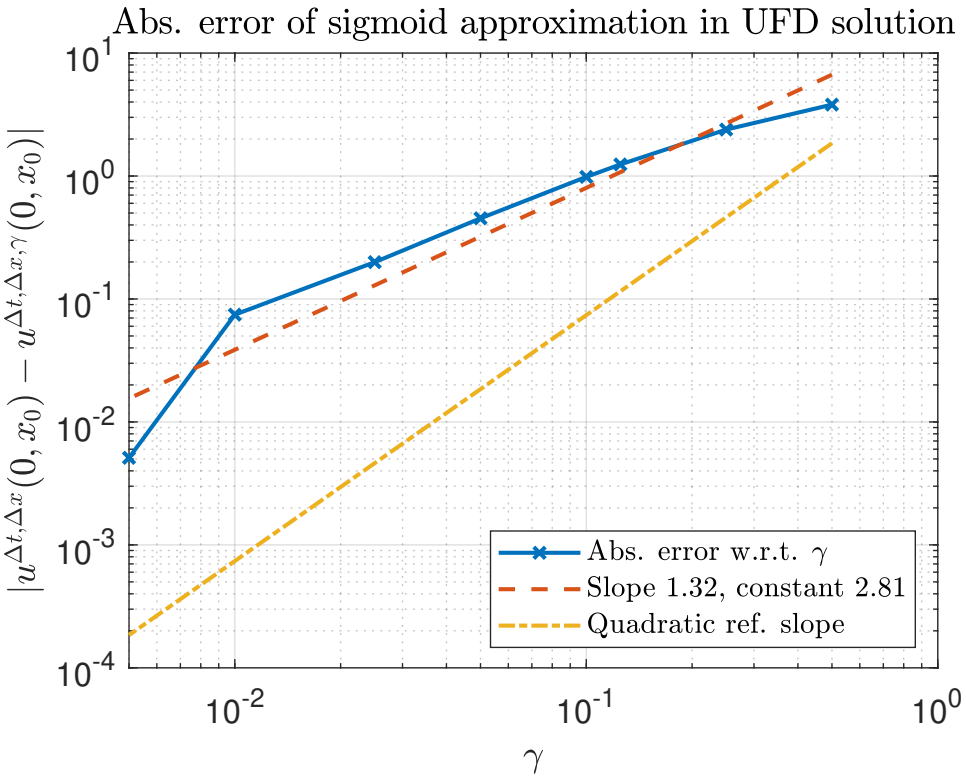


(a) Unconstrained Problem 8

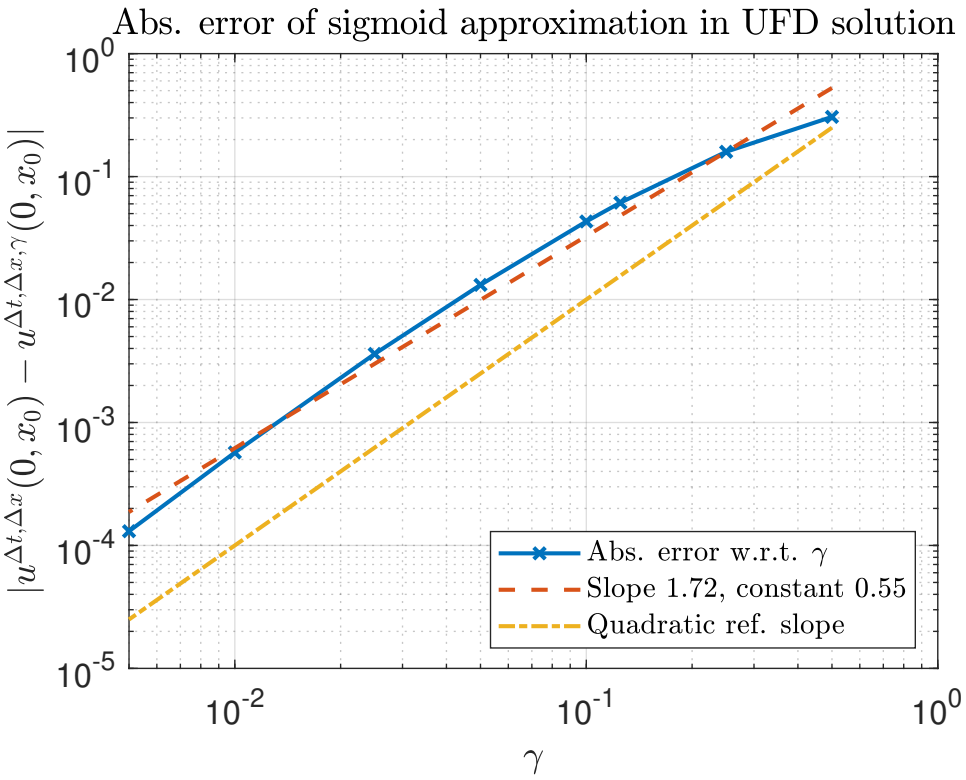


(b) Unconstrained Problem 9

Figure 4.3: 10 sample paths of the optimally controlled SDE on a discretized set of $N = 500$ time steps. The samples were computed using the procedure described in Equation (2.24) with $N = 500$ and the parameters from Table 4.1. The resulting optimally controlled sample paths for enforcing the constraints are shown in Figure 4.11b and 4.13b.



(a) Problem 8: $c_\gamma \approx 2.81$.



(b) Problem 9: $c_\gamma \approx 0.55$.

Figure 4.4: Estimation of the constant c_γ and rate of the induced smoothing using a sigmoid approximation with parameter γ.

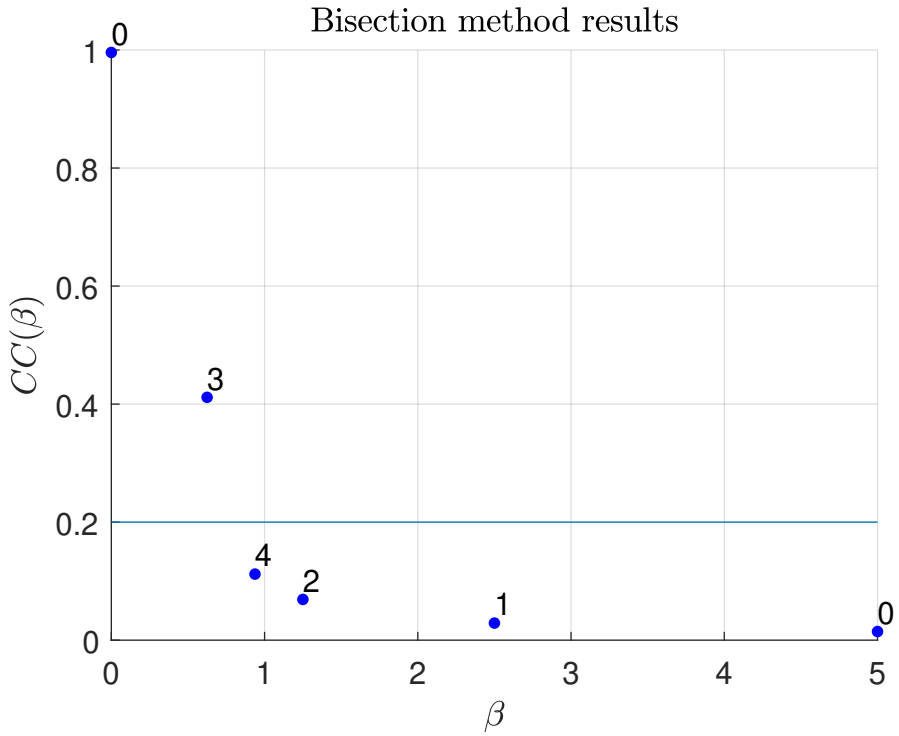
UFD approximation of the HJB equation. For Problem 8 this results in $N_1 = 298$, $M_1 = 419$ and for Problem 9 in $N_1 = 170$, $M_1 = 274$. The black horizontal line in Figure 4.3 indicates the threshold K of the, soon to be considered, chance constraint. Next, we need to estimate some constants used in the theorems of Section 4.6. First, we investigate the smoothing constant $c_\gamma > 0$. For that, we will solve Problem 8 and 9 using a fixed number of parameters and only vary the smoothing parameter from $\gamma = 0.5$ down to $\gamma = 0.005$. We compare the resulting solutions with a reference solution that uses a smaller parameter $\gamma = 10^{-4}$. The resulting rates are displayed in Figure 4.4. We can see that the numerical approximation of the rates are below the expected theoretical rate of two. This might be due to other numerical error sources and the fact that the reference $\gamma = 10^{-4}$ is too large w.r.t. the γ's that are used in Figure 4.4, i.e. the lowest is $\gamma = 5 \cdot 10^{-3}$ which might be already too close to the reference to get a reliable asymptotic behaviour representation. The results for the constants $\tilde{c}_1 > 0$ from the UFD induced error and $c_2, \bar{c}_2 > 0$ from the Crank-Nicolson induced error are analysed in Supplementary Material B.1.

4.7.1 Algorithm 5: Finding an initial LM λ_0

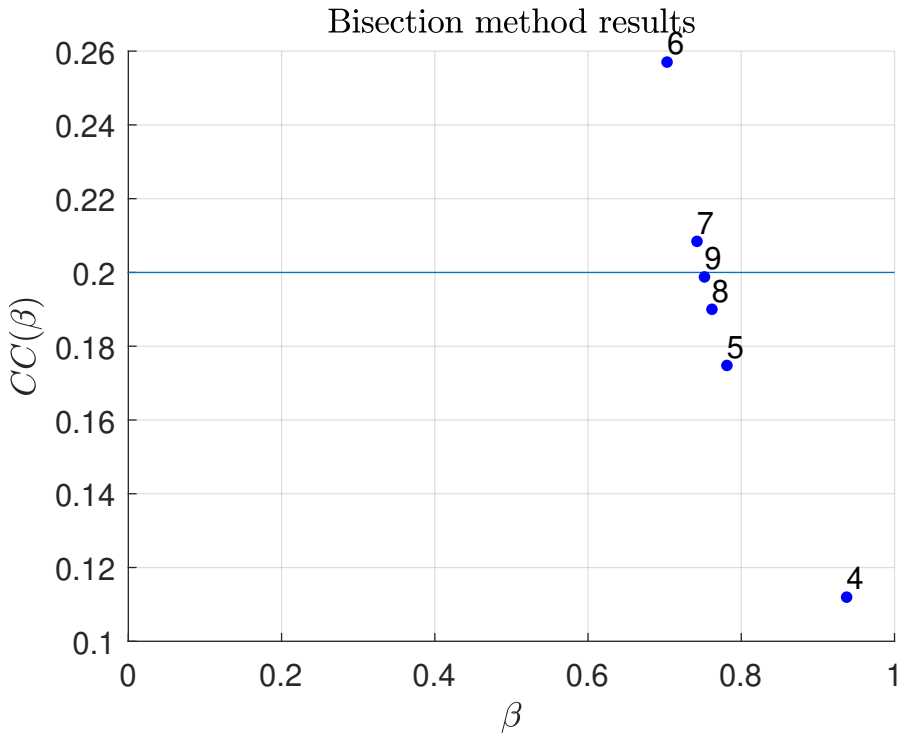
Next, we will investigate Algorithm 5. An example run of Algorithm 5 using the parameters in Table 4.1 with a violation probability of $\delta = 0.2$ and tolerances $\mathrm{TOL} = 10^{-2}$ and $\mathrm{TOL}_\mathrm{r} = 10^{-1}$ is displayed in Figure 4.5 for Problem 8 and in Figure 4.6 for Problem 9. The results for Problem 8 with $\mathrm{TOL}_\mathrm{b} = 10^{-2}$, resulting in $k^* = 9$, are displayed in Figure 4.5a and 4.5b and for Problem 9 with $\mathrm{TOL}_\mathrm{b} = 10^{-3}$, resulting in $k^* = 10$, in Figure 4.6a and 4.6b. We observe that the function $\mathrm{CC}_\mathrm{bisec}$ defined in (4.29) is very steep, i.e. small changes in the penalization parameter β result in large changes in the function $\mathrm{CC}_\mathrm{bisec}$. This is more evident in Problem 9. This suggests that we can use higher tolerances in Algorithm 5 to find a suitable initial LM $\lambda_0 \geq 0$ for Algorithm 4. The resulting solutions of Algorithm 5 can be used as initial values for Algorithm 4 to reduce the runtime of the subgradient method due to the lower value of R in Assumption 1 in Equation (2.33) which is crucial for the convergence of the SGM. An investigation of this is done in Section 4.7.3. Next, we would like to verifying the optimal work results from Theorem 4.6.3 numerically and see if the optimal work behaves as

$$\mathcal{W}^* = \mathcal{O}(\log(\mathrm{TOL}_\mathrm{b}^{-1})).$$

For that, we will fix the discretization tolerances for the UFD scheme $\mathrm{TOL} > 0$ and CN method $\mathrm{TOL}_\mathrm{r} > 0$. Note that these tolerances also appear in the optimal work result for the bisection method in Theorem 4.6.3 but act as constants when analysing the runtime of the algorithm with respect to $\mathrm{TOL}_\mathrm{b} > 0$. The results for Problem 8 and Problem 9 are displayed in Figure 4.7a and 4.7b, respectively.

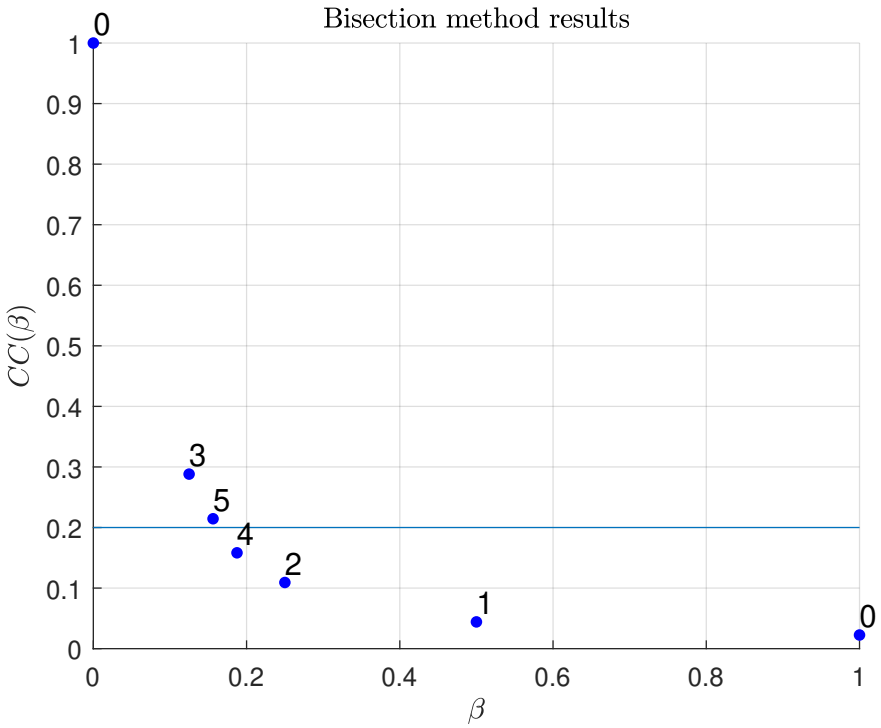


(a) Problem 8: iterations 1-4, $\bar{\beta}_0 = 5$

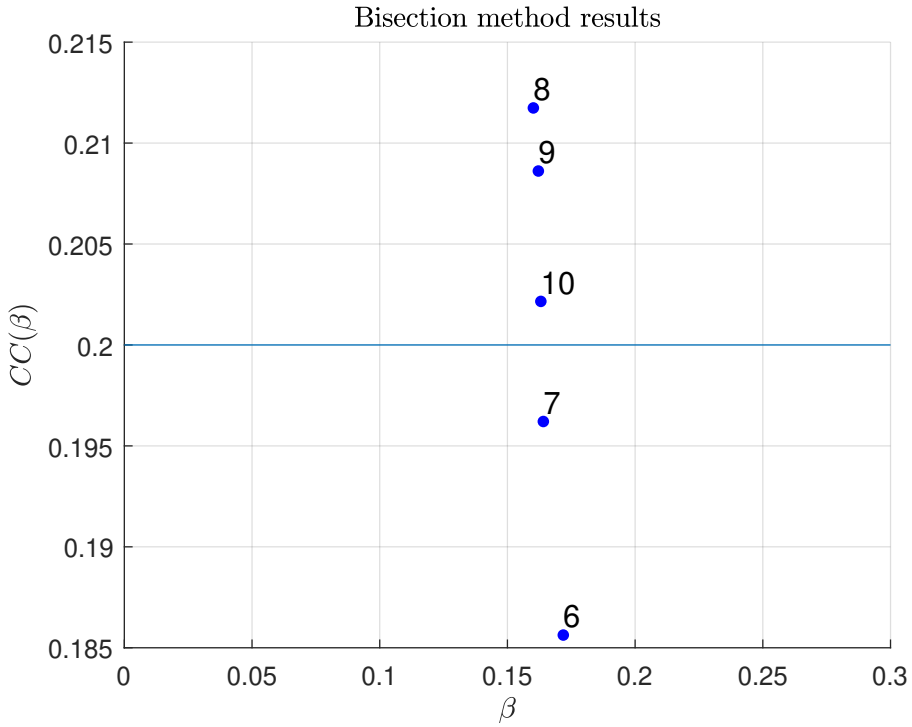


(b) Problem 8: iterations 5-9

Figure 4.5: Example run of Algorithm 5 for Problem 8 with $\mathrm{TOL}_\mathrm{b} = 10^{-2}$ 4.5a, 4.5b. The violation probability is $\delta = 0.2$ and $\mathrm{TOL} = 10^{-2}$ and $\mathrm{TOL}_\mathrm{r} = 10^{-1}$. The resulting penalization parameters are $\beta^* \approx 0.752$ with $\mathrm{CC}_{\mathrm{bisec}}(\beta^*) \approx 0.199$.

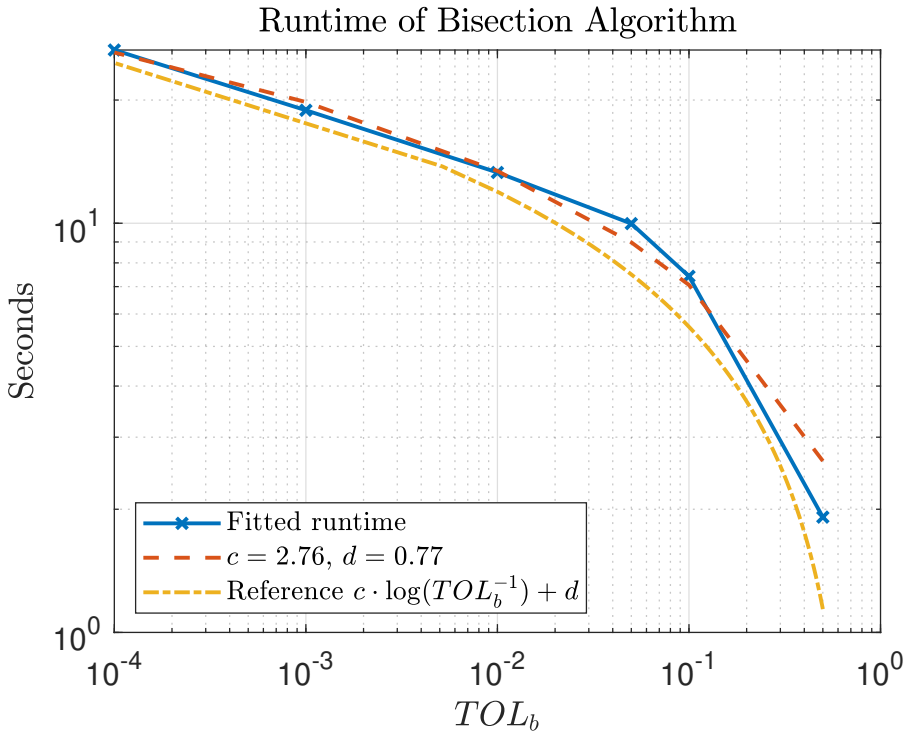


(a) Problem 9: iterations 1-5, $\bar{\beta}_0 = 1$

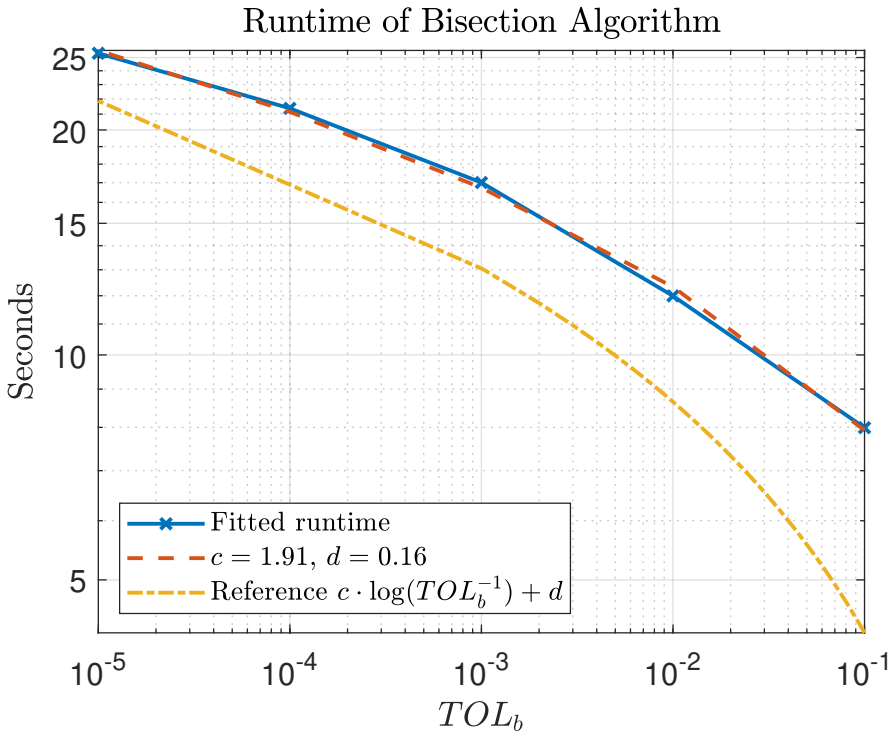


(b) Problem 9: iterations 6-10

Figure 4.6: Example run of Algorithm 5 for Problem 9 with $\mathrm{TOL}_\mathrm{b} = 10^{-3}$ 4.6a, 4.6b. The violation probability is $\delta = 0.2$ and $\mathrm{TOL} = 10^{-2}$ and $\mathrm{TOL}_\mathrm{r} = 10^{-1}$. The resulting penalization parameters are $\beta^* \approx 0.172$ with $\mathrm{CC}_{\mathrm{bisec}}(\beta^*) \approx 0.203$.



(a) Problem 8, TOL = 0.2, $\text{TOL}_\text{r} = 0.25$



(b) Problem 9, TOL = 0.05, $\text{TOL}_\text{r} = 0.1$

Figure 4.7: Runtime of the bisection method in Algorithm 5 for different tolerances $\text{TOL}_\text{b} > 0$.

4.7.2 Algorithm 3: Estimating the relative dual gap

In this section, we will illustrate the construction of the primal feasible solution η^{κ^*} in the relative dual gap estimation from Section 4.3 and show the asymptotic behaviour of the optimal work in Theorem 4.6.2 numerically. For an arbitrary LM $\lambda \geq 0$, we compute the solution to the HJB equation in (4.4) with terminal time boundary condition (4.5). We solve the equation and in return get the value function u^λ and optimal control α^λ on a grid of size $N_1 \times M_1$ of the truncated domain $[0,T] \times [c,d]$. When estimating the relative dual gap, we want to construct a primal feasible solution η^{κ^*} that satisfies the chance constraint and also be as close to the current dual solution α^λ as possible. This motivates, for a given λ, the objective in the auxiliary Problem 6

$$\min_{\eta \in \mathcal{A}_{\text{rel}}} \mathbb{E}\left[\int_0^T \left(\alpha^\lambda(s, X_s^\kappa) - \eta(s, X_s^\kappa)\right)^2 ds\right] + \kappa \cdot \mathbb{E}\left[\text{sig}_\gamma(X_T^\kappa - K)\right].$$

Since Algorithm 3 is an sub-procedure only, we do not want to solve it using a subgradient method itself but rather work heuristically. The procedure is described in Section 4.3. An illustrative example for Problem 8 with fixed LM $\lambda = 0.5$ is given in Figure 4.8. The resulting penalization parameter κ^* corresponding to the near-optimal solution η^{κ^*} of the auxiliary Problem 6 for the fixed $\lambda \geq 0$ satisfies

$$\text{CC}(\kappa^*) \approx \delta$$

and is shown in green in Figure 4.8a. It is the interception of the quadratic polynomial p_2 in Equation (4.23) and the violation probability δ as described in Section 4.3. A visualization of both current control α^λ and its corresponding primal feasible near-optimal control η^{κ^*} is given in Figure 4.8b. Computing the current HJB approximation for $\lambda = 0.5$ and its KBE approximation took 53 seconds in contrast to the computation of η^{κ^*} which in total took 909 seconds, more than 17 times as long. This might be caused by the interpolations needed to approximate the numerical approximation of the optimal control α^λ on a different grid than the computation of the optimal control corresponding to Problem 6. In theory, the relative dual gap estimation should only take about 5 times as long as one iteration of the SGM (in case one does not need to interpolate for different desired values). Using η^{κ^*}, we can compute an approximation of the primal costs using the PDE in (4.25) with terminal time boundary condition (4.26). We can then estimate the relative dual gap corresponding to the current LM $\lambda \geq 0$ by computing the quantity

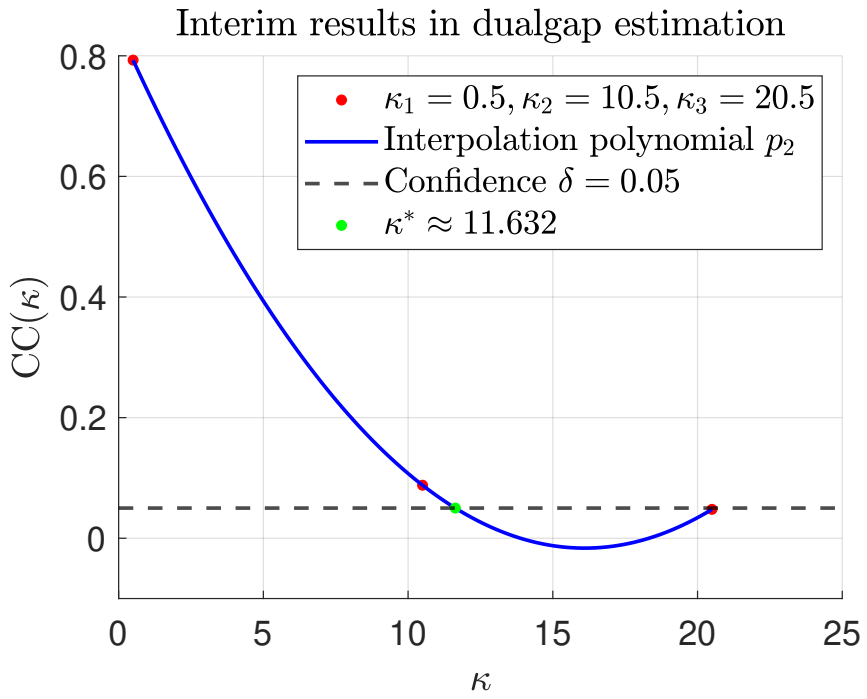
$$d^\lambda = \frac{w(0,x_0) - u^\lambda(0,x_0)}{u^\lambda(0,x_0)}.$$

For the optimal computational work, Theorem 4.6.2 yields an asymptotic runtime behaviour of order

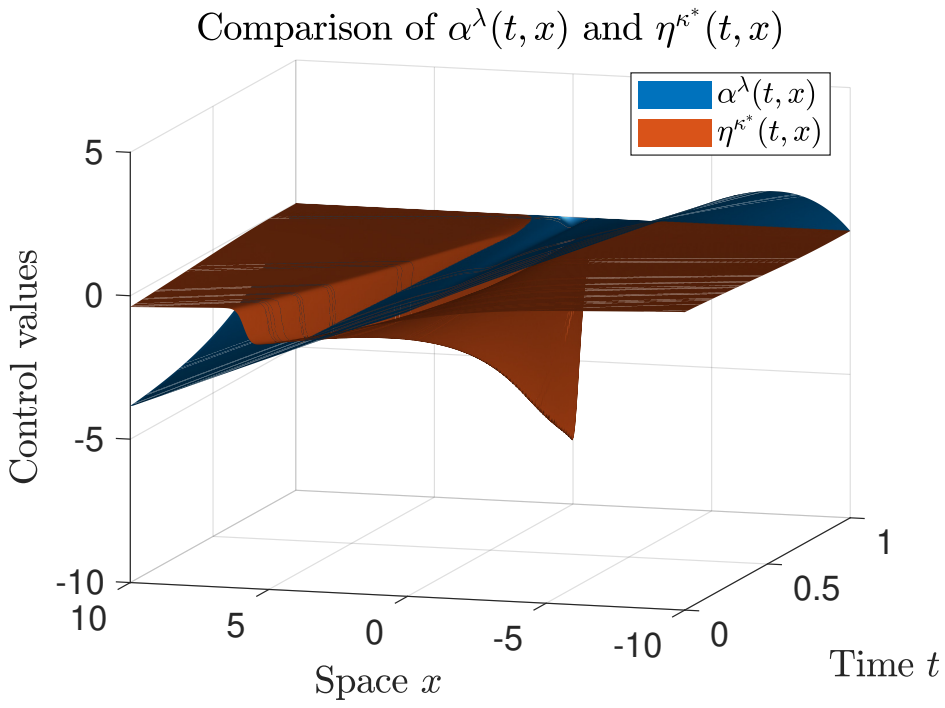
$$\mathcal{W}^* = \mathcal{O}\left(\text{TOL}^{-3} + (\text{TOL}_\text{r} \cdot \delta)^{-\frac{1+\Gamma}{2}}\right).$$

The work rate $\Gamma \in [1,3]$ for our implementation has been numerically verified to be $\Gamma \approx 1$. This will simplify the optimal work to

$$\mathcal{W}^* = \mathcal{O}\left(\text{TOL}^{-3} + (\text{TOL}_\text{r} \cdot \delta)^{-1}\right).$$



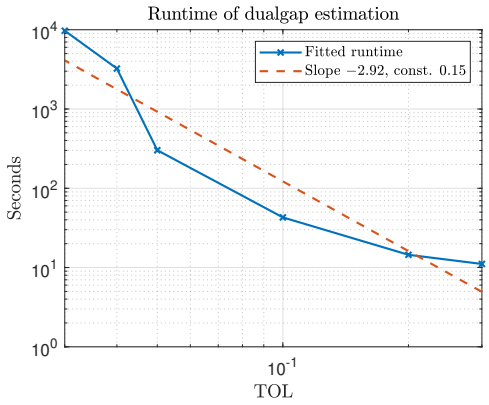
(a) $\lambda = 0.5$



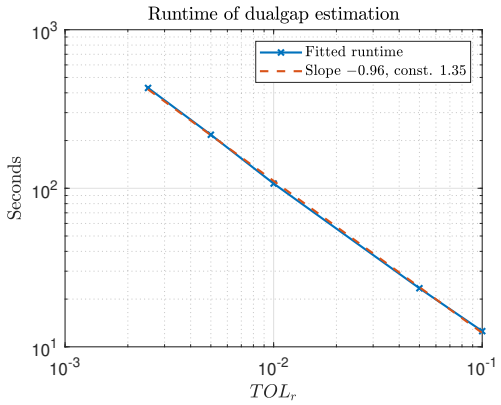
(b) Corresponding controls

Figure 4.8: Example visualization of Algorithm 3 as described in Section 4.3 to estimate the relative duality gap for Problem 8 for a violation probability of $\delta = 0.05$ and a discretization that is optimal w.r.t. Theorem 4.6.2 using TOL $= 0.1$, TOL$_\text{r} = 0.1$. The interpolation procedure for a fixed LM $\lambda = 0.5$ with $\tilde{c} = 1$ in Equation (4.20) is visualized in Figure 4.8a. The resulting optimal controls are displayed and compared in Figure 4.8b. The relative dual gap is $d^\lambda \approx 1.512$.

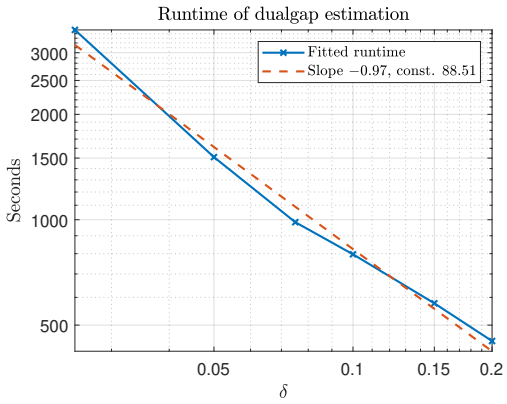
We can check the optimal work behaviour numerically by varying one of the parameters TOL, TOL_r or δ and leaving the rest constant. The results are shown in Figures 4.9a, 4.9b and 4.9c. Practically, the relative dual gap estimation procedure outweighs the computation of one iterative in the SGM, as seen in the example in Figure 4.8 where one iteration of the SGM took approximately 53 seconds in contrast to approximately 909 seconds for the relative dual gap estimation.



(a) TOL vs. Runtime



(b) TOL_r vs Runtime



(c) δ vs. Runtime

Figure 4.9: Runtime analysis of Algorithm 3 w.r.t. to the different parameters TOL, TOL_r and δ. The results show that, practically, our solver has a work rate of $\Gamma = 1$ which aligns with the results from Section 4.2.1 and Figure 4.1.

4.7.3 Algorithm 4: Subgradient method

In this section, we will perform the SGM in Algorithm 4 using all the previous results and comparing different scenarios. Due to the assumptions in Equation (2.33) and (2.34), we need to estimate the constants R and S respectively to get a theoretical bound on the runtime of Algorithm 4. This theoretical bound will be compared to runs using the relative dual gap from Section 4.3 as a stopping criterion, as well as using improved initial values λ_0 found in Section 4.7.1. In the 1D Problems 8 and 9, we can estimate the constant S by using the first subgradient norm, i.e.

$$\|v^{\lambda_k}(0, x_0)\|_2 \leq \|v^{\lambda_0}(0, x_0)\|_2 \quad \text{for all } k \in \mathbb{N}.$$

In general, we can bound the subgradient norm by

$$\|v^{\lambda_k}(0,x_0)\|_2 \in [-\delta, 1-\delta], \quad \text{for all } k \in \mathbb{N},$$

due to the fact that the subgradient is the value of constraint violation in λ_k which can only take values in $[-\delta, 1-\delta]$. For R, we can use the results from Section 4.5, namely $\bar{\beta}_0 = \mathcal{O}(\delta^{-2})$ which results in $\mathrm{CC}_{\mathrm{bisec}}(\bar{\beta}_0)$ and thus

$$\|\lambda^* - \lambda_0\|_2 \leq \|\bar{\beta}_0 - \lambda_0\|_2 =: R.$$

For a given tolerance $\mathrm{TOL}_{\mathrm{SG}} > 0$, we know from Section 2.3.2, Equation (2.39), that to reach $\mathrm{TOL}_{\mathrm{SG}}$-close to the optimal solution, we require at least

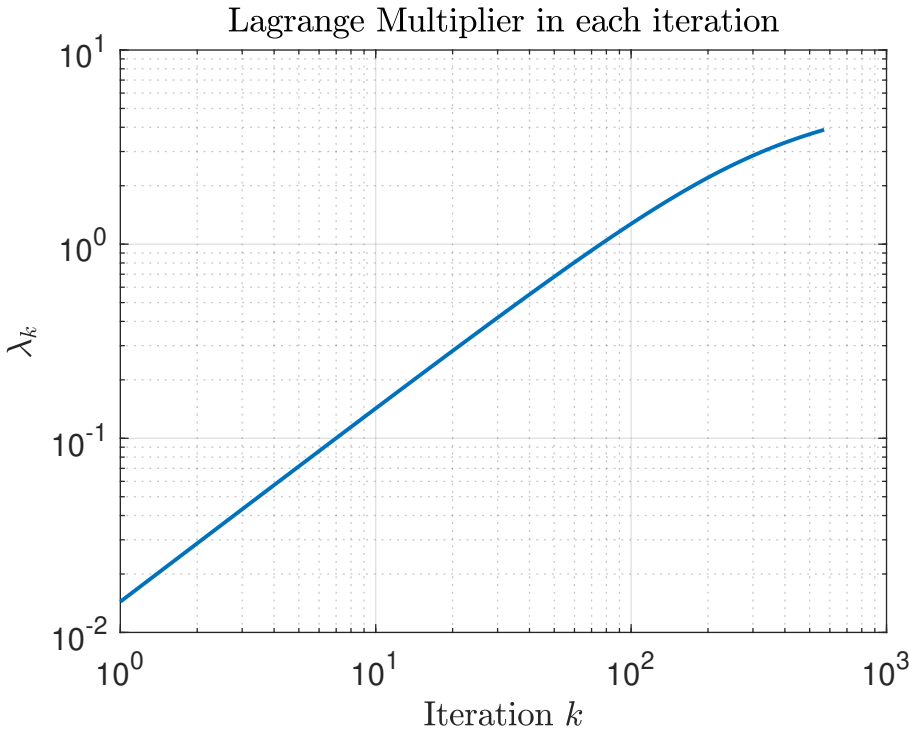
$$K^* = \left\lceil \left(\frac{\|v^{\lambda_0}(0,x_0)\|_2 \cdot \|\bar{\beta}_0 - \lambda_0\|_2}{\mathrm{TOL}_{\mathrm{SG}}} \right)^2 \right\rceil \geq \left(\frac{RS}{\mathrm{TOL}_{\mathrm{SG}}} \right)^2$$

iterations. We could numerically verify the asymptotic behaviour $\mathcal{O}(\mathrm{TOL}_{\mathrm{SG}}^{-2})$ using the theoretical results. This is not necessary due to the fact that we will instead use the relative dual gap in combination with the prescribed relative tolerance $\mathrm{TOL}_{\mathrm{r}} > 0$ of the chance constraint estimation in each iteration as a more informative stopping criterion. In Problem 8 and 9 the absolute value of the subgradient is the chance constraint value minus the violation probability such that we can use this information to stop the algorithm once $|v^{\lambda_k}(0,x_0)| < \mathrm{TOL}_{\mathrm{r}} \cdot \delta$. Let's consider an example run for both examples with an uninformed initial value of $\lambda_0 = 0$ and an aimed violation probability of $\delta = 0.2$:

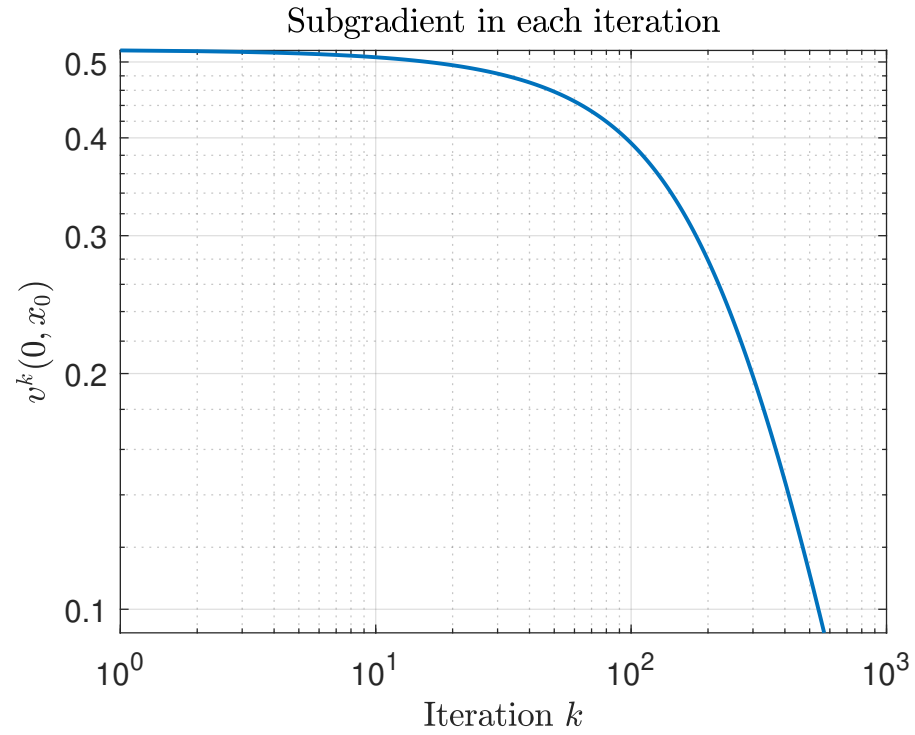
(i) Problem 8, Figure 4.10/4.11:
We aim to find λ^* for a tolerance of $\mathrm{TOL}_{\mathrm{SG}} = 10^{-2}$ and $\mathrm{TOL}_{\mathrm{r}} = 0.25$. We are using an optimal constant step size for the SGM $\tilde{t}_k = R/S\sqrt{K} \approx 0.0278$ for all iterations k with $R = 5$ and $S = 0.6$ as described in Section 2.3.2 which results in a theoretical number of iterations $K^* = 90000$. For the UFD approximation of the HJB equation and the smoothing error we impose an absolute tolerance of $\mathrm{TOL} = 0.25$. One iteration using these tolerances takes approximately 30 seconds resulting in a theoretical bound of around 750 hours. The actual time of the algorithm until $d_k < \mathrm{TOL}_{\mathrm{SG}}$ is around 4 and a half hours with $K = 571$ iterations. The resulting optimal LM is $\lambda^* \approx 3.91$ and 10 sample paths of the optimally controlled SDE using the dual solution with Lagrange Multiplier λ^* is displayed in Figure 4.11b.

(ii) Problem 9, Figure 4.12/4.13:
We aim to find λ^* for a tolerance of $\mathrm{TOL}_{\mathrm{SG}} = 10^{-2}$ and $\mathrm{TOL}_{\mathrm{r}} = 0.05$. We are using an optimal constant step size for the SGM $\tilde{t}_k = R/S\sqrt{K} \approx 0.0164$ for all iterations k with $R = 1$ and $S = 0.78$ as described in Section 2.3.2 which results in a theoretical number of iterations $K^* = 6084$. For the UFD approximation of the HJB equation and the smoothing error we impose an absolute tolerance of $\mathrm{TOL} = 0.1$. One iteration using these tolerances takes approximately 109 seconds resulting in a theoretical bound of around 184 hours. The actual time of the algorithm until $|v^{\lambda_k}(0,x_0)| < \mathrm{TOL}_{\mathrm{r}} \cdot \delta$ is around 10 hours with $K = 342$ iterations, see Figure 4.12b. The resulting optimal LM is $\lambda^* \approx 0.415$ and 10 sample paths of the optimally controlled SDE using the dual solution with Lagrange Multiplier λ^* is displayed in Figure 4.13b.

The high discrepancy between the actual runtime and theoretical upper bound on the runtime can be explained by the fact that the problems we are solving in our examples, i.e. Problem 8 and Problem 9, are relative smooth in comparison to the assumptions and general non-smooth functions considered in the theoretical analysis of the SGM in Section 2.3.2. The theoretical upper bound for the example in Figure 4.10/4.11 is about 166 times larger than the actual runtime. Similarly, for Problem 9 and Figure 4.12/4.13, the theoretical bound on the runtime is about 18 time larger than the actual runtime. This motivates the use of the relative dual gap estimation proposed in Section 4.3. We already discussed the fact that the subroutine of estimating the relative dual gap in each iteration, i.e. Algorithm 3, takes significantly longer than the computations of optimal control and subgradient in each iteration of the SGM itself. It is a trade-off that is beneficial in our case because we are still reducing the time by a great factor w.r.t. the theoretical bound on the runtime. The time to execute the subroutine may be reduced by taking increased tolerances for the HJB and KBE approximations within the subroutine, differing from the tolerances $\mathrm{TOL}, \mathrm{TOL}_\mathrm{r}$ used in the computation of optimal control α^λ and corresponding subgradient $v^\lambda(0, x_0)$. This is an open question for future research. In general, care must be taken that the tolerances are chosen so that the values of the solution of the HJB and KBE are well represented, e.g. if $u(0, x0)$ as in Problem 9 is close to zero, the tolerances must be chosen accordingly small. Further, the runtime can be even more decreased by using Algorithm 5 to compute an initial value $\lambda_0 \geq 0$ that is closer to the desired solution λ^*. We conducted the same procedure as in (ii) Problem 8, Figure 4.10/4.11 using the resulting initial LM $\lambda_0 = 0.752$ from Section 4.7.1 which resulted in a significant decrease in iterations needed to satisfy the condition $d_k < \mathrm{TOL}_\mathrm{SG}$. Namely, only $K = 14$ iterations with a total runtime of approximately 6 minutes is needed. This is a decrease in runtime with a factor of 7500 on the theoretical upper bound and a factor of 45 on the actual runtime with an initial LM $\lambda_0 = 0$. We may accelerate the procedure by varying the time steps and starting with a constant, possibly non-optimal, step size and use diminishing square summable but not summable time steps, see Corollary 2.3.1, to guarantee convergence.

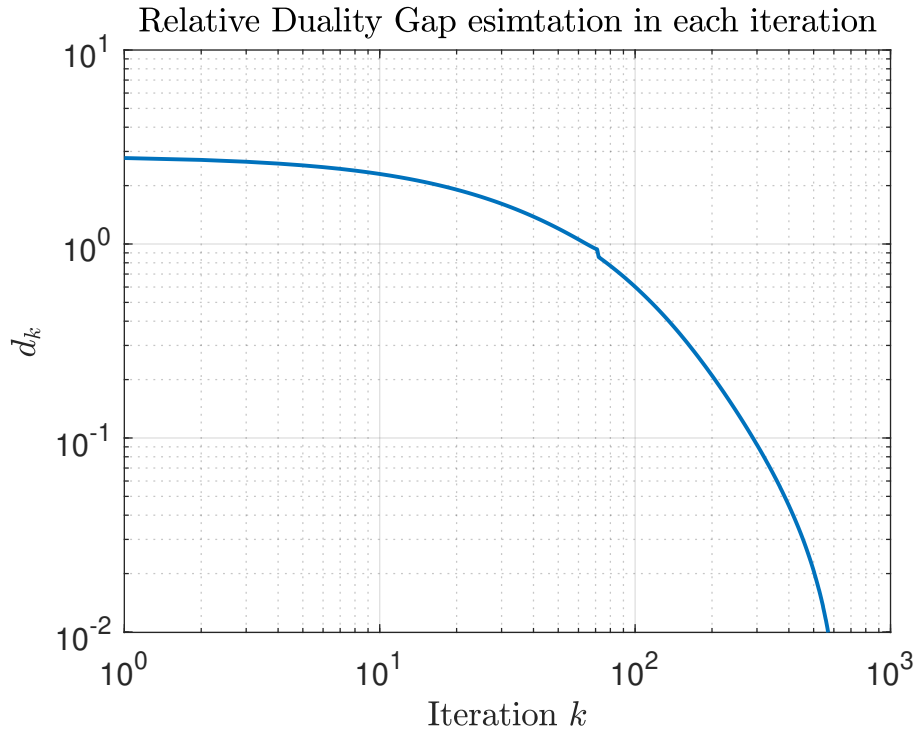


(a) Lagrange Multiplier in each iteration

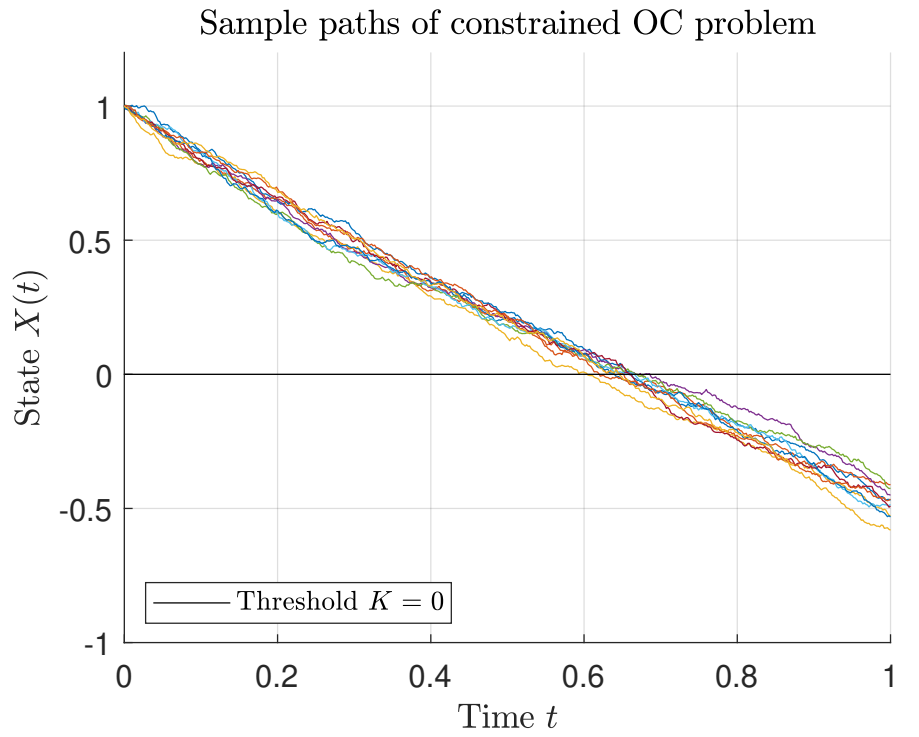


(b) Subgradient value in each iteration

Figure 4.10: An example run of Algorithm 4 for Problem 8.

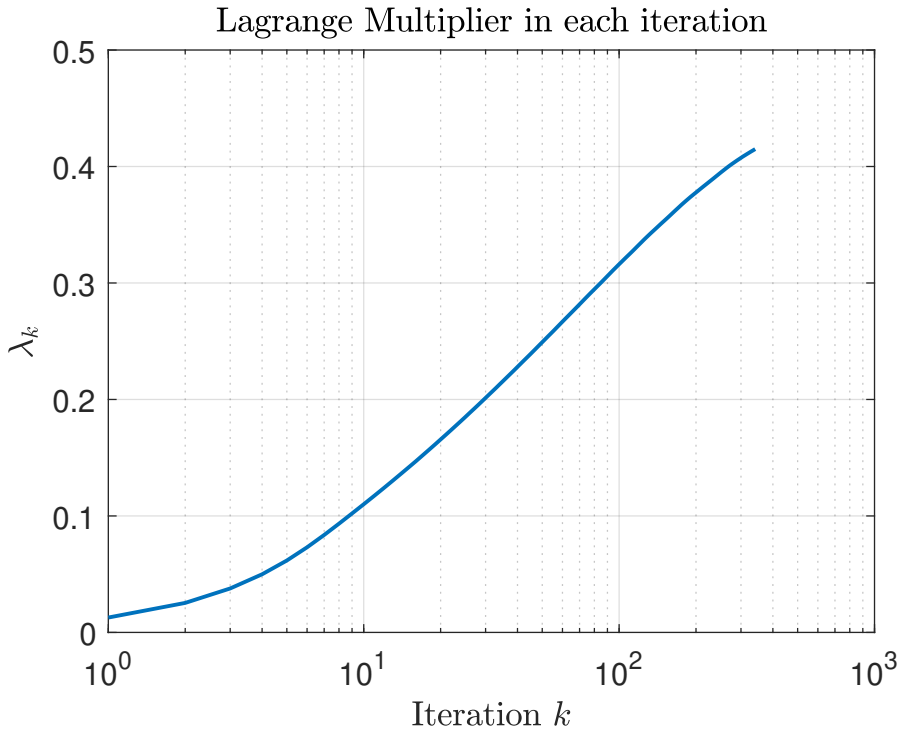


(a) Relative dual gap in each iteration

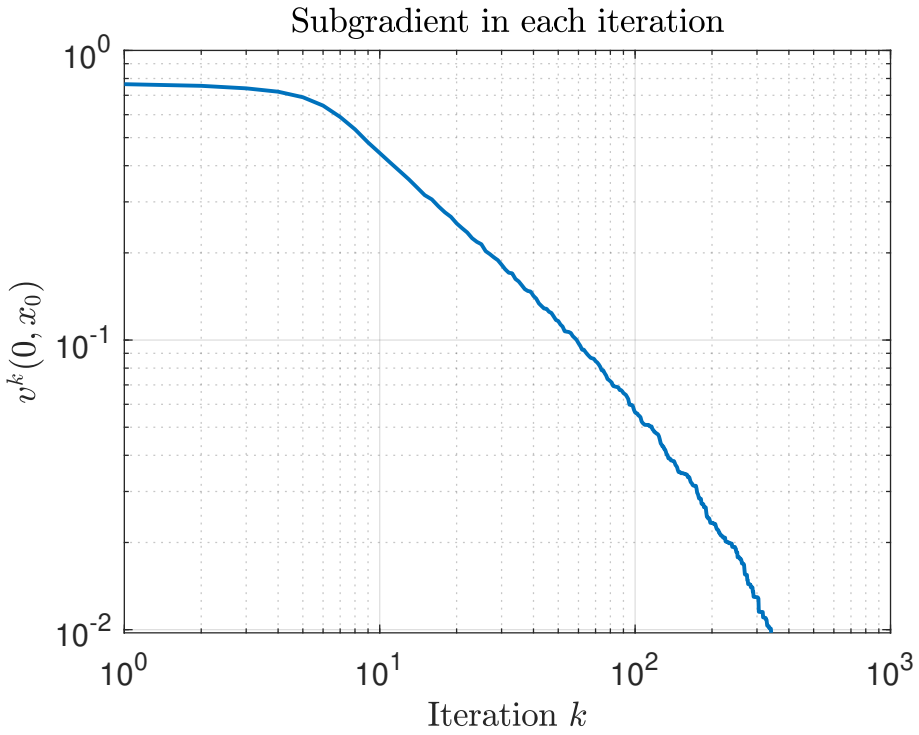


(b) Sample paths of opt. controlled SDE

Figure 4.11: An example run of Algorithm 4 for Problem 8.

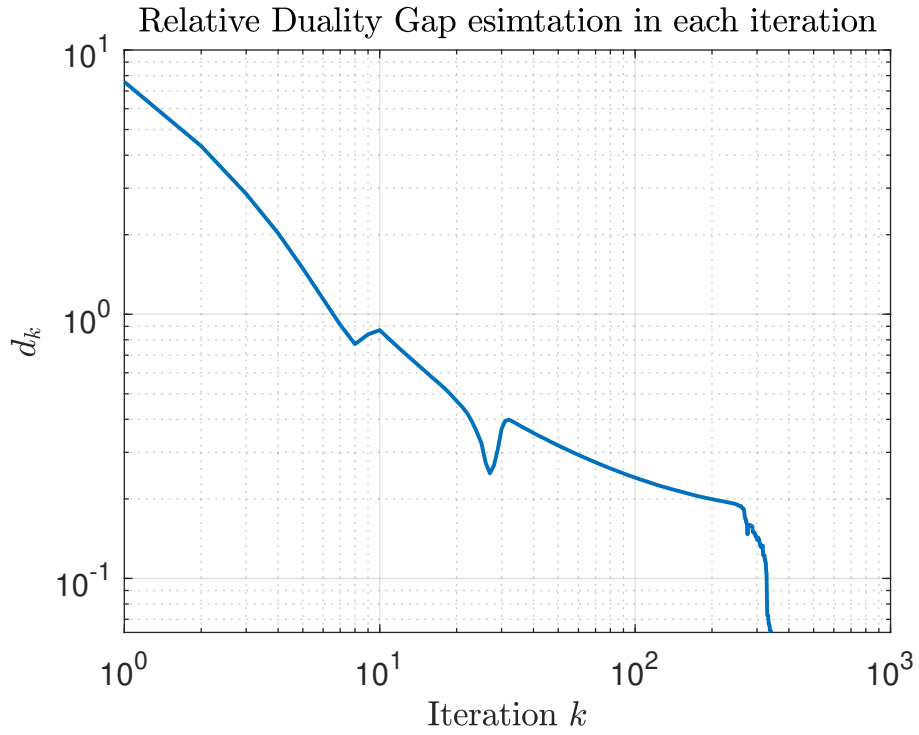


(a) Lagrange Multiplier in each iteration

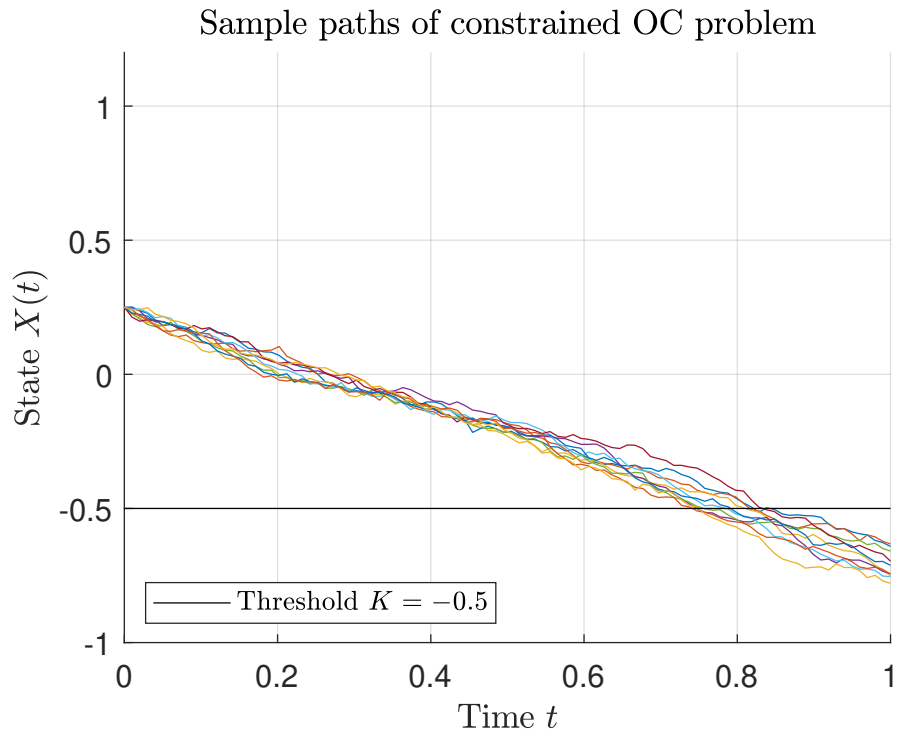


(b) Subgradient value in each iteration

Figure 4.12: An example run of Algorithm 4 for Problem 9.



(a) Relative dual gap in each iteration



(b) Sample paths of opt. controlled SDE

Figure 4.13: An example run of Algorithm 4 for Problem 9.

4.7.4 Cost reduction of optimal solution

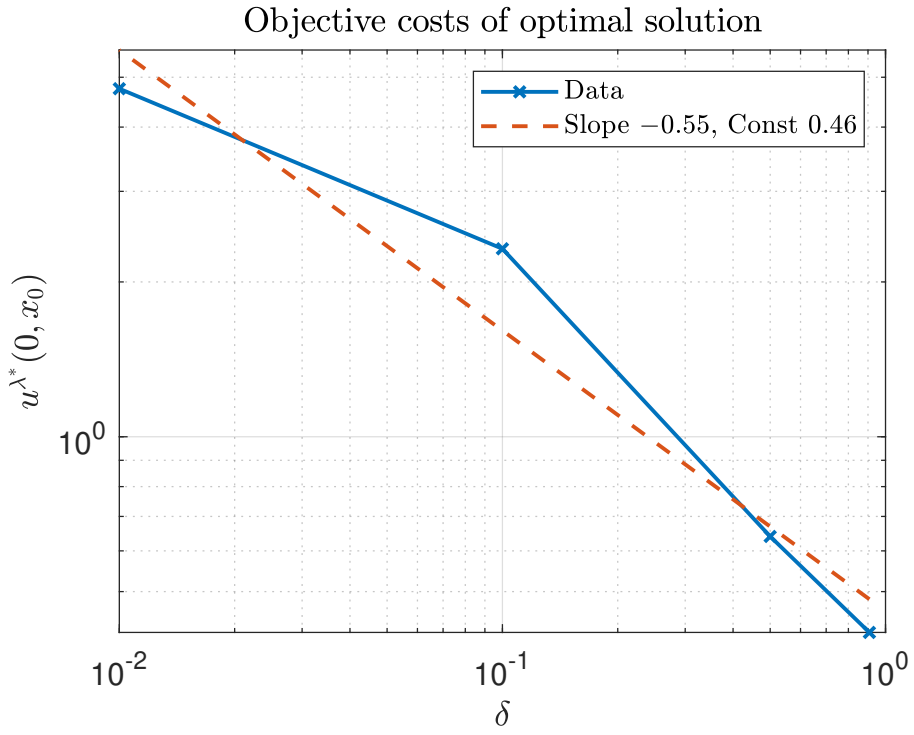
In this section, we investigate the reduction in the objective's costs for different violation probabilities $0 < \delta \ll 1$. This is motivated by almost sure constraints of the form

$$\mathbb{P}\left(G(X_T) > K\right) = 0 \quad \text{or} \quad \mathbb{P}\left(G(X_T) \leq K\right) \quad \mathbb{P} - \text{a.s}$$

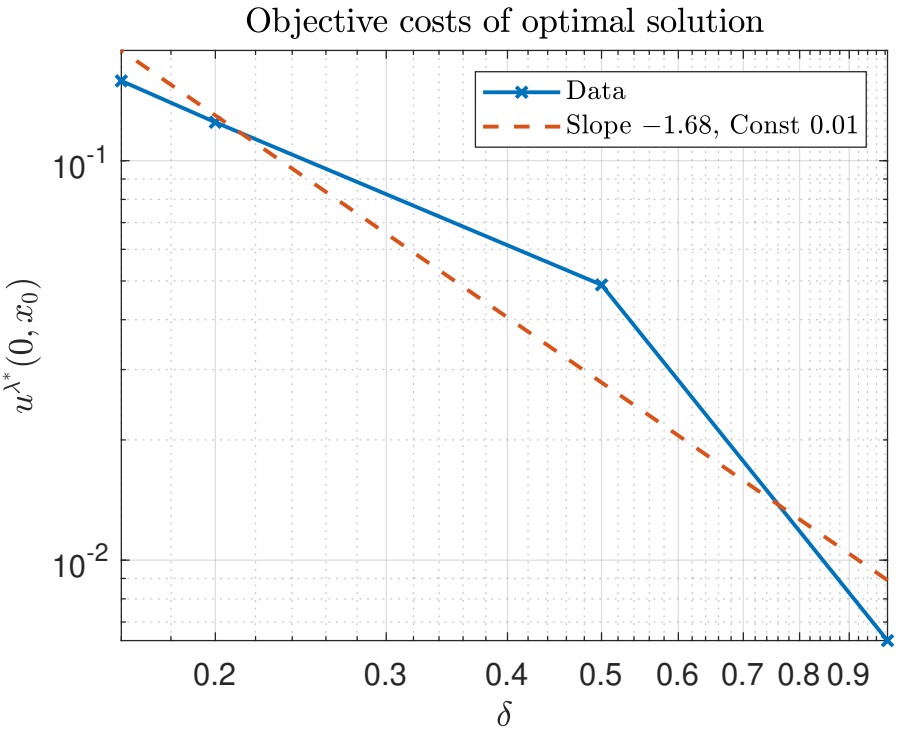
which, if the underlying probability density function is defined on all $\mathbb{R}$, is very expensive to enforce numerically. Analytically, these chance constraints have a very small probability in any case. Using a chance constraint, i.e. a relaxed version of the almost sure constraint above, we aim to reduce the costs of the solution. For this we would like to see the asymptotic behaviour of the optimal costs as the violation probability δ approaches zero. For Problem 8, we conducted the optimization procedure from Algorithm 4 to find λ^* for different violation probabilities δ, computed the value function $u^{\lambda^*}(t, x)$ and evaluated it at the initial values $(0, x_0)$ of the controlled SDE to get the objective costs for the different violation probabilities. A plot of the results is shown in Figure 4.14a. The approximate asymptotic behaviour with respect to the violation probability δ is $\mathcal{O}(\delta^{-0.55})$. The percentage increase in the objective costs for this example when going from a violation probability of $\delta = 10^{-1}$ to $\delta = 10^{-2}$ is approximately 231%. Similar, we conducted the experiment for Problem 9 in which an almost sure constraint with the usage of the parameters of Table 4.1 is not possible to achieve. Numerically, the constraint of Problem 9 was bounded below and above (by the unconstrained problem's constraint probability) by

$$0.1506 \leq \mathbb{P}\left(X_T > -0.5\right) \leq 0.99.$$

This results in a greater asymptotic behaviour of approximately $\mathcal{O}(\delta^{-1.68})$ towards the boundary of the constraint. The experiment results are shown in Figure 4.14b. The cost increase for going from a violation probability of $\delta \approx 0.99$ (unconstrained) to $\delta = 0.2$ is approximately 2500% which is inflated by the fact that the objective costs for the unconstrained Problem 9 is near zero. Regardless of this, both examples show that using chance constraints as a relaxation of a.s. constraints yields a significant decrease in the objective costs, making it a good alternative in practice.



(a) Problem 8



(b) Problem 9

Figure 4.14: A comparison of the different resulting costs in the objective for the optimally controlled dual problem solution for different violation probabilities δ. The resulting asymptotic behaviour of the costs for Problem 8 in Figure 4.14a is approximately of order $\mathcal{O}(\delta^{-0.55})$. The resulting asymptotic behaviour of the costs for Problem 9 in Figure 4.14b is approximately of order $\mathcal{O}(\delta^{-1.68})$.

Chapter 5

Conclusion

In this thesis, we developed a solution methodology for a class of chance-constrained optimal control problems governed by controlled stochastic differential equations in 1D. The core of our approach involved formulating a relaxed version of the constraint $\mathbb{P}(G(X_T) > K) \leq \delta$ for some violation probability $0 < \delta \ll 1$ and deriving a dual problem, which we then tackled numerically using a subgradient-based optimization method. To achieve stable numerical solvers, we used the theory from Itô calculus to reduce the problem to solutions of PDEs, namely the Hamilton-Jacobi-Bellman equation and the Kolmogorov Backward Equation. We constructed PDE solvers and proposed an efficient methodology for computing subgradients, enabling the use of more advanced non-smooth optimization techniques. Additionally, we introduced a novel approach for estimating the relative dual gap in each iteration of the subgradient method, allowing for a more informed stopping criterion. Furthermore, we proposed a bisection-based approach for initializing Lagrange multipliers for the case of a single scalar terminal state chance constraint, which can accelerate convergence due to the convergence assumptions in the subgradient methods. Finally, we analysed the sources of error in our numerical approach and established theoretical results that provide guidance on optimal parameter choices, ensuring efficient asymptotic computational work and stable execution of the proposed procedure. These contributions collectively advance the numerical treatment of chance-constrained optimal control problems and lay the groundwork for further research in this area.

5.1 Outlook

While this thesis provides a framework for solving chance-constrained optimal control problems governed by controlled stochastic differential equations, several open questions and potential extensions remain. These avenues for future research span theoretical analysis, methodological improvements, and practical applications.

- PDE solvers: One important aspect to explore is the impact of discontinuous boundary conditions, such as

  $$u(T, x) = \mathbb{1}_{\{G(x)>K\}},$$

  on the numerical stability and accuracy of the solution without the use of smoothing techniques. Additionally, a deeper analysis is required to understand the effects of singular control functions $\alpha(t, x)$ on the Kolmogorov-Backward Equation (KBE), particularly in cases where excessive singularity leads to numerical instabilities, despite the theoretical range of the solution to the KBE being between 0 and 1.

- Chance constraint types: Extending the framework to handle multiple constraints is a natural next step. This would introduce additional Lagrange multipliers. Standard numerical root-finding techniques to generalize the special proposal of Section 4.5 could be employed instead of simple bisection methods which works only with a single scalar chance constraint. Another promising direction is the formulation of chance constraints over the entire time horizon $t \in [0, T]$, which would require time-dependent Lagrange multipliers. A practical approach might be to approximate these multiplier functions using simple function representations, as suggested in the work [33].

- Applications: The methodology developed in this work could find practical applications in fields such as finance, logistics, power systems and more. An example application might be optimal power procurement for wireless network communication towers [33]. Investigating such real-world problems could provide valuable insights into the practical viability and scalability of the proposed approach.

- Computational Aspects: Finally, a key consideration for future work is the computational cost of estimating the relative dual gap. The auxiliary problem required for this estimation often exceeds the cost of a single iteration by a factor of five or more. Developing more efficient estimation techniques or alternative stopping criteria could significantly enhance the overall efficiency of the optimization process.

- Multiple dimensions: Increasing the dimension of the controlled SDE incorporates new numerical challenges such as the *curse of dimensionality* for solving the underlying PDEs in the algorithms. In higher dimensional problems, the computational complexity of our approach increases significantly. This may force us to consider statistical methods like stochastic-subgradient methods or Monte Carlo techniques for estimating quantities of interest.

- High Performance Computing: When increasing the dimensions it might come in handy to run simulations on clusters, i.e. investigate the possibility of including HPC, not only to solve the PDEs but also to improve the estimation of the constants in the optimal computational work theorems.

These open questions and extensions highlight the potential for further research in chance-constrained optimal control, bridging gaps between theory, computation, and practical applications.

Appendix A

Proofs

A.1 Proof of Example 2.1.1

1. The expected value, for a fixed t, is given by

$$\begin{aligned}\mathbb{E}[X(t)] &= \mathbb{E}\left[x_\infty + e^{-\theta t}(x_0 - x_\infty) + \sigma \int_0^t e^{-\theta(t-s)} dW(s)\right] \\ &= x_\infty + e^{-\theta t}(x_0 - x_\infty) + \sigma \cdot \underbrace{\mathbb{E}\left[\int_0^t e^{-\theta(t-s)} dW(s)\right]}_{=0} \\ &= x_\infty + e^{-\theta t}(x_0 - x_\infty).\end{aligned}$$

   The variance, for fixed t, can be written as

$$\begin{aligned}\mathrm{Var}[X(t)] &= \mathrm{Var}\left[x_\infty + e^{-\theta t}(x_0 - x_\infty) + \sigma \int_0^t e^{-\theta(t-s)} dW(s)\right] \\ &= \sigma^2 \cdot \mathrm{Var}\left[\int_0^t e^{-\theta(t-s)} dW(s)\right] \\ &= \sigma^2 \cdot \left(\underbrace{\mathbb{E}\left[\left(\int_0^t e^{-\theta(t-s)} dW(s)\right)^2\right]}_{\text{use Ito isometry}} - \underbrace{\mathbb{E}\left[\int_0^t e^{-\theta(t-s)} dW(s)\right]^2}_{=0} \right) \\ &= \sigma^2 \cdot \int_0^t \mathbb{E}[e^{-2\theta(t-s)}] ds = \sigma^2 \cdot \int_0^t e^{-2\theta(t-s)} ds \\ &= \frac{\sigma^2}{2\theta}\left[e^{-2\theta(t-s)}\right]_{s=0}^{s=t} = \frac{\sigma^2}{2\theta}\left(1 - e^{-2\theta t}\right)\end{aligned}$$

2. Taking the limit $t \to \infty$ for the expected value and variance results in

$$\begin{aligned}\lim_{t\to\infty} \mathbb{E}[X(t)] &= \lim_{t\to\infty} \left(x_\infty + e^{-\theta t}(x_0 - x_\infty)\right) = x_\infty, \\ \lim_{t\to\infty} \mathrm{Var}[X(t)] &= \lim_{t\to\infty} \frac{\sigma^2}{2\theta}(1 - e^{-2\theta t}) = \frac{\sigma^2}{2\theta}.\end{aligned}$$

3. The covariance can be written as

$$\mathrm{Cov}[X(t), X(t+\tau)] = \underbrace{\mathbb{E}[X(t) \cdot X(t+\tau)]}_{\text{(I)}} - \underbrace{\mathbb{E}[X(t)]\mathbb{E}[X(t+\tau)]}_{\text{(II)}}.$$

We will first look at (I):

$$\begin{aligned}
\mathbb{E}[X(t)\cdot X(t+\tau)] = \mathbb{E}\Bigg[& \left(x_\infty + e^{-\theta t}(x_0 - x_\infty) + \sigma \cdot \int_0^t e^{-a\theta(t-s)} dW(s)\right) \\
& \cdot \left(x_\infty + e^{-\theta(t+\tau)}(x_0 - x_\infty) + \sigma \cdot \int_0^{t+\tau} e^{-\theta(t+\tau-s)} dW(s)\right)\Bigg] \\
& = \underbrace{x_\infty^2 + x_\infty e^{-\theta(t+\tau)}(x_0 - x_\infty) + x_\infty e^{-\theta t}(x_0 - x_\infty) + (x_0 - x_\infty)^2 e^{-\theta(2t+\tau)}}_{=:c(t)} \\
& + x_\infty \cdot \sigma \cdot \underbrace{\mathbb{E}[\int_0^{t+\tau} e^{-\theta(t+\tau-s)} dW(s)]}_{=0} \\
& + e^{-\theta t}(x_0 - x_\infty) \cdot \sigma \cdot \underbrace{\mathbb{E}[\int_0^{t+\tau} e^{-\theta(t+\tau-s)} dW(s)]}_{=0} \\
& + x_\infty \cdot \sigma \cdot \underbrace{\mathbb{E}[\int_0^{t} e^{-\theta(t-s)} dW(s)]}_{=0} \\
& + e^{-\theta(t+\tau)}(x_0 - x_\infty) \cdot \sigma \cdot \underbrace{\mathbb{E}[\int_0^{t} e^{-\theta(t-s)} dW(s)]}_{=0} \\
& + \sigma^2 \cdot \mathbb{E}[\int_0^t e^{-\theta(t-s)} dW(s) \cdot \int_0^{t+\tau} e^{-\theta(t+\tau-s)} dW(s)],
\end{aligned}$$

such that we end up with the term

$$\mathbb{E}[X(t)\cdot X(t+\tau)] = c(t) + \sigma^2 \cdot \mathbb{E}[\int_0^t e^{-\theta(t-s)} dW(s) \cdot \int_0^{t+\tau} e^{-\theta(t+\tau-s)} dW(s)]. \quad \text{(A.1)}$$

Assuming that

$$\int_0^{t+\tau} e^{-\theta(t+\tau-s)} dW(s) = \int_0^t e^{-\theta(t+\tau-s)} dW(s) + \int_t^{t+\tau} e^{-\theta(t+\tau-s)} dW(s)$$

holds, we can rewrite the expectation in (A.1) as

$$\begin{aligned}
& \mathbb{E}[\int_0^t e^{-\theta(t-s)} dW(s) \cdot \int_0^{t+\tau} e^{-\theta(t+\tau-s)} dW(s)] \\
& = \mathbb{E}[\int_0^t e^{-\theta(t-s)} dW(s) \cdot \left(\int_0^t e^{-\theta(t+\tau-s)} dW(s) + \int_t^{t+\tau} e^{-\theta(t+\tau-s)} dW(s)\right)] \\
& = \mathbb{E}[\int_0^t e^{-\theta(t-s)} dW(s) \cdot \int_0^t e^{-\theta(t+\tau-s)} dW(s) + \int_0^t e^{-\theta(t-s)} dW(s) \cdot \int_t^{t+\tau} e^{-\theta(t+\tau-s)} dW(s)] \\
& = \underbrace{\mathbb{E}[\int_0^t e^{-\theta(t-s)} dW(s) \cdot \int_0^t e^{-\theta(t+\tau-s)} dW(s)]}_{\text{use Ito isometry}} + \mathbb{E}[\underbrace{\int_0^t e^{-\theta(t-s)} dW(s)}_{\text{depends on time } [0,t)} \cdot \underbrace{\int_t^{t+\tau} e^{-\theta(t+\tau-s)} dW(s)}_{\text{depends on time } [t,t+\tau]}] \\
& = \int_0^t \mathbb{E}[e^{-\theta(t-s)} e^{-\theta(t+\tau-s)}] ds + \underbrace{\mathbb{E}[\int_0^t e^{-\theta(t-s)} dW(s)]}_{=0} \cdot \underbrace{\mathbb{E}[\int_t^{t+\tau} e^{-\theta(t+\tau-s)} dW(s)]}_{=0} \\
& = \int_0^t e^{-\theta(2t-2s+\tau)} dW(s) = \frac{1}{2\theta}\left[e^{-\theta(2t-2s+\tau)}\right]_{s=0}^{s=t} \\
& = \frac{1}{2\theta}\left(e^{-\theta\tau} - e^{-\theta(2t+\tau)}\right).
\end{aligned}$$

Now we will look at (II)

$$\begin{aligned}\mathbb{E}[X(t)] \cdot \mathbb{E}[X(t+\tau)] &= \left(x_\infty + e^{-\theta t}(x_0 - x_\infty)\right) \cdot \left(x_\infty + e^{-\theta(t+\tau)}(x_0 - x_\infty)\right) \\ &= x_\infty^2 + x_\infty e^{-\theta(t+\tau)}(x_0 - x_\infty) + x_\infty e^{-\theta t}(x_0 - x_\infty) + (x_0 - x_\infty)^2 e^{-\theta(2t+\tau)} \\ &= c(t).\end{aligned}$$

Together, we get

$$\begin{aligned}\mathrm{Cov}[X(t), X(t+\tau)] &= \mathbb{E}[X(t)X(t+\tau)] - \mathbb{E}[X(t)] \cdot \mathbb{E}[X(t+\tau)] \\ &= c(t) + \frac{\sigma^2}{2\theta}\left(e^{-\theta\tau} - e^{-\theta(2t+\tau)}\right) - c(t) \\ &= \frac{\sigma^2}{2\theta}\left(e^{-\theta\tau} - e^{-\theta(2t+\tau)}\right).\end{aligned}$$

4. For a fixed t, the limit, as $\tau \to \infty$, can be calculated as

$$\lim_{\tau\to\infty} \mathrm{Cov}[X(t), X(t+\tau)] = \lim_{\tau\to\infty} \frac{\sigma^2}{2\theta}\left(e^{-\theta\tau} - e^{-\theta(2t+\tau)}\right) = 0.$$

A.2 Proof of Theorem 2.2.1

Step 1: Take the optimal control α^* from t to T to obtain

$$\begin{aligned}u(t,x) &= \min_{\alpha\in\mathcal{A}} \mathbb{E}\left[\int_t^T h(s, X_s, \alpha_s)ds + g(X_T)\right] \\ &= \mathbb{E}\left[\int_t^{t+\varepsilon} h(s, X_s, \alpha_s^*)ds\right] + \mathbb{E}\left[\int_{t+\varepsilon}^T h(s, X_s, \alpha_s^*)ds + g(X_T)\right] \\ &= \mathbb{E}[\int_t^{t+\varepsilon} h(s, X_s, \alpha_s^*)ds] + \mathbb{E}\left[\mathbb{E}\left[\int_{t+\varepsilon}^T h(s, X_s, \alpha_s^*)ds + g(X_T)\,\middle|\, X(t+\varepsilon)\right]\right] \\ &\overset{(1)}{\geq} \mathbb{E}\left[\int_t^{t+\varepsilon} h(s, X_s, \alpha_s^*)ds\right] + \mathbb{E}\left[u(t+\varepsilon, X(t+\varepsilon))\right] \\ &\overset{(2)}{\geq} \min_{\alpha\in\tilde{\mathcal{A}}} \mathbb{E}\left[\int_t^{t+\varepsilon} h(s, X_s, \alpha_s)ds + u(t+\varepsilon, X(t+\varepsilon))\right],\end{aligned}$$

where $\tilde{\mathcal{A}} := \{\alpha : [t, t+\varepsilon] \times \mathbb{R} \to A\}$ and (1) and (2) is due to the fact that the minimum from t to $t+\varepsilon$ or from $t+\varepsilon$ to T might differ from the minimum over all $[t, T]$, i.e. for example in (1) we have

$$\begin{aligned}&\mathbb{E}\left[\int_{t+\varepsilon}^T h(s, X_s, \alpha_s^*)ds + g(X_T)\,\middle|\, X(t+\varepsilon)\right] \\ &\geq \min_{\alpha\in\bar{\mathcal{A}}} \mathbb{E}\left[\int_{t+\varepsilon}^T h(s, X_s, \alpha_s^*)ds + g(X_T)\,\middle|\, X(t+\varepsilon)\right] \\ &= u(t+\varepsilon, X(t+\varepsilon)),\end{aligned}$$

where $\bar{\mathcal{A}} := \{\alpha : [t+\varepsilon, T] \times \mathbb{R} \to A\}$.

Step 2: Choose a control α^+ that is arbitrary from t to $t+\varepsilon$. Given the value $X(t+\varepsilon)$, choose the optimal control α^* from $t+\varepsilon$ to T. Denote this control by $\alpha' = (\alpha^+, \alpha^*)$. Definition (2.20) shows

$$\begin{aligned} u(t,x) \leq C_{t,x}(\alpha') &= \mathbb{E}\left[\int_t^T h(s, X_s, \alpha'_s)ds + g(X_T)\right] \\ &= \mathbb{E}\left[\int_t^{t+\varepsilon} h(s, X_s, \alpha_s^+)ds\right] + \mathbb{E}\left[\int_{t+\varepsilon}^T h(s, X_s, \alpha_s^*)ds + g(X_T)\right] \\ &= \mathbb{E}\left[\int_t^{t+\varepsilon} h(s, X_s, \alpha_s^+)ds\right] \\ &\quad + \mathbb{E}\left[\mathbb{E}\left[\int_{t+\varepsilon}^T h(s, X_s, \alpha_s^*)ds + g(X_T)\,\middle|\, X(t+\varepsilon)\right]\right] \\ &\overset{(3)}{=} \mathbb{E}\left[\int_t^{t+\varepsilon} h(s, X_s, \alpha_s^+)ds\right] + \mathbb{E}\left[u(t+\varepsilon, X(t+\varepsilon))\right], \end{aligned}$$

where (3) is justified due to α^* being the optimal control from $t+\varepsilon$ to T, i.e.

$$\begin{aligned} &\mathbb{E}\left[\int_{t+\varepsilon}^T h(s, X_s, \alpha_s^*)ds + g(X_T)\,\middle|\, X(t+\varepsilon)\right] \\ = \min_{\alpha\in\tilde{\mathcal{A}}} &\mathbb{E}\left[\int_{t+\varepsilon}^T h(s, X_s, \alpha_s)ds + g(X_T)\,\middle|\, X(t+\varepsilon)\right] = u(t+\varepsilon, X(t+\varepsilon)). \end{aligned}$$

Taking the minimum over all controls α^+ yields

$$u(t,x) \leq \min_{\alpha^+\in\tilde{\mathcal{A}}} \mathbb{E}\left[\int_t^{t+\varepsilon} h(s, X_s, \alpha_s^+)ds + u(t+\varepsilon, X(t+\varepsilon))\right].$$

A.3 Proof of Theorem 2.2.2

The proof is then divided into two steps:

1. Show that the optimal control α^* yields the equation

$$\partial_t u(t,x) + a^*(t,x)\partial_x u(t,x) + \frac{b^*(t,x)^2}{2}\partial_{xx}u(t,x) + h^*(t,x) = 0,$$

for arbitrary (t,x), where $a^*(t,x) = a(t,x,\alpha^*(t,x))$, $b^*(t,x) = b(t,x,\alpha^*(t,x))$ and $h^*(t,x) = h(t,x,\alpha^*(t,x))$.

2. Show that an arbitrary control α^+ implies the inequality

$$\partial_t u(t,x) + a^+(t,x)\partial_x u(t,x) + \frac{b^+(t,x)^2}{2}\partial_{xx}u(t,x) + h^+(t,x) \geq 0,$$

where $a^+(t,x) = a(t,x,\alpha^+(t,x))$, $b^+(t,x) = b(t,x,\alpha^+(t,x))$ and $h^+(t,x) = h(t,x,\alpha^+(t,x))$.

Both steps together imply Theorem 2.2.2.

Step 1: Choose α^* to be the optimal control. Then, by the dynamic programming principle from Theorem 2.2.1,

$$u(t, X(t)) = \mathbb{E}\left[\int_t^{t+\varepsilon} h(s, X(s), \alpha^*(s, X(s)))ds + u(t+\varepsilon, X(t+\varepsilon))\right].$$

Rearranging and dividing by $0 < \varepsilon < T - t$, we get

$$\begin{aligned} 0 &= \frac{1}{\varepsilon} \cdot \mathbb{E}\left[\int_t^{t+\varepsilon} h(s, X_s, \alpha_s^*) ds + (u(t+\varepsilon, X(t+\varepsilon)) - u(t,x))\right] \\ \iff \quad 0 &= \frac{1}{\varepsilon} \cdot \mathbb{E}\left[\int_t^{t+\varepsilon} h(s, X_s, \alpha_s^*) ds + \int_t^{t+\varepsilon} du(s, X(s))\right] \end{aligned}$$

When taking the limit $\varepsilon \to 0$, by continuity and conditioning on $X(t) = x$, we get

$$-h(t, x, \alpha^*(t,x))dt = \mathbb{E}\left[du(t, X(t)) \mid X(t) = x\right],$$

where the latter can be simplified using Itô's lemma 2.1.4, i.e.

$$\begin{aligned} &\mathbb{E}[du(t, X(t)) \mid X(t) = x] \\ &\quad = \mathbb{E}\left[\left(\partial_t u(t,x) + a^*(t,x)\partial_x u(t,x) + \frac{b^*(t,x)^2}{2}\partial_{xx} u(t,x)\right) \;\middle|\; X(t) = x\right]. \end{aligned}$$

Together, we have

$$0 = \partial_t u(t,x) + a^*(t,x)\partial_x u(t,x) + \frac{b^*(t,x)^2}{2}\partial_{xx} u(t,x) + h(t,x,\alpha^*(t,x))$$

The definition in Equation (2.20) shows

$$u(T,x) = g(x).$$

Step 2: Choose the control function α^+ to be arbitrary on the time interval $[t, t+\varepsilon]$. The function u satisfies

$$u(t,x) \leq \mathbb{E}\left[\int_t^{t+\varepsilon} h(s, X_s, \alpha_s^+) ds + u(t+\varepsilon, X(t+\varepsilon))\right].$$

Rearranging and dividing by $0 < \varepsilon < T - t$ yields

$$\begin{aligned} 0 &\leq \frac{1}{\varepsilon}\, \mathbb{E}\left[\int_t^{t+\varepsilon} h(s, X_s, \alpha_s^+) ds + u(t+\varepsilon, X(t+\varepsilon)) - u(t,x)\right] \\ \iff 0 &\leq \frac{1}{\varepsilon}\, \mathbb{E}\left[\int_t^{t+\varepsilon} h(s, X_s, \alpha_s^+) ds + \int_t^{t+\varepsilon} du(s, X(s))\right]. \end{aligned}$$

Taking the limit $\varepsilon \to 0$ and conditioning on $X(t) = x$ further yields

$$-h(t,x,\alpha^+(t,x))dt \leq \mathbb{E}[du(t, X(t)) \mid X(t) = x]$$

We know that for any given α^+, by Itô's formula,

$$\begin{aligned} &\mathbb{E}[du(t, X(t)) \mid X(t) = x] \\ &\quad = \mathbb{E}\left[\left(\partial_t u(t,x) + a^+(t,x)\partial_x u(t,x) + \frac{b^+(t,x)^2}{2}\partial_{xx} u(t,x\right) dt \;\middle|\; X(t) = x\right]. \end{aligned}$$

Therefore, for any arbitrary control α^+,

$$0 \leq \partial_t u(t,x) + a^+(t,x)\partial_x u(t,x) + \frac{b^+(t,x)^2}{2}\partial_{xx} u(t,x) + h(t,x,\alpha^+(t,x)),$$

which proves the second step.

Together both steps imply the result.

A.4 Proof of Theorem 3.3.1

Consider the Lagrangian

$$\mathcal{L}(N_1, M_1, \lambda_1, \lambda_2) = N_1 \cdot M_1 + \lambda_1 \left(\tilde{c}_1 M_1^{-1} - \mathrm{TOL}\right) + \lambda_2 \left(c_3 M_1 + \bar{c}_3 M_1^2 - N_1\right),$$

where

$$c_3 = \frac{T\|a\|_\infty}{|d-c|} > 0, \quad \bar{c}_3 = \frac{\sigma^2 T}{|d-c|^2} > 0$$

are constants. The partial derivatives and first order conditions for extrema are

$$\partial_{N_1}\mathcal{L} = M_1 - \lambda_2 = 0 \tag{A.2}$$
$$\partial_{M_1}\mathcal{L} = N_1 - \lambda_1 \tilde{c}_1 M_1^{-2} + \lambda_2 c_3 + 2\lambda_2 \bar{c}_3 M_1 = 0 \tag{A.3}$$
$$\partial_{\lambda_1}\mathcal{L} = \tilde{c}_1 \cdot M_1^{-1} - \mathrm{TOL} = 0 \tag{A.4}$$
$$\partial_{\lambda_2}\mathcal{L} = c_3 M_1 + \bar{c}_3 M_1^2 - N_1 = 0 \tag{A.5}$$

Equation (A.4) yields

$$M_1^* = \frac{\tilde{c}_1}{\mathrm{TOL}}.$$

Substituting this into Equation (A.5), yields

$$N_1^* = \frac{c_3 \tilde{c}_1}{\mathrm{TOL}} + \frac{\bar{c}_3 \tilde{c}_1^2}{\mathrm{TOL}^2}$$

The optimal work load can be computed using N_1^* and M_1^*

$$\mathcal{W}^* = N_1^* \cdot M_1^* = \left(\frac{c_3 \tilde{c}_1}{\mathrm{TOL}} + \frac{\bar{c}_3 \tilde{c}_1^2}{\mathrm{TOL}^2}\right) \frac{\tilde{c}_1}{\mathrm{TOL}} = \mathcal{O}(\mathrm{TOL}^{-3}).$$

A.5 Proof of Theorem 4.6.2

Consider the Lagrangian

$$\begin{aligned}\mathcal{L}(N_1, M_1, N_2, M_2, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = {} & N_1 \cdot M_1 + N_2 \cdot M_2^\Gamma \\ & + \lambda_1 \left(\tilde{c}_1 \cdot M_1^{-1} - \mathrm{TOL}\right) \\ & + \lambda_2 \left(c_3 M_1 + \bar{c}_3 M_1^2 - N_1\right) \\ & + \lambda_3 \left(c_2 N_2^{-2} + \bar{c}_2 \cdot M_2^{-2} - \mathrm{TOL}_r \cdot \delta\right) \\ & + \lambda_4 \left(c_\gamma \cdot \gamma^2 - \mathrm{TOL}\right)\end{aligned}$$

with partial derivatives and first order conditions for extrema

$$\partial_{N_1}\mathcal{L} = M_1 - \lambda_2 = 0 \tag{A.6}$$
$$\partial_{M_1}\mathcal{L} = N_1 - \lambda_1\tilde{c}_1 M_1^{-2} + \lambda_2\left(c_3 + 2\bar{c}_3 M_1\right) = 0 \tag{A.7}$$
$$\partial_{N_2}\mathcal{L} = M_2^{\Gamma} - 2\lambda_3 c_2 N_2^{-3} = 0 \tag{A.8}$$
$$\partial_{M_2}\mathcal{L} = N_2 M_2^{\Gamma-1}\Gamma - 2\lambda_3\bar{c}_2 M_2^{-3} = 0 \tag{A.9}$$
$$\partial_{\lambda_1}\mathcal{L} = \tilde{c}_1 \cdot M_1^{-1} - \mathrm{TOL} = 0 \tag{A.10}$$
$$\partial_{\lambda_2}\mathcal{L} = c_3 M_1 + \bar{c}_3 \cdot M_1^2 - N_1 = 0 \tag{A.11}$$
$$\partial_{\lambda_3}\mathcal{L} = c_2 N_2^{-2} + \bar{c}_2 \cdot M_2^{-2} - \mathrm{TOL}_r \cdot \delta = 0 \tag{A.12}$$
$$\partial_{\lambda_4}\mathcal{L} = c_\gamma \cdot \gamma^2 - \mathrm{TOL} = 0 \tag{A.13}$$

Using (A.10) in combination with yields

$$M_1^* = \frac{\tilde{c}_1}{\mathrm{TOL}} \overset{(A.11)}{\Longrightarrow} N_1^* = \frac{c_3\tilde{c}_1}{\mathrm{TOL}} + \bar{c}_3 \cdot \left(\frac{\tilde{c}_1}{\mathrm{TOL}^2}\right).$$

Combining Equation (A.8) and (A.9) gives a relation between N_2 and M_2

$$N_2 = \sqrt{\frac{c_2\Gamma}{\bar{c}_2}} \cdot M_2.$$

Substituting this relation into (A.12) yields

$$M_2^* = \sqrt{\bar{c}_2\frac{\Gamma+1}{\Gamma}} \cdot (\mathrm{TOL}_r \cdot \delta)^{-\frac{1}{2}} \implies N_2^* = \sqrt{c_2(1+\Gamma)} \cdot (\mathrm{TOL}_r \cdot \delta)^{-\frac{1}{2}}$$

Finally, using (A.13) we get

$$\gamma^* = \sqrt{\frac{\mathrm{TOL}}{c_\gamma}}.$$

Putting everything together yields the optimal computational work of

$$\mathcal{W}_{\mathrm{DG}}^* = N_1^* \cdot M_1^* + N_2^* \cdot (M_2^*)^{\Gamma} = \mathcal{O}\left(\mathrm{TOL}^{-3} + (\mathrm{TOL}_r \cdot \delta)^{-\frac{1+\Gamma}{2}}\right).$$

A.6 Proof of Theorem 4.6.3

Consider the Lagrangian

$$\begin{aligned}\mathcal{L}(N_1, M_1, N_2, M_2, k, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5) = {} & k \cdot \left(N_1 \cdot M_1 + N_2 \cdot M_2^{\Gamma}\right) \\ & + \lambda_1\left(\tilde{c}_1 \cdot M_1^{-1} - \mathrm{TOL}\right) \\ & + \lambda_2\left(c_3 M_1 + \bar{c}_3 M_1^2 - N_1\right) \\ & + \lambda_3\left(c_2 N_2^{-2} + \bar{c}_2 \cdot M_2^{-2} - \mathrm{TOL}_r \cdot \delta\right) \\ & + \lambda_4\left(c_\gamma \cdot \gamma^2 - \mathrm{TOL}\right) \\ & + \lambda_5\left(|\beta_0 - \bar{\beta}_0| \cdot 2^{-k} - \mathrm{TOL}_b\right)\end{aligned}$$

with partial derivatives and first order conditions for extrema

$$\partial_{N_1}\mathcal{L} = kM_1 - \lambda_2 = 0 \tag{A.14}$$
$$\partial_{M_1}\mathcal{L} = kN_1 - \lambda_1\tilde{c}_1 M_1^{-2} + \lambda_2\left(c_3 + 2\bar{c}_3 M_1\right) = 0 \tag{A.15}$$
$$\partial_{N_2}\mathcal{L} = kM_2^{\Gamma} - 2\lambda_3 c_2 N_2^{-3} = 0 \tag{A.16}$$
$$\partial_{M_2}\mathcal{L} = kN_2 M_2^{\Gamma-1}\Gamma - 2\lambda_3\bar{c}_2 M_2^{-3} = 0 \tag{A.17}$$
$$\partial_{k}\mathcal{L} = N_1 M_1 + N_2 M_2^{\Gamma} - \lambda_5 \cdot |\beta_0 - \bar{\beta}_0| \cdot \log(2) \cdot 2^{-k} = 0 \tag{A.18}$$
$$\partial_{\lambda_1}\mathcal{L} = \tilde{c}_1 \cdot M_1^{-1} - \mathrm{TOL} = 0 \tag{A.19}$$
$$\partial_{\lambda_2}\mathcal{L} = c_3 M_1 + \bar{c}_3 \cdot M_1^2 - N_1 = 0 \tag{A.20}$$
$$\partial_{\lambda_3}\mathcal{L} = c_2 N_2^{-2} + \bar{c}_2 \cdot M_2^{-2} - \mathrm{TOL}_r \cdot \delta = 0 \tag{A.21}$$
$$\partial_{\lambda_4}\mathcal{L} = c_\gamma \cdot \gamma^2 - \mathrm{TOL} = 0 \tag{A.22}$$
$$\partial_{\lambda_5}\mathcal{L} = |\beta_0 - \bar{\beta}_0| \cdot 2^{-k} - \mathrm{TOL}_b = 0 \tag{A.23}$$

Using (A.19), yields

$$M_1^* = \frac{\tilde{c}_1}{\mathrm{TOL}} \overset{(A.20)}{\Longrightarrow} N_1^* = \frac{c_3\tilde{c}_1}{\mathrm{TOL}} + \bar{c}_3 \cdot \left(\frac{\tilde{c}_1}{\mathrm{TOL}^2}\right).$$

Combining Equation (A.16) and (A.17) gives a relation between N_2 and M_2

$$N_2 = \sqrt{\frac{c_2\Gamma}{\bar{c}_2}} \cdot M_2.$$

Substituting this relation into (A.21) yields

$$M_2^* = \sqrt{\bar{c}_2\frac{\Gamma+1}{\Gamma}} \cdot (\mathrm{TOL}_r \cdot \delta)^{-\frac{1}{2}} \implies N_2^* = \sqrt{c_2(1+\Gamma)} \cdot (\mathrm{TOL}_r \cdot \delta)^{-\frac{1}{2}}$$

Using (A.22) we get

$$\gamma^* = \sqrt{\frac{\mathrm{TOL}}{c_\gamma}}$$

and (A.23) to get

$$k^* = \left\lceil \frac{1}{\log(2)} \cdot \log\left(\frac{|\beta_0 - \bar{\beta}_0|}{\mathrm{TOL}_b}\right) \right\rceil$$

Putting everything together yields the optimal computational work of

$$\mathcal{W}_{\mathrm{bisec}}^* = k^*\left(N_1^* \cdot M_1^* + N_2^* \cdot (M_2^*)^{\Gamma}\right) = \mathcal{O}\left(\log(\mathrm{TOL}_b^{-1})\left(\mathrm{TOL}^{-3} + (\mathrm{TOL}_r \cdot \delta)^{-\frac{1+\Gamma}{2}}\right)\right).$$

A.7 Proof of Theorem 4.6.4

Consider the Lagrangian

$$\begin{aligned}\mathcal{L}(N_1, M_1, N_2, M_2, K, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5) = K \cdot \left(N_1 \cdot M_1 + N_2 \cdot M_2^{\Gamma}\right) \\ + \lambda_1 \left(\tilde{c}_1 \cdot M_1^{-1} - \mathrm{TOL}\right) \\ + \lambda_2 \left(c_3 M_1 + \bar{c}_3 M_1^2 - N_1\right) \\ + \lambda_3 \left(c_2 N_2^{-2} + \bar{c}_2 \cdot M_2^{-2} - \mathrm{TOL}_r \cdot \delta\right) \\ + \lambda_4 \left(c_\gamma \cdot \gamma^2 - \mathrm{TOL}\right) \\ + \lambda_5 \left(R \cdot S \cdot K^{-\frac{1}{2}} - \mathrm{TOL}_{\mathrm{SG}}\right)\end{aligned}$$

with partial derivatives and first order conditions for extrema

$$\partial_{N_1}\mathcal{L} = kM_1 - \lambda_2 = 0 \tag{A.24}$$
$$\partial_{M_1}\mathcal{L} = kN_1 - \lambda_1 \tilde{c}_1 M_1^{-2} + \lambda_2 \left(c_3 + 2\bar{c}_3 M_1\right) = 0 \tag{A.25}$$
$$\partial_{N_2}\mathcal{L} = kM_2^{\Gamma} - 2\lambda_3 c_2 N_2^{-3} = 0 \tag{A.26}$$
$$\partial_{M_2}\mathcal{L} = kN_2 M_2^{\Gamma-1}\Gamma - 2\lambda_3 \bar{c}_2 M_2^{-3} = 0 \tag{A.27}$$
$$\partial_K\mathcal{L} = N_1 M_1 + N_2 M_2^{\Gamma} - \frac{1}{2}\lambda_5 \cdot RS \cdot K^{-\frac{3}{2}} = 0 \tag{A.28}$$
$$\partial_{\lambda_1}\mathcal{L} = \tilde{c}_1 \cdot M_1^{-1} - \mathrm{TOL} = 0 \tag{A.29}$$
$$\partial_{\lambda_2}\mathcal{L} = c_3 M_1 + \bar{c}_3 \cdot M_1^2 - N_1 = 0 \tag{A.30}$$
$$\partial_{\lambda_3}\mathcal{L} = c_2 N_2^{-2} + \bar{c}_2 \cdot M_2^{-2} - \mathrm{TOL}_r \cdot \delta = 0 \tag{A.31}$$
$$\partial_{\lambda_4}\mathcal{L} = c_\gamma \cdot \gamma^2 - \mathrm{TOL} = 0 \tag{A.32}$$
$$\partial_{\lambda_5}\mathcal{L} = R \cdot S \cdot K^{-\frac{1}{2}} - \mathrm{TOL}_{\mathrm{SG}} = 0 \tag{A.33}$$

Using (A.29), yields

$$M_1^* = \frac{\tilde{c}_1}{\mathrm{TOL}} \overset{(A.30)}{\Longrightarrow} N_1^* = \frac{c_3 \tilde{c}_1}{\mathrm{TOL}} + \bar{c}_3 \cdot \left(\frac{\tilde{c}_1}{\mathrm{TOL}^2}\right).$$

Combining Equation (A.26) and (A.27) gives a relation between N_2 and M_2

$$N_2 = \sqrt{\frac{c_2\Gamma}{\bar{c}_2}} \cdot M_2.$$

Substituting this relation into (A.31) yields

$$M_2^* = \sqrt{\bar{c}_2 \frac{\Gamma+1}{\Gamma}} \cdot (\mathrm{TOL}_r \cdot \delta)^{-\frac{1}{2}} \implies N_2^* = \sqrt{c_2(1+\Gamma)} \cdot (\mathrm{TOL}_r \cdot \delta)^{-\frac{1}{2}}$$

Using (A.32) we get

$$\gamma^* = \sqrt{\frac{\mathrm{TOL}}{c_\gamma}}$$

and (A.33) to get

$$K^* = \left\lceil \left(\frac{RS}{\mathrm{TOL}_{\mathrm{SG}}}\right)^2 \right\rceil.$$

Putting everything together yields the optimal computational work of

$$\mathcal{W}^*_{\mathrm{SG}} = K^* \left(N_1^* \cdot M_1^* + N_2^* \cdot (M_2^*)^{\Gamma} \right) = \mathcal{O}\left(\mathrm{TOL}_{\mathrm{SG}}^{-2} \left(\mathrm{TOL}^{-3} + (\mathrm{TOL}_r \cdot \delta)^{-\frac{1+\Gamma}{2}} \right) \right).$$

Appendix B

Supplementary Material

B.1 Convergence Analysis

In this section, we will numerically estimate the constants $\tilde{c}_1, c_2, \bar{c}_2$ in Theorem 3.3.1 and 4.2.1 arising from the errors in the UFD scheme of Algorithm 2 and Crank-Nicolson method from Section 4.2.1.

B.1.1 Algorithm 2: UFD scheme

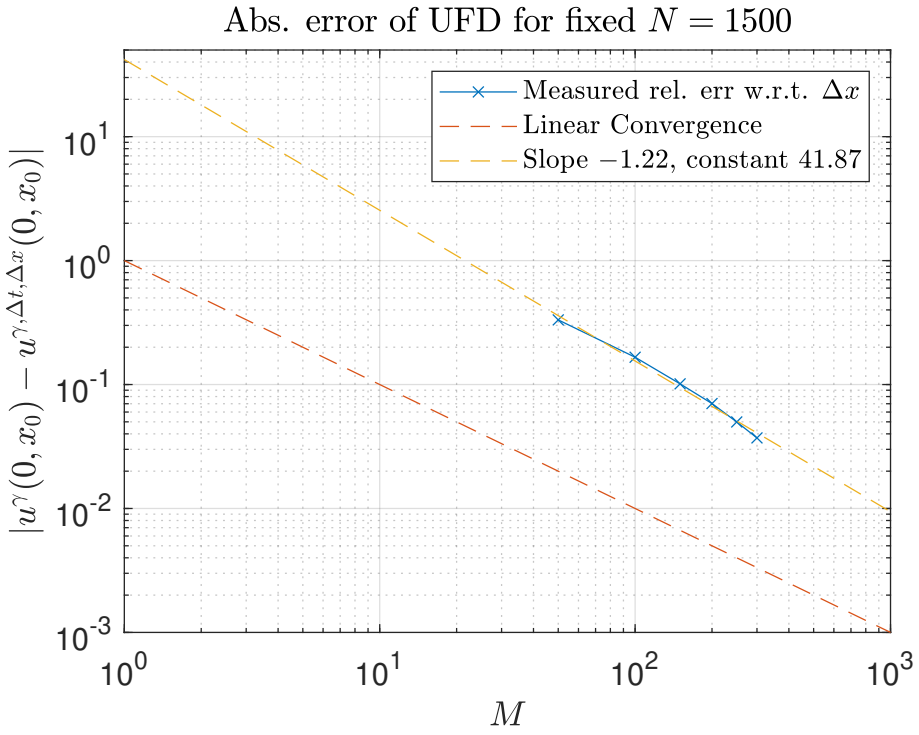
First, we will investigate the constants $\tilde{c}_1 > 0$ from the UFD scheme error in Theorem 3.3.1, i.e.

$$|u^{\gamma}(0, x_0) - u^{\gamma, \Delta t, \Delta x}(0, x_0)| \leq \frac{\tilde{c}_1}{M_1}$$

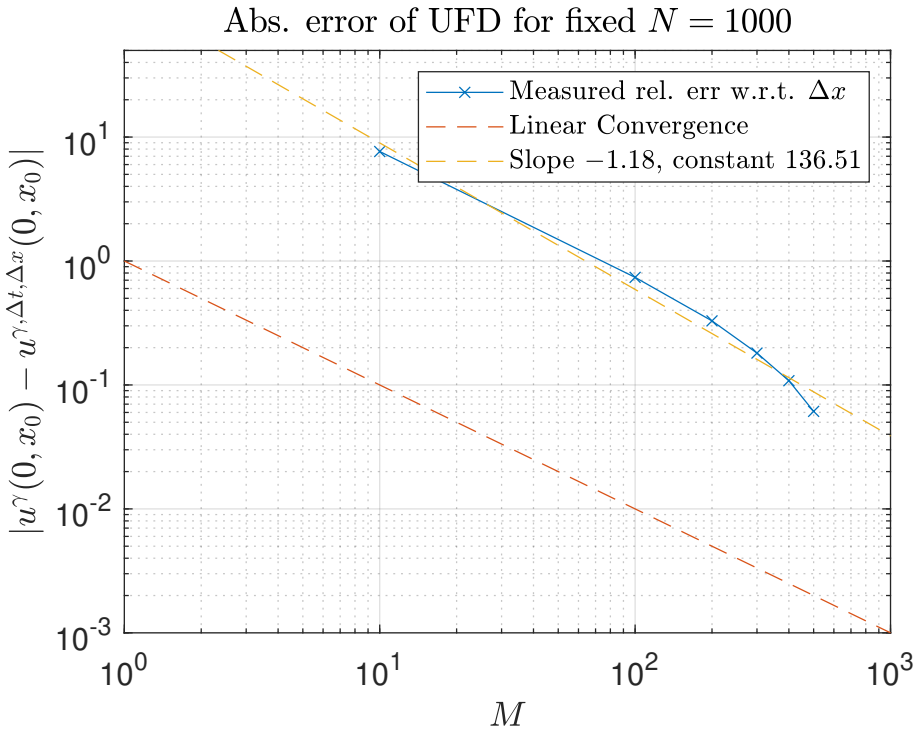
for the positive integer M_1 from the space discretization. We can choose a very high number of time discretization points $N_1 \gg M_1$ and increment the number of space discretization points M_1. Results for both running examples are shown in Figure B.1.

B.1.2 Crank-Nicolson method

In this section, the quadratic convergence rate of the Crank-Nicolson method is veryfied numerically for the KBE in (2.16). Similar to the UFD scheme's convergence analysis, this time with both N_2 and M_2, we can fix one parameter and increment the second one. Additionally, we use a reference solution of $N_2 = M_2 = 1200$ for Problem 8 and $N_2 = M_2 = 3000$ for Problem 9. For the control input in the KBE in (2.16), we used the optimal control α^{λ} resulting from the HJB equation (4.4) with fixed $\lambda = 10$ and $N_1 = M_1 = 100$ time and space steps. Additionally, we fixed the parameter $\gamma = 1/3$. In Figure B.2, the same procedure was performed again for Problem 8 but this time we took a higher value for $\lambda = 50$. This was done due to the fact that the optimal control in (3.6) is dependent on the x-gradient of the value function which has a singularity in the threshold point (T, K). To handle larger values for the Lagrange Multiplier, we redid the experiment. The optimal control was estimated using $N_1 = 500$, $M_1 = 200$ discretization points.

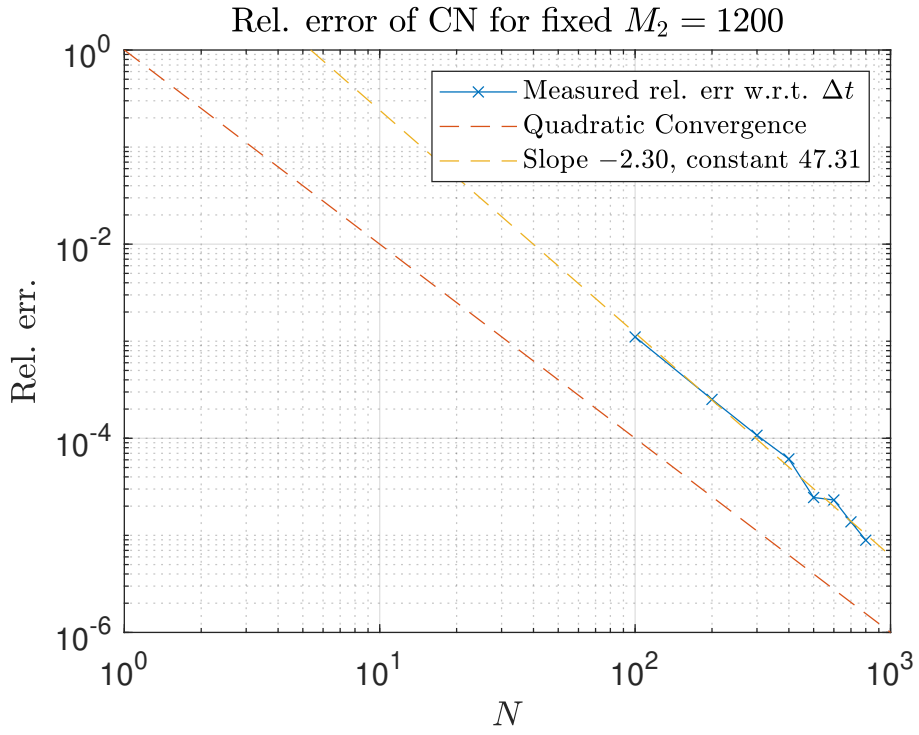


(a) Problem 8: $\tilde{c}_1 \approx 41.87$

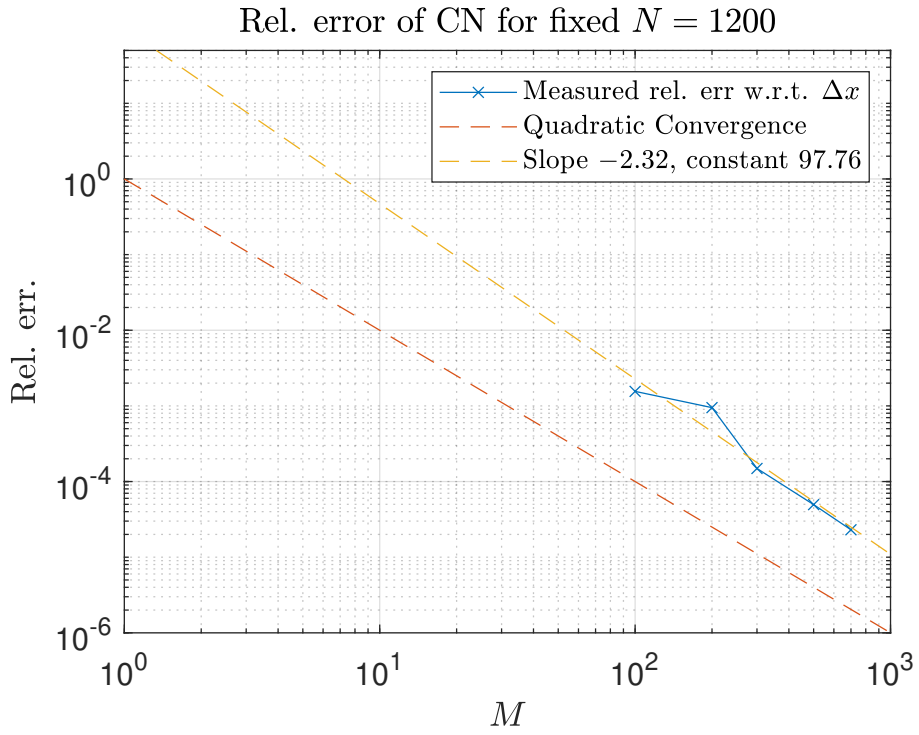


(b) Problem 9: $\tilde{c}_1 \approx 136.51$

Figure B.1: Estimation of $\tilde{c}_1$: Problem 8 on the domain $[0,1] \times [-10,10]$ with a reference solution using $N_1 = 1500$, $M_1 = 650$ steps. Problem 9 on the domain $[0,1] \times [-2,2]$ with a reference solution using $N_1 = M_1 = 1000$ steps.

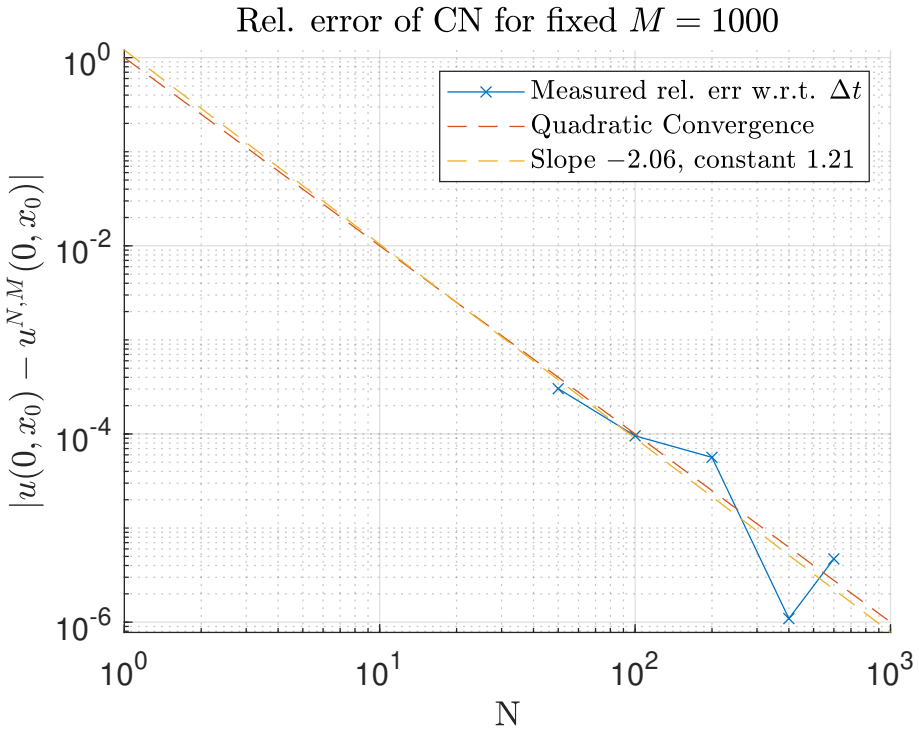


(a) $c_2 \approx 47.31$ using fixed $M_2 = 1200$ steps.

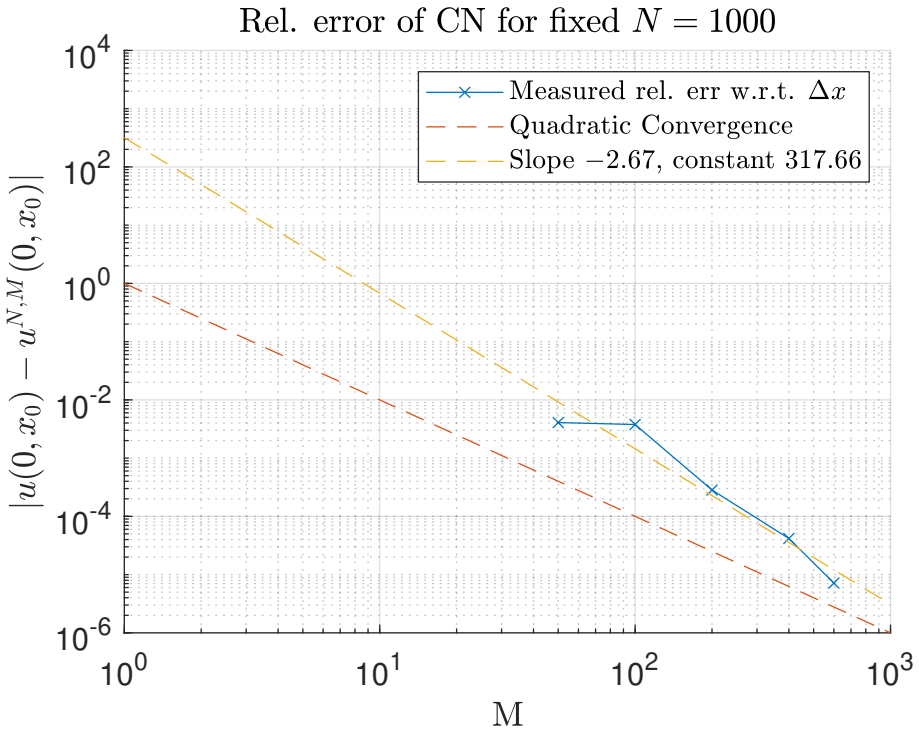


(b) $\bar{c}_2 \approx 97.76$ using fixed $N_2 = 1200$ steps.

Figure B.2: Estimation of the constant $c_2, \bar{c}_2$ and the rate of convergence of the Crank-Nicolson method for Problem 8 on the domain $[0, T] \times [-10, 10]$ with a reference solution using $N_2 = M_2 = 1200$ and Lagrange Multiplier $\lambda = 50$.



(a) $c_2 \approx 1.21$ using fixed $M_2 = 1000$ steps.



(b) $\bar{c}_2 \approx 317.66$ using fixed $N_2 = 1000$ steps.

Figure B.3: Estimation of the constant $c_2, \bar{c}_2$ and the rate of convergence of the Crank-Nicolson method for Problem 9 on the domain $[0,T] \times [-2,2]$ with a reference solution using $N_2 = M_2 = 3000$ steps.

B.2 Ricatti Ansatz Analysis

Continuing the analysis of the usage of the Ricatti Ansatz in Section 3.4, we can also analyse the solution to the optimally controlled process defined by the SDE in Equation (3.29). The process is an Ornstein-Uhlenbeck process which is a Gaussian process. Gaussian processes can be defined by their mean function $m(t)$ and covariance function $k(t,t')$. We denote by

$$m(t) := \mathbb{E}[X(t)] \quad \text{and} \quad k(t,t') := \mathbb{E}[X(t)X(t')] - \mathbb{E}[X(t)]\mathbb{E}[X(t')] \tag{B.1}$$

the mean and covariance functions of the process

$$X(t) \sim \mathcal{GP}(m(t), k(t,t')). \tag{B.2}$$

Let

$$r(t) = -\theta - \frac{f(t)}{\mu}, \quad I(t) = \exp\left(-\int_0^t r(s)ds\right).$$

We can integrate the SDE defined by (3.29) and take the expected value to get

$$\begin{aligned}
& dX(t) = r(t)X(t)dt + \sigma dW(t) \\
\implies & \int_0^t dX(s) = \int_0^t r(s)X(s)ds + \int_0^t \sigma dW(s) \\
\implies & X(t) - X(0) = \int_0^t r(s)X(s)ds + \int_0^t \sigma dW(s) \\
\implies & \mathbb{E}[X(t)] - x_0 = \int_0^t r(s)\mathbb{E}[X(s)]ds + \mathbb{E}\left[\int_0^t \sigma dW(s)\right] \\
\implies & \frac{dm(t)}{dt} = r(t)m(t), \quad m(0) = x_0.
\end{aligned} \tag{B.3}$$

Equation (B.3) is a separable ordinary differential equation with solution

$$m(t) = x_0 \cdot \exp\left(-\theta t - \frac{1}{\mu}\int_0^t f(s)ds\right).$$

We can find an explicit solution to the SDE defined in Equation (3.29). First, we multiply both sides of the SDE by $I(t)$ and rearrange to get

$$\begin{aligned}
& I(t)dX(t) = I(t)\left(-\theta - \frac{f(t)}{\mu}\right)X(t)dt + I(t)\sigma dW(t) \\
\iff & d\left(I(t)X(t)\right) = I(t)\sigma dW(t)
\end{aligned}$$

Integrating both sides from 0 to t and solving for $X(t)$ results in

$$\begin{aligned}
& I(t)X(t) = x_0 + \sigma\int_0^t I(s)dW(s) \\
\implies & X(t) = x_0 I^{-1}(t) + \sigma I^{-1}(t)\int_0^t I(s)dW(s)
\end{aligned} \tag{B.4}$$

where

$$I^{-1}(t) = \exp\left(\int_0^t r(s)ds\right).$$

The mean of the SDE in Equation (3.29) can then be determined by taking the expectation on both sides of the explicit solution representation in Equation (B.4) and using the martingale property of Itô-integrals, see Theorem 2.1.2 (ii), yielding

$$m(t) = \mathbb{E}\left[X(t)\right] = \mathbb{E}\left[x_0 I^{-1}(t) + \sigma I^{-1}(t)\int_0^t I(s)dW(s)\right] = x_0 I^{-1}(t).$$

For the covariance function $k(t,t')$, we compute the quantity $\mathbb{E}\left[X(t)X(t')\right]$, for some $t < t'$, using the properties of expectations and Itô-isometry, see Theorem 2.1.2 (iii), giving us

$$\begin{aligned}
\mathbb{E}\left[X(t)X(t')\right] &= \mathbb{E}\left[\left(x_0 I^{-1}(t) + \sigma I^{-1}(t)\int_0^t I(s)dW(s)\right)\cdot\left(x_0 I^{-1}(t') + \sigma I^{-1}(t')\int_0^{t'} I(s)dW(s)\right)\right] \\
&= x_0^2 I^{-1}(t)I^{-1}(t') + \sigma^2 I^{-1}(t)I^{-1}(t')\mathbb{E}\left[\int_0^t I(s)dW(s)\int_0^{t'} I(s)dW(s)\right] \\
&= x_0^2 I^{-1}(t)I^{-1}(t') + \sigma^2 I^{-1}(t)I^{-1}(t')\int_0^{\min(t,t')} I(s)^2 ds.
\end{aligned}$$

We can write the covariance function of the Gaussian process defining the optimally controlled SDE using the Ricatti Ansatz as

$$\begin{aligned}
k(t,t') &= \mathbb{E}[X(t)X(t')] - \mathbb{E}[X(t)]\mathbb{E}[X(t')] \\
&= x_0^2 I(t)^{-1}I(t')^{-1} + \sigma^2 I^{-1}(t)I^{-1}(t')\int_0^{\min(t,t')} I(s)^2 ds - x_0 I(t)^{-1}\cdot x_0 I(t')^{-1} \\
&= \sigma^2 I^{-1}(t)I^{-1}(t')\int_0^{\min(t,t')} I(s)^2 ds
\end{aligned}$$

For completeness, we are interested in the integral $I(t)$ for any $t \in [0,T]$. Denote by

$$q(t) := w_1(t) - \frac{\lambda_1}{\lambda_2} e^{T(\lambda_1-\lambda_2)} w_2(t)$$

the denominator of Equation (3.30). The nominator in Equation (3.30) is the derivative of q. Thus, we can write the desired integral as

$$\begin{aligned}
I(t) &= \exp\left(\theta t - \int_0^t \frac{q'(s)}{q(s)}ds\right) = \exp\left(\theta t - \log\left(\frac{|w_1(t) - \frac{\lambda_1}{\lambda_2}e^{T(\lambda_1-\lambda_2)}w_2(t)|}{|1 - \frac{\lambda_1}{\lambda_2}e^{T(\lambda_1-\lambda_2)}|}\right)\right) \\
&= e^{\theta t}\cdot\left(\frac{1 - \frac{\lambda_1}{\lambda_2}e^{T(\lambda_1-\lambda_2)}}{w_1(t) - \frac{\lambda_1}{\lambda_2}e^{T(\lambda_1-\lambda_2)}w_2(t)}\right)
\end{aligned}$$

with $\lambda_1 = \theta + \sqrt{\theta^2 + \frac{1}{\mu}}$ and $\lambda_2 = \theta + \sqrt{\theta^2 + \frac{1}{\mu}}$.

Bibliography

[1] Huyên Pham. *Continuous-time stochastic control and optimization with financial applications*. Vol. 61. Springer Science & Business Media, 2009.

[2] Priscilla E Greenwood and Lawrence M Ward. *Stochastic neuron models*. Vol. 1. Springer, 2016.

[3] Etienne Pardoux. *Probabilistic models of population evolution*. 2015.

[4] Lars Blackmore. "Robust path planning and feedback design under stochastic uncertainty". In: *AIAA Guidance, Navigation and Control Conference and Exhibit*. 2008, p. 6304.

[5] Michael P Vitus. *Stochastic control via chance constrained optimization and its application to unmanned aerial vehicles*. Stanford University, 2012.

[6] Kengy Barty, Pierre Carpentier, and Pierre Girardeau. "Decomposition of large-scale stochastic optimal control problems". In: *RAIRO-Operations Research* 44.3 (2010), pp. 167–183.

[7] Kengy Barty et al. "Price decomposition in large-scale stochastic optimal control". In: *arXiv preprint arXiv:1012.2092* (2010).

[8] Lars Blackmore. "A probabilistic particle control approach to optimal, robust predictive control". In: *AIAA Guidance, Navigation, and Control Conference and Exhibit*. 2006, p. 6240.

[9] Lars Blackmore and Masahiro Ono. "Convex chance constrained predictive control without sampling". In: *AIAA guidance, navigation, and control conference*. 2009, p. 5876.

[10] L Jeff Hong, Yi Yang, and Liwei Zhang. "Sequential convex approximations to joint chance constrained programs: A Monte Carlo approach". In: *Operations Research* 59.3 (2011), pp. 617–630.

[11] Masahiro Ono and Brian C Williams. "Decentralized chance-constrained finite-horizon optimal control for multi-agent systems". In: *49th IEEE Conference on Decision and Control (CDC)*. IEEE. 2010, pp. 138–145.

[12] James Luedtke and Shabbir Ahmed. "A sample approximation approach for optimization with probabilistic constraints". In: *SIAM Journal on Optimization* 19.2 (2008), pp. 674–699.

[13] Dennis H Van Hessem, Carsten W Scherer, and Okko H Bosgra. "LMI-based closed-loop economic optimization of stochastic process operation under state and input constraints". In: *Proceedings of the 40th IEEE Conference on Decision and Control (Cat. No. 01CH37228)*. Vol. 5. IEEE. 2001, pp. 4228–4233.

[14] Dennis H van Hessem and Okko H Bosgra. "Closed-loop stochastic dynamic process optimization under input and state constraints". In: *Proceedings of the 2002 American Control Conference (IEEE Cat. No. CH37301)*. Vol. 3. IEEE. 2002, pp. 2023–2028.

[15] Dennis H van Hessem and Okko H Bosgra. "A full solution to the constrained stochastic closed-loop MPC problem via state and innovations feedback and its receding horizon implementation". In: *42nd IEEE International Conference on Decision and Control (IEEE Cat. No. 03CH37475)*. Vol. 1. IEEE. 2003, pp. 929–934.

[16] Dennis H van Hessem and Okko H Bosgra. "Closed-loop stochastic model predictive control in a receding horizon implementation on a continuous polymerization reactor example". In: *Proceedings of the 2004 American Control Conference*. Vol. 1. IEEE. 2004, pp. 914–919.

[17] Dennis Harald Van Hessem. "Stochastic inequality constrained closed-loop model predictive control with application to chemical process operation". In: (2004).

[18] Patrick Piprek, Sébastien Gros, and Florian Holzapfel. "Rare event chance-constrained optimal control using polynomial chaos and subset simulation". In: *Processes* 7.4 (2019), p. 185.

[19] Arkadi Nemirovski and Alexander Shapiro. "Convex approximations of chance constrained programs". In: *SIAM Journal on Optimization* 17.4 (2007), pp. 969–996.

[20] Niklas Schmid et al. "Computing optimal joint chance constrained control policies". In: *arXiv preprint arXiv:2312.10495* (2023).

[21] Laurent Pfeiffer, Xiaolu Tan, and Yu-Long Zhou. "Duality and approximation of stochastic optimal control problems under expectation constraints". In: *SIAM Journal on Control and Optimization* 59.5 (2021), pp. 3231–3260.

[22] Laurent Pfeiffer. "Sensitivity analysis for optimal control problems. Stochastic optimal control with a probability constraint". PhD thesis. Ecole Polytechnique X, 2013.

[23] Laurent Pfeiffer. "Two approaches to constrained stochastic optimal control problems". In: *SFB Report* 7 (2015), p. 2015.

[24] Yuk-Loong Chow, Xiang Yu, and Chao Zhou. "On dynamic programming principle for stochastic control under expectation constraints". In: *Journal of Optimization Theory and Applications* 185.3 (2020), pp. 803–818.

[25] Caroline Geiersbach and Michael Hintermüller. "Optimality conditions and Moreau–Yosida regularization for almost sure state constraints". In: *ESAIM: Control, Optimisation and Calculus of Variations* 28 (2022), p. 80.

[26] Caroline Geiersbach, René Henrion, and Pedro Pérez-Aros. "Numerical solution of an optimal control problem with probabilistic and almost sure state constraints". In: *arXiv preprint arXiv:2311.15115* (2023).

[27] Caroline Geiersbach and René Henrion. "Optimality Conditions in Control Problems with Random State Constraints in Probabilistic or Almost Sure Form". In: *Mathematics of Operations Research* (2024).

[28] Bruno Bouchard, Romuald Elie, and Cyril Imbert. "Optimal control under stochastic target constraints". In: *SIAM Journal on Control and Optimization* 48.5 (2010), pp. 3501–3531.

[29] J Frédéric Bonnans and Francisco J Silva. "First and second order necessary conditions for stochastic optimal control problems". In: *Applied Mathematics & Optimization* 65 (2012), pp. 403–439.

[30] Nicole El Karoui, Monique Jeanblanc, and Vincent Lacoste. "Optimal portfolio management with American capital guarantee". In: *Journal of Economic Dynamics and Control* 29.3 (2005), pp. 449–468.

[31] R Tyrrell Rockafellar, Stanislav Uryasev, et al. "Optimization of conditional value-at-risk". In: *Journal of risk* 2 (2000), pp. 21–42.

[32] Phelim Boyle and Weidong Tian. "Portfolio management with constraints". In: *Mathematical Finance* 17.3 (2007), pp. 319–343.

[33] Nadhir Ben Rached, Shyam Mohan Subbiah Pillai, and Raúl Tempone. "Optimal power procurement for green cellular wireless networks under uncertainty and chance constraints". In: *Entropy* 27.3 (2025), p. 308.

[34] Dimitri Bertsekas. *Dynamic programming and optimal control: Volume I*. Vol. 4. Athena scientific, 2012.

[35] Apurva Patil et al. "Chance-constrained stochastic optimal control via path integral and finite difference methods". In: *2022 IEEE 61st Conference on Decision and Control (CDC)*. IEEE. 2022, pp. 3598–3604.

[36] Jesper Carlsson et al. "Stochastic differential equations: Models and numerics". In: *Lecture notes* (2010).

[37] Grigori N Milstein and Michael V Tretyakov. *Stochastic numerics for mathematical physics*. Vol. 39. Springer, 2004.

[38] Bernt Oksendal. *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, 2013.

[39] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[40] Stephen Boyd, Lin Xiao, and Almir Mutapcic. "Subgradient methods". In: *lecture notes of EE392o, Stanford University, Autumn Quarter* 2004.01 (2003).

[41] Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczynski. *Lectures on stochastic programming: modeling and theory*. SIAM, 2021.

[42] Oldrich Vasicek. "An equilibrium characterization of the term structure". In: *Journal of financial economics* 5.2 (1977), pp. 177–188.

[43] Laurent Badel et al. "Dynamic IV curves are reliable predictors of naturalistic pyramidal-neuron voltage traces". In: *Journal of Neurophysiology* 99.2 (2008), pp. 656–666.

[44] George E Uhlenbeck and Leonard S Ornstein. "On the theory of the Brownian motion". In: *Physical review* 36.5 (1930), p. 823.

[45] Christopher KI Williams and Carl Edward Rasmussen. *Gaussian processes for machine learning*. Vol. 2. 3. MIT press Cambridge, MA, 2006.

[46] Simo Särkkä and Lennart Svensson. *Bayesian filtering and smoothing*. Vol. 17. Cambridge university press, 2023.

[47] Ivar Ekeland and Roger Temam. *Convex analysis and variational problems*. SIAM, 1999.

[48] Bing Sun and Bao-Zhu Guo. "Convergence of an upwind finite-difference scheme for Hamilton–Jacobi–Bellman equation in optimal control". In: *IEEE Transactions on automatic control* 60.11 (2015), pp. 3012–3017.

[49] Iain Smears. "Hamilton-jacobi-bellman equations, analysis and numerical analysis". In: *Univ. Durham, Durham, UK. Available online: http://fourier. dur. ac. uk/Ug/projects/highlights/PR4/SmearsHJBreport. pdf (accessed on 2 February 2018)* (2018).

[50] Helfried Peyrl, Florian Herzog, and Hans P Geering. "Numerical solution of the Hamilton–Jacobi–Bellman equation for stochastic optimal control problems". In: *Proc. 2005 WSEAS International Conference on Dynamical Systems and Control*. 2005, pp. 489–497.

[51] Adil M Bagirov et al. *Numerical nonsmooth optimization*. Springer, 2020.

[52] Omowo Babajide Johnson and Longe Idowu Oluwaseun. "Crank-Nicolson and modified Crank-Nicolson scheme for one dimensional parabolic equation". In: *International Journal of Applied Mathematics and Theoretical Physics* 6.3 (2020), pp. 35–40.

[53] Wolfgang Dahmen, Arnold Reusken, and Arnold Reusken. *Numerik für Ingenieure und Naturwissenschaftler*. Vol. 2. Springer, 2008.

[54] Michael B Giles. "MLMC techniques for discontinuous functions". In: *International Conference on Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing*. Springer. 2022, pp. 33–47.