Numerical Schemes for Stochastic Hybrid Control Problems in Finance

Yufei Zhang

The Queen's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2021

Acknowledgements

First and foremost, I would like to express my sincere gratitude towards my supervisor Prof. Christoph Reisinger, for his indispensable guidance, constructive critique, and consistent encouragement throughout my doctoral studies. I am particularly grateful for the absolute freedom he gave me to explore my scientific interests and for his invaluable feedback which propelled me forward. It has been an honour and great pleasure to be one of his students.

Moreover, I want to express my appreciation to Prof. Kazufumi Ito from North Carolina State University for introducing me to control theory, for teaching me advanced functional analysis and for the fruitful collaboration on Chapter 4.

I would like to thank my colleagues Vadim Kaushansky, Zhenru Wang, Wolfgang Stockinger, Leandro Sanchez Betancourt, Tanut Treetanthiploet, Theerawat Bhudisaksang and Christina Zou, with whom I have shared this enriching DPhil experience. I wish also to thank my friends Chong Liu, Renyuan Xu, Sai-Mang Pun, Tinyi Chu and Zhicheng Xu, for sharing my happinesses and frustrations.

Above all, I could not thank enough my father Yongcun Zhang and my mother Lili Zhai who believed in my DPhil studies from beginning to end. Despite being separated by eight time zones, they never doubted my choices and tried their best to ensure my studies were enjoyable. I am also infinitely grateful to my wife Yuqian Ou, who has accompanied me throughout the DPhil, and provided me with a much appreciated stability. Her unconditional support and unwavering love helped me get through many difficult times, and make this journey full of pleasant memories. Without my parents and my wife, none of this would have been possible.

Statement of Originality

I hereby declare that this thesis contains no material which has been accepted or is currently being submitted for any other degree, diploma, certificate or other qualifications at the University of Oxford or elsewhere. To the best of my knowledge, this thesis contains no material previously published and precise reference is made when a previously published result is used or discussed.

This thesis includes three papers published in peer-reviewed journals, namely the SIAM Journal on Numerical Analysis [88] (Chapter 2), the SIAM Journal on Control and Optimization [89] (Chapter 3) and the Foundations of Computational Mathematics [60] (Chapter 4). The first two are co-authored with my supervisor Prof. Christoph Reisinger, and the last one is co-authored with Prof. Christoph Reisinger and Prof. Kazufumi Ito from North Carolina State University.

The ideas, development and writing up of the thesis, and the papers included in it, were the principal responsibility of myself, Yufei Zhang, working under the supervision of Prof. Christoph Reisinger.

Abstract

In this thesis, we propose a class of numerical schemes for weakly coupled systems of Hamilton-Jacobi-Bellman quasi-variational inequalities (HJBQVIs) arising from stochastic hybrid control problems of regime-switching models with both continuous and impulse controls. By penalizing the difference between the solution and the obstacles, we reduce the HJBQVIs into a sequence of HJB equations, whose solutions converge monotonically to those of the HJBQVIs. We show that the penalty scheme is first-order accurate for HJBQVIs with discrete state spaces, half-order accurate for HJBQVIs with Lipschitz coefficients, and first-order accurate for equations with more regular coefficients. We also demonstrate the convergence of monotone discretizations of the penalized equations, and establish that policy iteration applied to the discrete equation is monotonically convergent with an arbitrary initial guess in an infinite dimensional setting. Numerical examples for infinite-horizon optimal switching problems are presented to illustrate that the penalty scheme along with a continuation procedure in the penalty parameter is significantly more efficient than the conventional direct control scheme for solving HJBQVIs.

We further propose an efficient neural network-based policy iteration algorithm for solving semilinear Hamilton-Jacobi-Bellman-Isaacs (HJBI) boundary value problems, which not only can be combined with the penalty approach to solve high-dimensional stochastic hybrid control problems, but is also applicable to high-dimensional stochastic games of diffusion processes with controlled drift. The algorithm exploits policy iteration to reduce the semilinear problem into a sequence of linear Dirichlet problems, which are subsequently approximated by a multilayer feedforward neural network ansatz. We establish that the numerical solutions converge globally and further demonstrate that this convergence is superlinear, by interpreting the algorithm as an inexact Newton iteration for the HJBI equation. Moreover, we construct the optimal feedback controls from the numerical value functions and deduce convergence. Numerical experiments on the stochastic Zermelo navigation problem are presented to illustrate the theoretical results and to demonstrate the effectiveness of the method.

Contents

List of Figures

Chapter 1

Introduction

The topic of the thesis is the construction of efficient and provably convergent numerical methods for solving Hamilton-Jacobi-Bellman quasi-variational inequalities (HJBQVIs) associated with multi-dimensional stochastic hybrid control problems. In this first chapter, we briefly describe stochastic hybrid control problems arising from mathematical finance, and their connection to Hamilton-Jacobi-Bellman quasi-variational inequalities (HJBQVIs). We shall also outline the structure of this thesis, and highlight the main contributions to the field.

1.1 A motivating example: control of foreign exchange rates

We start by introducing a classical stochastic hybrid control problem on controlling the exchange rate between two currencies, which has been studied in [70, 81, 31, 6]. In this model, a government influences the foreign exchange (FX) rate of its domestic currency by combining the following two approaches: (1) the government chooses a domestic interest rate for all times (a regular control strategy); (2) the government picks specific intervention times to buy or sell foreign currency in large quantities (an impulse control strategy).

For any domestic interest rate process $(r_t)_{t\geq 0}$ chosen by the government, and for any sequence of intervention times $0 = \tau_0 \leq \tau_1 \leq \tau_2 \leq \ldots < \infty$ such that at each τ_k, $k \in \mathbb{N}$, the government will influence the FX market with an amount $\xi_k \in \mathbb{R}$, the state process $(X_t)_{t\geq 0}$ (i.e., the log of the FX rate) would evolve according to the following controlled dynamics: for all $k \in \mathbb{N} \cup \{0\}$,

$$dX_t = -a(r_t - \bar{r})\,dt + \sigma\,dW_t, \quad \tau_k < t < \tau_{k+1}; \quad X_{\tau_{k+1}} = X_{\tau_{k+1}-} + \xi_{k+1},$$

where $\bar{r}$ is the given foreign interest rate, $(W_t)_{t\geq 0}$ is a standard Brownian motion, and $a > 0$ represents the effect of the difference between the domestic and foreign interest rates on the FX rate.

The government's objective is to keep X close to a given target rate $x^\star$ by minimizing the following cost functional

$$\mathbb{E}\left[\int_0^\infty e^{-\beta t}\big((X_t - x^\star)^2 + \rho(r_t - \bar{r})^2\big)\, dt + \sum_{k=1}^\infty e^{-\beta \tau_k}(\kappa|\xi_k| + c)\bigg| X_0 = x\right],$$

over all admissible control strategies $(r_t)_{t\geq 0}$ and $(\tau_1, \xi_1; \tau_2, \xi_2; \ldots)$. Here, $\beta > 0$ is the discount factor, the term $(X_t - x^\star)^2$ penalizes the deviation of the FEM rate from the target, the term $\rho(r_t - \bar{r})^2$ with $\rho \geq 0$ penalizes the government for choosing a domestic interest rate that is far away from the foreign interest rate, and the constants $k, c > 0$ are the fixed and proportional costs paid by the government to intervene the FX market instantaneously.

1.2 General form of stochastic hybrid control problems and HJBQVIs

The general form of stochastic hybrid control problems could involve state dynamics given by controlled regime-switching models and admissible control strategies given by both continuous and impulse controls.

More precisely, let $T > 0$ be a given terminal time and $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a complete filtered probability space. We assume the filtration $\mathbb{F}$ satisfies the usual conditions, whose precise choice will be specified subsequently. Let α be an adapted stochastic control process, and let $\gamma = (\tau_1, \xi_1; \tau_2, \xi_2; \ldots)$ be an impulse control strategy consisting of a sequence of impulse times $0 = \tau_0 \leq \tau_1 \leq \tau_2 \leq \ldots$, and adapted impulse controls $(\xi_1, \xi_2, \ldots)$. Between impulse times, we assume the state process X follows a controlled regime-switching process defined as follows: $X_0 = x \in \mathbb{R}^d$, $I_0 = i \in \mathcal{I}$, and for all $k \in \mathbb{N} \cup \{0\}$,

$$dX_t = b(\alpha_t, I_t, X_t)\, dt + \sigma(\alpha_t, I_t, X_t)\, dW_t, \quad \tau_k < t < \tau_{k+1},$$

where W is a standard Brownian motion, and I is a continuous-time Markov chain with values in the finite set $\mathcal{I}$, which represents the uncertainty in the environment and randomly switches among $M = |\mathcal{I}|$ states, governed by a controlled Markov transition matrix $(d_{ij}^{\alpha_t}(X_t))_{i,j\in\mathcal{I}}$. At an impulse time τ_k, the impulse control ξ_k taking value in the action set $Z(I_{\tau_k^-}, X_{\tau_k^-})$ is applied and instantaneously changes the state into $X_{\tau_k} = \Gamma(I_{\tau_k^-}, X_{\tau_k^-}, \xi_k)$. We assume that the Brownian motion W and the Markov chain I are independent, and the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is generated by W and I and is augmented by the $\mathbb{P}$-null sets.

The aim is to minimize the expected cost over all admissible strategies (α, γ) by considering the following value function:

$$u_i(x) := \inf_{\alpha,\gamma} \mathbb{E}\left[\int_0^\infty \ell(\alpha_t, I_t, X_t) e^{-c(I_t,X_t)t}\, dt + \sum_{k=1}^\infty K(I_{\tau_k^-}, X_{\tau_k^-}, \xi_k) e^{-c(I_{\tau_k^-}, X_{\tau_k^-})\tau_k} \right] \tag{1.2.1}$$

for each $x \in \mathbb{R}^d$ and $i \in \mathcal{I}$, where ℓ and K are the running cost and the impulse cost, respectively.

Such hybrid control problems appear in mathematical finance, such as in the following applications: portfolio optimization with transaction costs [70, 81, 6], control of exchange rates [70, 81, 31, 6], credit securitization [92], inventory control and dividend control [70, 8]. It is well-known that under suitable assumptions, one can apply the dynamic programming principle to (1.2.1) and characterize the value functions $(u_i)_{i\in\mathcal{I}}$ by the viscosity solution to the following weakly coupled system of degenerate Hamilton-Jacobi-Bellman quasi-variational inequalities (HJBQVIs) (see e.g. [100, 98]): for all $i \in \mathcal{I} := \{1, \dots, M\}$,

$$\min\left\{ \inf_{\alpha\in\mathcal{A}_i} \mathcal{L}_i^\alpha(x, u(x), Du_i(x), D^2u_i(x));\ (\mathcal{M}_i u - u_i)(x) \right\} = 0, \quad x \in \mathbb{R}^d, \tag{1.2.2}$$

where $u = (u_i)_{i\in\mathcal{I}}$ denotes the unknown solution, $(\mathcal{L}_i^\alpha)_{i\in\mathcal{I}}$ is a family of second order differential operators corresponding to the generators of the state processes, and $\mathcal{M}_i$ is an intervention operator of the following form:

$$(\mathcal{M}_i u)(x) = \min_{z\in Z_i(x)} \{u_i(\Gamma_i(x, z)) + K_i(x, z)\}.$$

Note that due to the random switching process I, each operator $\mathcal{L}_i^\alpha$ involves all components of the solution u, which leads us to a weakly coupled system of HJBQVIs (see e.g. [103, 23]).

It is clear that for each $i \in \mathcal{I}$, the solution to (1.2.2) is separated into two sets (by free boundaries), i.e., the continuation region where an HJB equation is satisfied, and the action region where $u_i = \mathcal{M}_i u$. As it is usually difficult to obtain analytically the solutions and free boundaries, it is necessary to design efficient and robust numerical methods for solving these fully nonlinear equations.

1.3 Outline and contribution

Throughout this thesis, we provide efficient and provably convergent numerical methods for solving (1.2.2). While cross-referencing and avoiding redundancies, the chapters are kept largely self-contained and allow for separate reading.

We start by studying (1.2.2) with spatial dimension $d \leq 3$. Such an equation can be efficiently discretized by monotone schemes as in [6, 39, 5], which results in a nonlinear discrete equation of HJB type. However, due to the presence of the obstacle term $\mathcal{M}_i u - u_i$, the discrete systems arising from most sensible discretizations of (1.2.2) do not fulfil the convergence conditions of policy iteration (see e.g. [19]), which is the state-of-the-art algorithm to solve discrete HJB equations. In fact, even for some simple intervention operators, policy iteration for discrete QVIs may not be well-defined for an arbitrary initial guess due to the possible singularity of the matrix iterates (see Remark 3.6.1 in Chapter 3 for details).

In Chapter 2, we present a novel penalty approach for a class of discrete QVIs, which includes those arising from a sensible discretization of (1.2.2) associated to hybrid control problems with switching controls. A major advantage of the penalty approach is that policy iteration applied to penalized equations converges superlinearly.

We show that for any given positive switching cost, the solutions of the penalized equations converge monotonically to those of the QVIs. We estimate the penalization errors and are able to deduce that the optimal switching regions are constructed exactly. We further demonstrate that as the switching cost tends to zero, the QVI degenerates into an equation of HJB type, which is approximated by the penalized equation at the same order (up to a log factor) as that for positive switching cost. Numerical experiments on optimal switching problems are presented to illustrate the theoretical results and to demonstrate the effectiveness of the method. To the best of our knowledge, this is the first work which proposes penalty approximations for discrete QVIs in such a generality and presents rigorous error estimates for the penalization errors.

In Chapter 3, we extend the penalty schemes of discrete systems presented in Chapter 2 to HJBQVIs arising from general stochastic hybrid control problems. The penalty approximation reduces the HJBQVI into a sequence of HJB equations, which can then be solved by monotone schemes if the spatial dimension is less than 3, or mesh-free methods in a high-dimensional setting.

We show that the solutions of the penalized equations converge monotonically to those of the HJBQVIs. We further establish that the scheme is half-order accurate for HJBQVIs with Lipschitz coefficients, and first-order accurate for equations with more regular coefficients. Moreover, we construct the action regions and optimal impulse controls based on the error estimates and the penalized solutions. The penalty schemes and convergence results are then extended to HJBQVIs with possibly negative impulse costs. Numerical examples for infinite-horizon optimal switching problems are presented to illustrate the effectiveness of the penalty schemes over the conventional direct control scheme. To the best of our knowledge, this is the first work which quantifies the precise penalization error for systems of degenerate HJBQVIs, depending on the regularity of the coefficients.

In Chapter 4, we propose a class of neural network-based numerical schemes for solving semilinear Hamilton-Jacobi-Bellman-Isaacs (HJBI) boundary value problems, which include penalized HJBQVIs as special cases. We exploit policy iteration to reduce the semilinear problem into a sequence of linear Dirichlet problems, which are subsequently approximated by a multilayer feedforward neural network ansatz. We establish that the numerical solutions converge globally in the H^2-norm, and further demonstrate that this convergence is superlinear, by interpreting the algorithm as an inexact Newton iteration for the HJBI equation. Moreover, we construct the optimal feedback controls from the numerical value functions and deduce convergence. Numerical experiments on the stochastic Zermelo navigation problem are presented to illustrate the theoretical results and to demonstrate the effectiveness of the method. To the best of our knowledge, this is the first work which demonstrates the global superlinear convergence of policy iteration for nonconvex HJBI equations and proposes convergent neural network-based numerical methods for solving the solutions of nonlinear boundary value problems and their derivatives.

Chapter 2

Penalty schemes for discrete quasi-variational inequalities

2.1 Introduction

In this chapter, we consider systems of discrete quasi-variational inequalities (QVIs) stemming from optimal switching control problems of Markov chain models. They also arise naturally from a sensible discretization (i.e., stable and convergent) of elliptic and parabolic QVIs associated to hybrid optimal control problems of jump-diffusion processes with switching controls (see e.g. [83, 6, 33, 39]). Numerical methods for general hybrid control problems will be studied in Chapter 3, whose analyses involve technical regularity analyses based on viscosity solution theory.

More precisely, we consider the following discrete QVIs:

Problem 2.1.1. Find $u = (u^1, \ldots, u^d) \in \mathbb{R}^{N\times d}$, such that

$$G_i(u) = \min(F_i(u),\ u^i - \mathcal{M}_i u) = 0, \quad i \in \mathcal{I} \coloneqq \{1, \cdots, d\}, \tag{2.1.1}$$

where $d \in \mathbb{N} \cap [2, \infty)$, $\mathcal{M}_i u \coloneqq \max_{j\neq i}(u^j - c^{i,j})$ with $c^{i,j} \in (0, \infty)$ for all $j \neq i$, and $F = (F_i)_{i\in\mathcal{I}} : \mathbb{R}^{N\times d} \to \mathbb{R}^{N\times d}$ is a continuous function satisfying the following monotonicity condition:

- there exists a constant $\gamma > 0$ such that for any $u, v \in \mathbb{R}^{N\times d}$, $i \in \mathcal{I}$ and $l \in \mathcal{N} \coloneqq \{1, \ldots, N\}$ with $u^i_l - v^i_l = \max_{j\in\mathcal{I},k\in\mathcal{N}}(u^j_k - v^j_k) \geq 0$, we have

$$F_i(u)_l - F_i(v)_l \geq \gamma(u^i_l - v^i_l). \tag{2.1.2}$$

Following [33], we will refer to $\mathcal{M}u$ as an interconnected obstacle, where $\mathcal{M}$ is a special form of the intervention operator in [6]. Similar to [23] in the continuous context, we shall also refer to a continuous function $F = (F_i)_{i\in\mathcal{I}}$ with condition (2.1.2) as a monotone system, which in particular implies for all $u, v \in \mathbb{R}^{N\times d}$ and $\delta \geq 0$, $F(u+\delta) \geq F(u) + \gamma\delta$ and if $F(u) \leq F(v)$ then $u \leq v$. Since each F_i could depend on all components of the solution u, the system F in Problem 2.1.1 not only includes systems of Isaacs equations [64] or possibly nonlocal Hamilton-Jacobi-Bellman (HJB) variational inequalities [102, 87], but also coupled systems stemming from regime switching models [7] and non-coorperative games [34]. To simplify the presentation, we shall focus on the case with constant switching cost $c^{i,j} \equiv c > 0$ for all $j \neq i$, $i \in \mathcal{I}$, but the results and analysis extend naturally to the cases with general switching costs $c_l^{i,j} > 0$ for all $j \neq i$.

By rewriting the interconnected obstacle $u^i - \mathcal{M}_i u$ as $\min_{j\neq i}(P^j u + c)$, one can easily see that the matrices $(P^j)_{j\neq i}$ involved are neither M-matrices[1] nor weakly chained diagonally dominant matrices ([6]). Hence a direct application of policy iteration could fail due to the possible singularity of the matrix iterates. However, as we shall see later, penalty schemes are always applicable (even for zero switching cost) and it is straightforward to derive convergent iterative solvers for the penalized equations, which make penalty schemes more appealing for solving the QVI (2.1.1). Thus it is important to design efficient penalty schemes for general monotone systems with interconnected obstacles.

Moreover, the implementation of penalty schemes, in particular the choice of penalty parameters, depends greatly on the accuracy of the penalty approximation with a given penalty parameter, hence it is practically important to quantify the penalty errors. However, the non-diagonal dominance of the matrices $(P^j)_{j\neq i}$ poses a significant challenge for estimating the penalization errors. In fact, an essential step in estimating the penalty error for standard obstacle problems is to show that if u^ρ solves the penalized equation with the penalty parameter $\rho \geq 0$, then $u^\rho + C/\rho$ is a feasible solution to the obstacle problem (i.e., it satisfies the constraint) for a large enough constant C independent of the penalty parameter (see e.g. [87]). But this is clearly false in the current setting since the interconnected obstacles remain invariant under any vertical shift of the solutions. We shall overcome these difficulties by introducing certain regularization procedures, which consist of approximating the switching problem by a sequence of obstacle problems involving diagonally dominant matrices, and recover the same convergence rates (up to a log factor) as those for conventional obstacle problems.

[1] A real square matrix A is said to be an M-matrix if off-diagonal entries of A are all non-positive, A is invertible and $A^{-1} \geq 0$ componentwise.

Finally, we remark that the monotonicity condition (2.1.2) is essentially different from the monotonicity in the Euclidean norm discussed in [51], which enables us to consider penalty methods for the QVIs of fully nonlinear degenerate equations including Isaacs equations. To the best of our knowledge, there is no published penalty scheme with rigorous error estimates covering such general QVIs. However, there is a vast literature on penalty methods for variational inequalities (see e.g. [41, 59, 51]). For works covering specific extensions, we refer the reader to [101] for penalty approximations to HJB equations, to [102, 87] for applying policy iteration together with penalization to solve HJB variational inequalities, and to [14, 6, 5] for an application of penalty schemes to classical HJBQVIs (without error estimates).

The main contributions of this chapter are:

- We propose penalty schemes for discrete monotone systems with interconnected obstacles. We present a novel analysis technique for the well-posedness of the penalized equations with a general class of penalty terms by smoothing the monotone systems. We further demonstrate that the solution of the penalized equation converges to the solution of (2.1.1) monotonically from below, which subsequently gives a constructive proof for the existence of solutions to Problem 2.1.1.

- Based on regularizations of the interconnected obstacles, we estimate the penalization error for monotone systems with concave nonlinearity, which include HJBQVIs as special cases (see the discussions below Assumption 1). We introduce two iterative regularization procedures, namely the iterated optimal stopping approximation and the time-marching iteration, which enable us to demonstrate that for any given positive switching cost, the penalty approximation using a penalty term with degree $\sigma > 0$ enjoys convergence of order $\mathcal{O}(\rho^{-\sigma} \ln \rho)$ as the penalty parameter $\rho \to \infty$, independent of the number of switching regimes d. We emphasize that, unlike the error estimates for HJBQVIs in [21, 39], our analysis does not require the running reward functions or the solutions to have a unique sign. Moreover, our error estimate also enables us to exactly construct the switching regions of Problem 2.1.1, which to the best of our knowledge is new, even for classical HJBQVIs.

- We further investigate the limiting case with zero switching cost, where Problem 2.1.1 degenerates to an equation of HJB type, i.e., Problem 2.4.1 below. In this case, the penalty scheme of (2.1.1) leads to a novel penalty scheme for such equations, which admits the same convergence rates (up to a log factor) as those for fixed positive switching cost when the penalty parameter tends to infinity. We remark that this error estimate applies to non-convex/non-concave systems, such as systems of Isaacs equations.

- Contrary to [6, 5], the penalty is applied to each component of the system, which enables us to derive easily implementable and efficient iterative schemes for penalized equations without taking the pointwise maximum over all switching components.

- Numerical examples for infinite-horizon optimal switching problems in the two-regime case and the three-regime case are included to illustrate the theoratical results for the asymptotic behaviours of the penalty errors with respect to the penalty parameter and switching cost.

The remainder of this chapter is organized as follows. We shall propose a class of penalty approximations to Problem 2.1.1 in Section 2.2, and demonstrate its well-posedness and monotone convergence. Then we construct two regularization procedures for Problem 2.1.1 in Section 2.3.1, which enable us to obtain an error estimate of the penalization error for QVIs with positive switching cost in Section 2.3.2. We then proceed to estimate the penalization errors for QVIs with vanishing switching cost in Section 2.4. Numerical examples for two-regime and three-regime optimal switching problems are presented in Section 2.5 to illustrate the effectiveness of our algorithms.

2.2 Penalty approximations of QVIs

In this section, we discuss how Problem 2.1.1 can be approximated by a sequence of penalized equations. The well-posedness of the penalized equations and their monotone convergence shall be established, which subsequently lead to a constructive proof for the well-posedness of (2.1.1).

We start by collecting some useful notation. For every $\delta \in \mathbb{R}$ and $A, B \in \mathbb{R}^{d_1 \times d_2}$ with $d_1, d_2 \in \mathbb{N}$, we denote by $A \geq B$ the relation $A_{ij} \geq B_{ij}$ for all indices i, j, by $A+\delta$ the matrix of elements $A_{ij}+\delta$, by $\min(A,B)$ the matrix of elements $\min(A_{ij}, B_{ij})$, by $A^+ = \max(A,0)$ (resp. $A^- = \max(-A,0)$) the (element-wise) positive (resp. negative) part of A, and by $\|A\|$ the usual sup-norm $\|A\| = \max_{i,j} |A_{ij}|$. Moreover, if $(x^\alpha)_{\alpha\in\mathcal{A}}$ is a bounded subset of $\mathbb{R}^{d_1}$ with $d_1 \in \mathbb{N}$, we denote by $\min_{\alpha\in\mathcal{A}} x^\alpha$ (resp. $\max_{\alpha\in\mathcal{A}} x^\alpha$) the vector of components $\min_{\alpha\in\mathcal{A}} x_i^\alpha$ (resp. $\max_{\alpha\in\mathcal{A}} x_i^\alpha$).

Before introducing the penalty equations, we first adapt the non-loop arguments in [64] to the current discrete setting and establish a comparison theorem of the QVI (2.1.1).

Proposition 2.2.1. *Suppose $c > 0$ and $u = (u^i)_{i\in\mathcal{I}}$ (resp. $v = (v^i)_{i\in\mathcal{I}}$) satisfies*

$$\min(F_i(u),\ u^i - \mathcal{M}_i u) \le 0 \quad (\text{resp.} \ge 0), \quad i \in \mathcal{I},$$

with $(F_i)_{i\in\mathcal{I}}$ and $(\mathcal{M}_i)_{i\in\mathcal{I}}$ given in Problem 2.1.1, then we have $u \le v$.

Proof. Let $M := \max_{j,k}(u_k^j - v_k^j) = u_l^i - v_l^i$. Suppose that $u_l^j \le (\mathcal{M}_j u)_l$ for all $j \in \Gamma$, where $\Gamma := \{j \mid u_l^j - v_l^j = M\}$. Pick $i_1 \in \Gamma$ such that $u_l^{i_1} \le (\mathcal{M}_{i_1} u)_l = u_l^{i_2} - c$ for some $i_2 \ne i_1$. Since $v_l^{i_1} \ge (\mathcal{M}_{i_1} v)_l \ge v_l^{i_2} - c$, we have

$$u_l^{i_1} - u_l^{i_2} \le -c \le v_l^{i_1} - v_l^{i_2},$$

which implies $u_l^{i_2} - v_l^{i_2} \ge u_l^{i_1} - v_l^{i_1} = M$. By the maximality of M, the previous inequality is in fact an equality and hence i_2 is in Γ. Continuing this way, we can pick indices $i_3, i_4, \dots$ in Γ such that $u_l^{i_k} - u_l^{i_{k+1}} \le -c$ for all $k \ge 1$, which further implies that

$$u_l^{i_1} - u_l^{i_n} = \sum_{k=1}^{n-1} u_l^{i_k} - u_l^{i_{k+1}} \le -(n-1)c < 0, \quad \text{for } n > 1.$$

Since Γ is finite, we can use the pigeonhole principle to find $n > 1$ with $i_1 - i_n = 0$, arriving at a contradiction.

The above argument establishes that $u_l^{i_0} > (\mathcal{M}_{i_0} u)_l$ for some $i_0 \in \Gamma$. Consequently, we have $F_{i_0}(u)_l \le 0 \le F_{i_0}(v)_l$. Combining this with the monotonicity (2.1.2) of F, we have $M \le 0$. □

A direct consequence of Proposition 2.2.1 is the uniqueness of solutions to (2.1.1). The existence of solutions to Problem 2.1.1 shall be established constructively via penalty approximations below (see Remark 2.2.2).

Now we are ready to propose the penalty approximation of the QVI (2.1.1), which is an extension of the ideas used for HJB obstacle problems in [102, 87]. For any given parameter $\rho \ge 0$, we shall consider the following penalized problem:

Problem 2.2.1. Find $u^\rho = (u^{\rho,i})_{i\in\mathcal{I}} \in \mathbb{R}^{N\times d}$ such that

$$G_i^\rho(u^\rho)_l := F_i(u^\rho)_l - \rho \sum_{j\ne i} \pi(u_l^{\rho,j} - c - u_l^{\rho,i}) = 0, \quad i \in \mathcal{I},\ l \in \mathcal{N}, \tag{2.2.1}$$ [2]

[2] For notational simplicity, for every $x \in \mathbb{R}^N$, we denote by $\pi(x)$ the vector with components $\pi(x_\ell) \in \mathbb{R}^N$ in the subsequent analysis.

where the penalty term $\pi : \mathbb{R} \to \mathbb{R}$ is a continuous non-decreasing function satisfying $\pi|_{(-\infty,0]} = 0$ and $\pi|_{(0,\infty)} > 0$.

Remark 2.2.1. In (2.2.1), the penalty is applied to each component of the switching system, thus efficient iterative schemes for penalized equations can be implemented without taking the pointwise maximum over all switching components at each index $l \in \mathcal{N}$. However, all the statements below can be shown to hold for penalized problems with penalty terms involving the maximum of all switching components, such as $\max_{j\neq i} \pi(u^{\rho,j} - c - u^{\rho,i})$ or $\pi(\mathcal{M}_i u^\rho - u^{\rho,i})$.

Due to the fact that the control j takes only d distinct values, we can apply the penalty term finitely many times (once per value). This is not directly possible in the framework of [5, 6], where the number of attainable values for the control in the intervention operator grows unbounded as the meshing parameter in the approximation of an infinite control set approaches zero (see [68] for an extension of such a penalty scheme to general intervention operators with an infinite number of control values: the summation is replaced by an integral, which might subsequently be approximated by quadrature).

The following result shows a comparison principle for the penalized equation with a fixed penalty parameter ρ, which not only implies the uniqueness of solutions to the penalized equations, but also plays a crucial role in the convergence analysis of the penalty approximations.

Proposition 2.2.2. *For any given penalty parameter $\rho \geq 0$ and switching cost $c \geq 0$, suppose $u^\rho = (u^{\rho,i})_{i\in\mathcal{I}}$ (resp. $v^\rho = (v^{\rho,i})_{i\in\mathcal{I}}$) satisfies*

$$F_i(u^\rho) - \rho \sum_{j\neq i} \pi(u^{\rho,j} - c - u^{\rho,i}) \leq 0 \quad (\text{resp.} \geq 0), \quad i \in \mathcal{I},$$

then we have $u^\rho \leq v^\rho$.

Proof. Let $M := \max_{j,k} u_k^{\rho,j} - v_k^{\rho,j} = u_l^{\rho,i} - v_l^{\rho,i}$. Then we have $u_l^{\rho,j} - c - u_l^{\rho,i} \leq v_l^{\rho,j} - c - v_l^{\rho,i}$ for all $j \neq i$, and hence $\sum_{j\neq i} \pi(u_l^{\rho,j} - c - u_l^{\rho,i}) \leq \sum_{j\neq i} \pi(v_l^{\rho,j} - c - v_l^{\rho,i})$. This, along with the fact that

$$F_i(u^\rho)_l - F_i(v^\rho)_l - \rho\left(\sum_{j\neq i} \pi(u_l^{\rho,j} - c - u_l^{\rho,i}) - \sum_{j\neq i} \pi(v_l^{\rho,j} - c - v_l^{\rho,i})\right) \leq 0, \quad (2.2.2)$$

leads to $F_i(u^\rho)_l - F_i(v^\rho)_l \leq 0$. Then we can conclude from the monotonicity of F that $M \leq 0$. □

The next result presents an a priori estimate of the solution to the penalized equations, independent of the penalty parameter ρ and switching cost c.

Lemma 2.2.3. *Suppose u^ρ solves Problem 2.2.1 with given penalty parameter $\rho \geq 0$ and switching cost $c \geq 0$, then $\|u^\rho\| \leq \|F(0)\|/\gamma$.*

Proof. Let $|u_l^{\rho,i}| = \|u^\rho\|$. Suppose that $u_l^{\rho,i} \geq 0$, then $u_l^{\rho,j} - c - u_l^{\rho,i} \leq 0$ for all $j \neq i$, hence we deduce from (2.1.2) that

$$\gamma(u_l^\rho - 0) \leq F_i(u^\rho)_l - F_i(0)_l = \rho \sum_{j \neq i} \pi(u_l^{\rho,j} - c - u_l^{\rho,i}) - F_i(0)_l = -F_i(0)_l,$$

thus $\|u^\rho\| \leq \|F(0)\|/\gamma$. On the other hand, suppose that $u_l^{\rho,i} < 0$, we can obtain directly from (2.1.2) and the non-negativity of π that

$$\gamma(0 - u_l^\rho) \leq F_i(0)_l - F_i(u^\rho)_l \leq F_i(0)_l,$$

hence $\|u^\rho\| \leq \|F(0)\|/\gamma$, which leads us to the desired estimate. □

Now we are ready to conclude the well-posedness of the penalized problem (2.2.1). The following lemma has been proved in [82, Theorem 5.3.9], which is of crucial importance for the existence of solutions to the penalized equations.

Lemma 2.2.4. *Suppose that $F : \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable on $\mathbb{R}^n$, and $\nabla F(x)$ is nonsingular for all $x \in \mathbb{R}^n$. Then F is a homeomorphism from $\mathbb{R}^n$ onto $\mathbb{R}^n$ if and only if $\lim_{\|x\| \to \infty} \|F(x)\| = \infty$.*

Theorem 2.2.5. *For any given penalty parameter $\rho \geq 0$ and switching cost $c \geq 0$, Problem 2.2.1 admits a unique solution u^ρ satisfying $\|u^\rho\| \leq \|F(0)\|/\gamma$.*

Proof. The uniqueness and the a priori bound have been established in Proposition 2.2.2 and Lemma 2.2.3 respectively. Now we shall prove the existence of solutions by approximating the penalized equation G^ρ with a sequence of smooth equations.

Consider a family of smooth functions $\delta_m : \mathbb{R}^{N \times d} \to [0, \infty)$ supported in $B(0, 1/m)$ with unit mass, we define the smooth functions $G^{\rho,m} := G^\rho * \delta_m$, where the convolution is applied elementwise. The continuity of G^ρ implies that $G^{\rho,m}$ converges to G^ρ uniformly on compact sets as $m \to \infty$. Moreover, one can deduce from (2.2.2) that G^ρ satisfies the monotonicity condition (2.1.2) with the same constant γ. This implies that for any $u, v \in \mathbb{R}^{N \times d}$, $i \in \mathcal{I}$ and $l \in \mathcal{N}$ with $u_l^i - v_l^i = \max_{j \in \mathcal{I}, k \in \mathcal{N}}(u_k^j - v_k^j) \geq 0$, we have $G_i^\rho(u - y)_l - G_i^\rho(v - y)_l \geq \gamma(u_l^i - v_l^i)$ for all $y \in \mathbb{R}^{N \times d}$, which along with the definition of $(G^{\rho,m})_{m\in\mathbb{N}}$ shows that $(G^{\rho,m})_{m\in\mathbb{N}}$ satisfies (2.1.2) with the constant γ, as $\delta_m(y)\geq 0$ for all $y\in\mathbb{R}^{N\times d}$ and $\int_{\mathbb{R}^{N\times d}}\delta_m(y)\,dy=1$.

For any given $m\in\mathbb{N}$, we shall now apply Lemma 2.2.4 to establish that $G^{\rho,m}$ is a homeomorphism from $\mathbb{R}^{Nd}$ onto $\mathbb{R}^{Nd}$, which implies the equation $G^{\rho,m}=0$ has a solution. More precisely, we shall show (1) the Jacobian matrix $\nabla G^{\rho,m}(u)$ is nonsingular for any given $u\in\mathbb{R}^{Nd}$ and (2) $\lim_{\|u\|\to\infty}\|G^{\rho,m}(u)\|=\infty$. To prove (1), suppose $\nabla G^{\rho,m}(u)x=0$ for some $u,x\in\mathbb{R}^{Nd}$ and let $|x^i_l|=\|x\|\geq 0$ for some $i\in\mathcal{I}$ and $l\in\mathcal{N}$. If $x^i_l\geq 0$, we can deduce from the differentiability and monotonicity of $G^{\rho,m}$ that

$$\gamma(hx^i_l-0)\leq G^{\rho,m}_i(u+hx)_l-G^{\rho,m}_i(u)_l-h(\nabla G^{\rho,m}_i(u)x)_l=\mathcal{O}(h^2),\quad \text{as } h\to 0,$$

which implies $\|x\|=0$. The same conclusion can be drawn for the case with $x^i_l\leq 0$, which implies $x=0$ and consequently the non-singularity of $\nabla G^{\rho,m}(u)$. To prove (2), let $u\in\mathbb{R}^{Nd}$ and $|u^i_l|=\|u\|$ for some $i\in\mathcal{I}$ and $l\in\mathcal{N}$. If $u^i_l\geq 0$, we can obtain from the monotonicity of $G^{\rho,m}$ that

$$\|G^{\rho,m}(u)\|\geq G^{\rho,m}_i(u)_l\geq\gamma(u^i_l)+G^{\rho,m}_i(0)_l\geq\gamma\|u\|-\|G^{\rho,m}(0)\|,$$

where the same estimate can be derived similarly for the case $u^i_l\leq 0$. Therefore, we can conclude the existence of a solution $u^{\rho,m}$ to $G^{\rho,m}=0$. Since $G^{\rho,m}$ satisfies (2.1.2) with the same constant γ, one can deduce from Lemma 2.2.3 and the continuity of $G^{\rho,m}$ that its solution is uniformly bounded, i.e., $\|u^{\rho,m}\|\leq\|G^{\rho,m}(0)\|/\gamma\leq L$ independent of $m\in\mathbb{N}$.

Lastly, let $(u^{\rho,m_k})_{k\in\mathbb{N}}$ be a convergent subsequence of $(u^{\rho,m})_{m\in\mathbb{N}}$ with a limit u^ρ. Note that

$$|G^\rho(u^\rho)-G^{\rho,m_k}(u^{\rho,m_k})|\leq|G^\rho(u^\rho)-G^\rho(u^{\rho,m_k})|+|G^\rho(u^{\rho,m_k})-G^{\rho,m_k}(u^{\rho,m_k})|\to 0,$$

as $k\to\infty$, due to the continuity of G^ρ and the uniform convergence (on compact sets) of $G^{\rho,m}$ to G^ρ. Therefore, u^ρ is a solution of the penalized equation (2.2.1) $G^\rho=0$. □

We end this section with the following monotone convergence result of the penalty approximations.

Theorem 2.2.6. *For any fixed switching cost $c\geq 0$, the solution to Problem 2.2.1 converges monotonically from below to a function $u\in\mathbb{R}^{N\times d}$ as the penalty parameter $\rho\to\infty$. Moreover, u solves Problem 2.1.1 if the switching cost c is positive.*

Proof. It is straightforward to verify that if u^{ρ_1} satisfies (2.2.1) with the parameter ρ_1 and $\rho_1 \geq \rho_2 \geq 0$, then $G^{\rho_2}(u^{\rho_1}) \geq 0$. Hence one can deduce from Proposition 2.2.2 that $u^{\rho_1} \geq u^{\rho_2}$, which together with Lemma 2.2.3 implies u^ρ converges to some function $u \in \mathbb{R}^{N\times d}$ as $\rho \to \infty$. Owing to the fact that the solution of (2.1.1) is unique for positive switching cost (see Proposition 2.2.1), it suffices to show u solves Problem 2.1.1.

Let $i \in \mathcal{I}$ be fixed. Since $\|u^\rho\| \leq \|F(0)\|/\gamma$, we see that $\sum_{j\neq i} \pi(u_l^{\rho,j} - c - u_l^{\rho,i}) \leq C/\rho$ for all $l \in \mathcal{N}$, with a constant C defined as:

$$C = \sup_{\|u\|\leq\|F(0)\|/\gamma} \|F(u)\| < \infty, \tag{2.2.3}$$

which is finite due to the continuity of F. Hence the limiting function u satisfies $u^i \geq \max_{j\neq i}(u^j - c)$ and $F_i(u) = \lim_{\rho\to\infty} F_i(u^\rho) \geq 0$. Moreover, suppose $u_l^i - (\mathcal{M}_i u)_l > 0$ at the index $l \in \mathcal{N}$, we can deduce that $F_i(u^\rho)_l = 0$ for all large enough ρ, which further implies $F_i(u)_l = 0$ and completes our proof. □

Remark 2.2.2. We point out that, unlike Problem 2.1.1, the well-posedness of Problem 2.2.1, and the monotone convergence of their solutions hold for any non-negative switching cost c, which enables us to study penalty schemes with zero switching cost (see Section 2.4).

Moreover, based on the penalized equations, Theorems 2.2.5 and 2.2.6 explicitly construct the solution to Problem 2.1.1, which is uniformly bounded by $\|F(0)\|/\gamma$ for all positive switching costs.

2.3 Penalization errors for positive switching cost

In this section, we shall proceed to analyze the convergence rate of the penalty approximation for Problem 2.1.1 with a fixed positive switching cost. As discussed in Section 2.1, it is not easy to construct a supersolution of Problem 2.1.1 from the solution of Problem 2.2.1 due to the non-diagonal dominance of the interconnected obstacles. We shall overcome this difficulty by regularizing the obstacles and establish the convergence rates of the penalty approximations with respect to the penalty parameter.

In order to obtain error estimates of the regularization procedures, we impose the following concavity condition on the monotone system:

Assumption 1. *The function F in Problem 2.1.1 is concave in the sense that: for any given $i \in \mathcal{I}$, $u, v \in \mathbb{R}^{N\times d}$, $\theta \in [0, 1]$, we have $F_i(\theta u + (1-\theta)v) \geq \theta F_i(u) + (1-\theta)F_i(v)$.*

Assumption 1 will only be used in Section 2.3 to quantify the regularization errors (not for the well-posedness or the monotone convergence of the regularization procedures). It is well-known that a concave function can be equivalently represented as the infimum of a family of affine functions, i.e., $F_i(u) = \inf_{\alpha\in\mathcal{A}_i} B_i(\alpha)u - b_i(\alpha)$ for some set $\mathcal{A}_i$ and coefficients $B_i : \mathcal{A}_i \to \mathbb{R}^{Nd\times Nd}$ and $b_i : \mathcal{A}_i \to \mathbb{R}^{Nd}$, hence our error estimates apply to the HJBQVIs studied in [21, 92, 93, 6, 39, 5]. However, our setting significantly extends the classical HJBQVIs in the following important aspects: (1) F_i can depend on all components of the solutions to the switching systems, (2) the control set $\mathcal{A}_i$ can be non-compact and coefficients B_i, b_i can be discontinuous, (3) b_i does not necessarily have a unique sign.

2.3.1 Regularizations of the QVIs

In this section we discuss how to approximate Problem 2.1.1 by variational inequalities with diagonally dominant obstacle terms. We shall propose two regularization procedures, namely an iterated optimal stopping approximation and a novel time-marching iteration, and estimate the regularization errors, which are essential for analyzing the penalization error of Problem 2.2.1.

Similar error estimates of the iterated optimal stopping approximation have been obtained in [20, 39] for continuous (scalar-valued) elliptic HJBQVIs with positive running costs, finite control sets, and sufficiently regular coefficients. Here, we relax these conditions and obtain regularization errors for general discrete monotone systems satisfying Assumption 1. The time-marching regularization leads to a more accurate approximation to Problem 2.1.1 than the iterated optimal stopping regularization, especially when the switching cost is small.

Let us start with the iterated optimal stopping approximation (see [92, 39] for its applications to the classical HJBQVIs), which approximates Problem 2.1.1 as follows: find $u^0 \in \mathbb{R}^{N\times d}$ satisfying $F_i(u^0) = 0$, $i \in \mathcal{I}$, and for each $n \geq 1$, given $u^{n-1} \in \mathbb{R}^{N\times d}$, find $u^n \in \mathbb{R}^{N\times d}$ such that $u^n = Qu^{n-1}$, where for any given $u \in \mathbb{R}^{N\times d}$, we define $Qu := ((Qu)^1, \dots, (Qu)^d) \in \mathbb{R}^{N\times d}$ to be the quantity which satisfies the following obstacle problem:

$$\min(F_i(Qu),\ (Qu)^i - \mathcal{M}_i u) = 0, \quad i \in \mathcal{I}. \tag{2.3.1}$$

By extending the arguments in Section 2.2, one can show that the above procedure is well-defined. Moreover, it is not difficult to establish the following comparison principle for (2.3.1): for any fixed $w \in \mathbb{R}^{N\times d}$, if $u \in \mathbb{R}^{N\times d}$ satisfies

$$\min(F_i(u),\ u^i - \mathcal{M}_i w) \leq 0, \quad i \in \mathcal{I},$$

and $v \in \mathbb{R}^{N\times d}$ satisfies

$$\min(F_i(v),\ v^i - \mathcal{M}_i w) \geq 0, \quad i \in \mathcal{I},$$

then $u \leq v$. In fact, for any fixed $w \in \mathbb{R}^{N\times d}$, we can show that the system $F_w = (F_{w,i})_{i\in\mathcal{I}}$ such that $F_{w,i}(u) := \min(F_i(u),\ u^i - \mathcal{M}_i w)$, $i \in \mathcal{I}$, satisfies the monotone condition (2.1.2) with the constant $\min(\gamma, 1)$, which subsequently implies the above comparison principle.

The next result presents some important properties of the operator Q.

Lemma 2.3.1. *The operator Q is monotone, i.e., $Qu \geq Qv$ provided that $u \geq v$, and satisfies the a priori estimate:*

$$\|Qu\| \leq \max(\|F(0)\|/\gamma,\ \|u\|) \quad \forall u \in \mathbb{R}^{N\times d}.$$

If we further suppose Assumption 1 holds, then the operator Q is convex.

Proof. If $u \geq v$, then $-\mathcal{M}_i u \leq -\mathcal{M}_i v$ due to the monotonicity of $\mathcal{M}_i$. Thus we have

$$\min(F_i(Qu),\ (Qu)^i - \mathcal{M}_i v) \geq \min(F_i(Qu),\ (Qu)^i - \mathcal{M}_i u) = 0, \quad i \in \mathcal{I},$$

which together with the comparison principle of (2.3.1) shows that $Qu \geq Qv$.

For the a priori estimate, we suppose that $|(Qu)^i_l| = \|Qu\|$. If $(Qu)^i_l = (\mathcal{M}_i u)_l$ and $(\mathcal{M}_i u)_l \geq 0$, then we have $\|Qu\| = (\mathcal{M}_i u)_l \leq \|u\|$. Otherwise, we can adapt the arguments of Lemma 2.2.3 to show $\|Qu\| \leq \|F(0)\|/\gamma$. Finally, for any given $u, v \in \mathbb{R}^{N\times d}$ and $\theta \in [0, 1]$, one can deduce from the concavity of F_i and $-\mathcal{M}_i$ that $\theta Qu + (1-\theta)Qv$ is a supersolution to (2.3.1) with the obstacle $\mathcal{M}_i(\theta u + (1-\theta)v)$, hence the comparison principle leads us to the desired result. □

The above lemma directly implies the monotone convergence of the iterates $(u^n)_{n\in\mathbb{N}}$.

Proposition 2.3.2. *For any given positive switching cost $c > 0$, the iterates $(u^n)_{n\in\mathbb{N}}$ satisfy $\|u^n\| \leq \|F(0)\|/\gamma$ for all $n \in \mathbb{N}$, and converge monotonically from below to the solution u of Problem 2.1.1 as $n \to \infty$.*

Proof. The bound of u^0 follows from Lemma 2.2.3 (with $\rho = 0$), while the uniform bound of $(u^n)_{n\in\mathbb{N}}$ follows from Lemma 2.3.1. Moreover, since $F(u^0) = 0$ and $F(u^1) \geq 0$ implies $u^1 \geq u^0$ by the comparison principle of F, we can show by an inductive argument and the monotonicity of the operator Q that $(u^n)_{n\in\mathbb{N}}$ monotonically increases to some vector u, which solves Problem 2.1.1 due to the continuity of (2.3.1). □

Now we proceed to estimate the difference $u - u^n$, where u and u^n solve Problem 2.1.1 and the equation (2.3.1), respectively. We shall first introduce the concept of strict supersolution, which was used in [54, 92, 93] to study impulse control problems.

Definition 2.3.1. A vector $w \in \mathbb{R}^{N\times d}$ is said to be a strict supersolution of Problem 2.1.1 if there exists a constant $\kappa > 0$, such that $G_i(w) = \kappa$ for all $i \in \mathcal{I}$.

For any any given $0 < \kappa < c$, by applying Theorem 2.2.5 to the problem

$$\min(F_i(u) - \kappa,\ u^i - \mathcal{M}_i^\kappa u) = 0, \quad \text{with } \mathcal{M}_i^\kappa u := \max_{j\neq i}(u^j - (c-\kappa)),$$

we can show that Problem 2.1.1 admits a unique strict supersolution satisfying the bound $\|w\| \leq (\|F(0)\| + \kappa)/\gamma$. For convenience, we shall assume without loss of generality that $\|F(0)\| > 0$ in the remaining part of this chapter, which excludes the trivial case where 0 is the unique solution to Problems 2.1.1 and 2.2.1.

The next lemma shows a contractive property of the operator Q. A similar result has been shown in [92] for a classical (continuous in time and space) HJBQVI via a control-theoretic approach. Here we shall present a simpler proof for our discrete setting based on the comparison principle, which can be easily extended to other regularization methods. For any given $\kappa \in (0, c)$, we introduce the following constant L_κ, which will be used frequently in the subsequent analysis:

$$L_\kappa := (2\|F(0)\| + \kappa)/\gamma. \tag{2.3.2}$$

Lemma 2.3.3. *Suppose Assumption 1 holds. Let w be the strict supersolution to Problem 2.1.1 with $\kappa \in (0, c)$. If $u^n - u^{n-1} \leq \lambda(w - u^{n-1})$ for some $\lambda \in [0, 1]$ and $n \in \mathbb{N}$, then we have $u^{n+1} - u^n \leq \lambda(1-\mu)(w - u^n)$ with*

$$\mu = \min\left(1, \frac{\gamma\kappa}{2\|F(0)\| + \kappa}\right) \in (0, 1]. \tag{2.3.3}$$

Proof. One can deduce from the convexity and monotonicity of the operator Q that

$$Qu^n \leq Q(\lambda w + (1-\lambda)u^{n-1}) \leq \lambda Qw + (1-\lambda)Qu^{n-1} = \lambda Qw + (1-\lambda)u^n,$$

hence it suffices to show $Qw \leq (1-\mu)w + \mu u^n$. Note that for any given $i \in \mathcal{I}$, we can obtain from the concavity of F_i that $F_i((1-\mu)w + \mu u^n) \geq 0$, and also for $\mu \in [0, 1]$,

$$(1-\mu)w^i + \mu u^{n,i} - \mathcal{M}_i w \geq \kappa - \mu(w^i + (u^{n,i})^-) \geq \kappa - \mu(\|w\| + \|(u^n)^-\|) \geq 0,$$

provided that $\mu \leq \frac{\kappa}{\|w\|+\|(u^n)^-\|}$. Since we have $\|w\| + \|(u^n)^-\| \leq L_\kappa$ for all $n \in \mathbb{N}$, setting $\mu = \min(1, \kappa/L_\kappa)$ gives us that

$$\min(F_i((1-\mu)w + \mu u^n),\ (1-\mu)w^i + \mu u^{n,i} - \mathcal{M}_i w) \geq 0, \quad i \in \mathcal{I},$$

which subsequently enables us to conclude the desired result from the comparison principle. □

Now we are ready to present the error estimate of the iterated optimal stopping approximation.

Theorem 2.3.4. *Suppose Assumption 1 holds, u solves Problem 2.1.1, and $(u^n)_{n\in\mathbb{N}}$ are recursively defined by the equation (2.3.1). Then we have for any given $\kappa \in (0,c)$ that*

$$0 \leq u - u^n \leq L_\kappa (1-\mu)^n/\mu, \quad \forall n \geq 0,^3$$

where μ and L_κ are defined as in (2.3.3) and (2.3.2) respectively.

Proof. Let w be the strict supersolution to Problem 2.1.1 with parameter κ. Since $(u^n)_{n\in\mathbb{N}}$ converge to u monotonically from below, the comparison principle for Problem 2.1.1 applied to u and w implies that $u^0 \leq u^1 \leq \cdots \leq u \leq w$. Hence $u^1 - u^0 \leq \lambda(w - u^0)$ with $\lambda = 1$, which along with Lemma 2.3.3 gives that $u^2 - u^1 \leq \lambda(1-\mu)(w-u^1)$. Inductively, we have

$$0 \leq u^{n+1} - u^n \leq (1-\mu)^n(w-u^n) \leq (1-\mu)^n(w-u^0), \quad n \geq 0.$$

Now summing the above inequality and employing $w - u^0 \leq L_\kappa$, we obtain the desired estimate:

$$0 \leq u - u^n = \sum_{k=n}^{\infty} u^{k+1} - u^k \leq (w-u^0)\sum_{k=n}^{\infty}(1-\mu)^k \leq L_\kappa(1-\mu)^n/\mu, \quad n \geq 0. \quad \square$$

Remark 2.3.1. Similar geometric convergence rates have been establish in [21, 39] for classical HJBQVIs, i.e., $F_i(u) = \inf_{\alpha\in\mathcal{A}_i} B_i(\alpha)u^i - b_i(\alpha)$ for all $i \in \mathcal{I}$, under the assumptions that $\mathcal{A}_i$ is a compact (or finite) set and $b_i(\alpha) \geq 0$ for all $\alpha \in \mathcal{A}_i$. Here we remove these restrictions.

Theorem 2.3.4 suggests the iterated optimal stopping only gives a good approximation to Problem 2.1.1 for sufficiently large switching cost c. For small enough switching cost c, we have $1-\mu \approx 1 - \gamma c/(2\|F(0)\|)$, which converges to 1 as $c \to 0$.

Due to the slow convergence rate of the iterated optimal stopping approximation for small switching cost, let us now discuss another regularization method, called the time-marching iteration (see [61]), which introduces an additional pseudo-time parameter ε to the interconnected obstacle, and gives an accurate approximation to Problem 2.1.1 even for small switching cost.

[3] Here we follow the convention that $0^0 = 1$.

For any given parameter $\varepsilon > 0$, the time-marching iteration is given as follows: find $u^0 \in \mathbb{R}^{N\times d}$ satisfying $F_i(u^0) = 0$, $i \in \mathcal{I}$, and for each $n \geq 1$, given $u^{n-1} \in \mathbb{R}^{N\times d}$, find $u^n \in \mathbb{R}^{N\times d}$ such that $u^n = Tu^{n-1}$, where for any given $u \in \mathbb{R}^{N\times d}$, we define $Tu := ((Tu)^1, \dots, (Tu)^d) \in \mathbb{R}^{N\times d}$ to be the quantity which satisfies the following obstacle problem:

$$\min\big(F_i(Tu),\ (Tu)^i - \mathcal{M}_i(Tu) + \varepsilon((Tu)^i - u^i)\big) = 0, \quad i \in \mathcal{I}. \tag{2.3.4}$$

The operator T enjoys properties analogous to the operator Q, i.e., Lemma 2.3.1 and Proposition 2.3.2, whose proofs are similar and details are omitted. Moreover, similar to (2.3.1), we can show (2.3.4) admits the following comparison principle: for any fixed $w \in \mathbb{R}^{N\times d}$, if $u \in \mathbb{R}^{N\times d}$ satisfies

$$\min\big(F_i(u),\ u^i - \mathcal{M}_i u + \varepsilon(u^i - w^i)\big) \leq 0, \quad i \in \mathcal{I}.$$

and $v \in \mathbb{R}^{N\times d}$ satisfies

$$\min\big(F_i(v),\ v^i - \mathcal{M}_i v + \varepsilon(v^i - w^i)\big) \geq 0, \quad i \in \mathcal{I}.$$

then $u \leq v$. The next theorem presents the convergence rate of the time-marching iteration.

Theorem 2.3.5. *Suppose Assumption 1 holds, u solves Problem 2.1.1, and $(u^n)_{n\in\mathbb{N}}$ are recursively defined by the equation (2.3.4) with the pseudo-time parameter $\varepsilon > 0$. Then we have for any $\kappa \in (0, c)$ that*

$$0 \leq u - u^n \leq L_\kappa(1-\mu)^n/\mu, \quad n \geq 0,$$

where L_κ is defined as in (2.3.2) and $\mu = \kappa/(\kappa + \varepsilon L_\kappa)$.

Proof. Let w be the strict supersolution to Problem 2.1.1 with parameter κ. Following the proofs of Lemma 2.3.3 and Theorem 2.3.4, we see it is essential to obtain $\mu \in (0, 1]$ such that $Tw \leq (1-\mu)w + \mu u^n$ for all $n \in \mathbb{N}$, which leaves us to show that for suitable u we have

$$\begin{aligned}
&(1-\mu)w^i + \mu u^{n,i} - \mathcal{M}_i[(1-\mu)w + \mu u^n] + \varepsilon((1-\mu)w^i + \mu u^{n,i} - w^i) \\
\geq&(1-\mu)(w^i - \mathcal{M}_i w) + \mu[u^{n,i} - \mathcal{M}_i u^n + \varepsilon(u^{n,i} - u^{n-1,i})] - \varepsilon\mu(w^i - u^{n-1,i}) \\
\geq&(1-\mu)\kappa - \varepsilon\mu(\|w\| + \|(u^{n-1})^-\|) \geq (1-\mu)\kappa - \varepsilon\mu L_\kappa \geq 0,
\end{aligned}$$

for L_κ defined as in (2.3.2). Thus by setting $\mu = \kappa/(\kappa + \varepsilon L_\kappa)$, we see the above inequality holds, and one can deduce the desired result following similar arguments as the proof of Theorem 2.3.4. □

Remark 2.3.2. Through the choice of the pseudo-time parameter ε, the time-marching iteration gives a more accurate approximation to Problem 2.1.1 than the iterated optimal stopping approximation for small switching cost c. In fact, it holds for the time-marching iteration that

$$1-\mu = 1-\frac{\kappa}{\kappa+\varepsilon(2\|F(0)\|+\kappa)/\gamma} = 1-\frac{1}{1+2\varepsilon\|F(0)\|/(\kappa\gamma)+\varepsilon/\gamma}.$$

Therefore, taking $c \to 0$ and $\varepsilon \to 0$ such that $\varepsilon/c \to 0$, we get $1-\mu \to 0$. However, as we shall see in Section 2.3.2, the error of the penalty approximation to (2.3.4) grows proportionally to $1/\varepsilon$, hence after minimizing over ε, both the iterated optimal stopping approximation and the time-marching iteration lead to the same error estimate for Problem 2.2.1.

2.3.2 Convergence order of penalty methods

In this section, we shall use the regularization procedures proposed in Section 2.3.1 to demonstrate that for fixed positive switching cost and a penalty function with degree $\sigma > 0$, the approximation error of Problem 2.2.1 is bounded above by the quantity $C_0\rho^{-\sigma}\ln\rho$ for some constant C_0, which depends only on the function F and is independent of the number of switching regimes d. Since both regularization procedures lead to the same error estimate (see Remark 2.3.2), we shall focus on the regularization by the iterated optimal stopping, and only outline the essential results for the time-marching iteration.

To quantify the penalty error of Problem 2.2.1 with a fixed parameter $\rho \geq 0$, we introduce the following sequence of auxiliary problems: find $u^{\rho,0} \in \mathbb{R}^{N\times d}$ satisfying $F_i(u^0) = 0$, $i \in \mathcal{I}$, and for each $n \geq 1$, given $u^{\rho,n-1} \in \mathbb{R}^{N\times d}$, find $u^{\rho,n} \in \mathbb{R}^{N\times d}$ such that $u^{\rho,n} = Q^\rho u^{\rho,n-1}$, where for any given $u \in \mathbb{R}^{N\times d}$, we define $Q^\rho u := ((Q^\rho u)^1, \dots, (Q^\rho u)^d) \in \mathbb{R}^{N\times d}$ to be the quantity which satisfies the following penalized equation:

$$F_i(Q^\rho u) - \rho\sum_{j\neq i}\pi(u^j - c - (Q^\rho u)^i) = 0, \quad i \in \mathcal{I}. \tag{2.3.5}$$

Note that the above auxiliary problem has the same initialization as the iterated optimal stopping approximation to Problem 2.1.1, i.e., $u^{\rho,0} = u^0$. Moreover, for any fixed $w \in \mathbb{R}^{N\times d}$, we can consider the system $G_w = (G_{w,i})_{i\in\mathcal{I}}$ such that $G_{w,i}(u) := F_i(u) - \rho\sum_{j\neq i}\pi(w^j - c - u^i)$, $i \in \mathcal{I}$, which satisfies the monotonicity condition (2.1.2) with the constant γ due to the facts that F satisfies the monotonicity condition (2.1.2) with γ and π is non-decreasing. Consequently, we have the following comparison principle for (2.3.5): for any fixed $w \in \mathbb{R}^{N\times d}$, if $u \in \mathbb{R}^{N\times d}$ satisfies $G_{w,i}(u) \le 0$, $i \in \mathcal{I}$, and $v \in \mathbb{R}^{N\times d}$ satisfies $G_{w,i}(v) \ge 0$, $i \in \mathcal{I}$, then $u \le v$. Therefore, we can easily establish the well-posedness of (2.3.5) by adapting the proof of Theorem 2.2.5.

The following result summarizes the essential properties of the operator Q^ρ and the iterates $(u^{\rho,n})_{n\in\mathbb{N}}$.

Proposition 2.3.6. *The operator Q^ρ is monotone, satisfies the a priori estimate:*

$$\|Q^\rho u\| \le \max(\|F(0)\|/\gamma, \|u\|) \quad \forall u \in \mathbb{R}^{N\times d},$$

and is Lipschitz continuous with constant 1, i.e., $\|Q^\rho u - Q^\rho v\| \le \|u - v\|$ for all $u, v \in \mathbb{R}^{N\times d}$. Consequently, for any given penalty parameter $\rho \ge 0$ and switching cost $c \ge 0$, the iterates $(u^{\rho,n})_{n\in\mathbb{N}}$ converge monotonically from below to the solution u^ρ of Problem 2.2.1 as $n \to \infty$.

Proof. The a priori bound can be obtain exactly as Lemma 2.3.1. For the monotonicity and Lipschitz continuity of Q^ρ, it suffices to show for any given $u, v \in \mathbb{R}^{N\times d}$, we have $Q^\rho u - Q^\rho v \le \|(u - v)^+\|$.

For any given $u, v \in \mathbb{R}^{N\times d}$, we introduce the quantity $\hat{u} := Q^\rho v + \|(u - v)^+\|$. It is important to observe that for any given $L \ge 0$ and $u \in \mathbb{R}^{N\times d}$, the monotonicity (2.1.2) of F implies

$$F_i(u + L) - F_i(u) \ge \gamma L \quad \forall i \in \mathcal{I}, \tag{2.3.6}$$

which along with the fact that

$$\pi(u^j - c - \hat{u}^i) = \pi(u^j - \|(u - v)^+\| - c - (Q^\rho v)^i) \le \pi(v^j - c - (Q^\rho v)^i), \quad j \ne i,$$

enables us to conclude the desired result through the following estimate: for any $i \in \mathcal{I}$,

$$F_i(\hat{u}) - \rho \sum_{j\ne i} \pi(u^j - c - \hat{u}^i) \ge F_i(Q^\rho v) + \gamma\|(u - v)^+\| - \rho \sum_{j\ne i} \pi(v^j - c - (Q^\rho v)^i) \ge 0.$$

The above implies that $\hat{u}$ is a supersolution to (2.3.5) with the input u. By the comparison principle for (2.3.5), $Q^\rho u \le \hat{u}$ and hence $Q^\rho u - Q^\rho v \le \|(u - v)^+\|$ as desired. Then the monotone convergence of $(u^{\rho,n})_{n\in\mathbb{N}}$ follows from similar arguments as those in Proposition 2.3.2. □

The next result provides an upper bound of the term $u^n - u^{\rho,n}$, where u^n and $u^{\rho,n}$ solve the equations (2.3.1) and (2.3.5), respectively.

Proposition 2.3.7. *For any given penalty parameter $\rho \geq 0$ and switching cost $c > 0$, let $(u^n)_{n\in\mathbb{N}}$ and $(u^{\rho,n})_{n\in\mathbb{N}}$ be recursively defined by the equations (2.3.1) and (2.3.5), respectively. Suppose that there exist positive constants τ and σ such that $\pi(y) \geq \tau y^{1/\sigma}$ for all $0 \leq y \leq 2\|F(0)\|/\gamma$. Then we have*

$$\|u^n - u^{\rho,n}\| \leq \left(\frac{C}{\tau\rho}\right)^{\sigma} n, \quad n \geq 0,$$

where $C = \sup_{\|u\|\leq\|F(0)\|/\gamma} \|F(u)\|$.

Proof. The Lipschitz continuity of Q^ρ implies that for any $n \in \mathbb{N}$,

$$\|u^n - u^{\rho,n}\| = \|u^n - Q^\rho u^{n-1}\| + \|Q^\rho u^{n-1} - Q^\rho u^{\rho,n-1}\| \leq \|u^n - Q^\rho u^{n-1}\| + \|u^{n-1} - u^{\rho,n-1}\|. \tag{2.3.7}$$

Now we bound $u^n - Q^\rho u^{n-1}$ for any given $n \in \mathbb{N}$. From the a priori bounds of u^n (Proposition 2.3.2) and Q^ρ (Proposition 2.3.6), we know $\|Q^\rho u^{n-1}\| \leq \|F(0)\|/\gamma$ for all $\rho \geq 0$ and $n \in \mathbb{N}$. Moreover, by using the comparison principle for (2.3.5) and a modification of the arguments in Theorem 2.2.6, we can deduce for any given $n \in \mathbb{N}$ that $(Q^\rho u^{n-1})_{\rho>0}$ converges monotonically from below to u^n as $\rho \to \infty$. This implies that $\|\rho\pi(u^{n-1,j} - c - (Q^\rho u^{n-1})^i)\| \leq C$ for all $j \neq i$, $i \in \mathcal{I}$, where C is defined as in (2.2.3). Therefore, we have

$$\begin{aligned}
(Q^\rho u^{n-1})^i + \left(\frac{C}{\tau\rho}\right)^{\sigma} - \mathcal{M}_i u^{n-1} &= \min_{j\neq i} \frac{1}{\rho^\sigma}\left(\left(\frac{C}{\tau}\right)^{\sigma} - \rho^\sigma(u^{n-1,j} - c - (Q^\rho u^{n-1})^i)\right) \\
&\geq \frac{1}{(\rho\tau)^\sigma} \min_{j\neq i}\left(C^\sigma - \|\rho\pi(u^{n-1,j} - c - (Q^\rho u^{n-1})^i)\|^\sigma\right) \geq 0.
\end{aligned}$$

Moreover, by applying (2.3.6) with $u = Q^\rho u^{n-1}$ and $L = (C/(\tau\rho))^\sigma$, we can obtain that

$$F_i\left(Q^\rho u^{n-1} + \left(\frac{C}{\tau\rho}\right)^{\sigma}\right) \geq F_i(Q^\rho u^{n-1}) + \gamma\left(\frac{C}{\tau\rho}\right)^{\sigma} \geq 0, \quad \forall i \in \mathcal{I},$$

which implies $Q^\rho u^{n-1} + \big(C/(\tau\rho)\big)^\sigma$ is a supersolution to (2.3.1). Consequently, we obtain $0 \leq u^n - Q^\rho u^{n-1} \leq (C/(\tau\rho))^\sigma$ for all $n \in \mathbb{N}$, and conclude the desired result from $u^{\rho,0} = u^0$ and (2.3.7). ☐

Remark 2.3.3. This proposition greatly extends the results in [102] (even for the case with $\sigma = 1$) by removing the continuous differentiability assumption of the penalty function π. In practice, one can choose $\pi(y) = (y^+)^{1/\sigma}$ as the penalty function. Since π is semismooth if $\sigma = 1$, a direct application of semismooth Newton methods allows us to solve Problem 2.2.1 efficiently (see [102, 87]). A penalty term with $\sigma > 1$ needs an additional smoothing for the application of Newton methods and usually requires a larger number of Newton iterations to solve the penalized equation [51]. Though the higher convergence rate allows us to use a relatively small value of ρ to achieve the desired accuracy, which could avoid the numerical instability caused by the usage of a large penalty parameter according to [51], we do not discover any problem using the penalty function $\pi(y) = y^+$ in our numerical experiments.

Now we are ready to conclude the penalty error of Problem 2.2.1 to Problem 2.1.1. The following result has been proved in [20, Lemma 6.1], and will be used in our error estimates.

Lemma 2.3.8. *Let* $\phi : \mathbb{R} \to \mathbb{R}$, $\phi(x) = \nu a^x + bx$, *where* $0 < a < 1$, $0 < b < \infty$ *and* $\nu > 0$. *Let* $m := \min_{n\in\mathbb{N}} \phi(n)$. *Then we have*

$$m \le \begin{cases} \nu, & -b/(\nu \ln a) \ge 1, \\ -ab/(\ln a) + b[\log_a(b/(\nu \ln a)) + 1], & \text{otherwise.} \end{cases}$$

Theorem 2.3.9. *For any given switching cost* $c > 0$, *let* u *and* u^ρ *solve Problem 2.1.1 and 2.2.1, respectively. Suppose that Assumption 1 and the assumptions in Proposition 2.3.7 hold. Then if* $c > 2\|F(0)\|/\gamma$, *we have* $u^\rho = u$ *for all* $\rho \ge 0$, *and if* $c \le 2\|F(0)\|/\gamma$, *we have for any* $\kappa \in (0, c)$,

$$\|u - u^\rho\| \le f(\rho), \quad \text{where} \quad f \sim -\frac{\sigma(C/\tau)^\sigma}{\ln(1 - \kappa/L_\kappa)} \rho^{-\sigma} \ln \rho, \quad \text{as } \rho \to \infty,^4$$

with the constants C *and* L_κ *defined as in* (2.2.3) *and* (2.3.2) *respectively.*

Proof. Suppose that $c > 2\|F(0)\|/\gamma$. Theorem 2.2.5 shows that $\|u^\rho\| \le \|F(0)\|/\gamma$ for all $\rho \ge 0$, which implies that

$$u_l^{\rho,j} - c - u_l^{\rho,i} \le 2\|F(0)\|/\gamma - c < 0, \quad \forall i \in \mathcal{I},\ j \ne i,\ l \in \mathcal{N}.$$

Thus we have $G_i^\rho(u^\rho) = F_i(u^\rho) = 0$ for all $i \in \mathcal{I}$ and $\rho \ge 0$. Similarly, we can obtain by using $\|u\| \le \|F(0)\|/\gamma$ (see Remark 2.2.2) that $u^i - \mathcal{M}_i u > 0$, which implies that $F_i(u) = 0$ for all $i \in \mathcal{I}$. The comparison principle of F gives us that $u = u^\rho$ for all $\rho \ge 0$.

Now we assume that $c \le 2\|F(0)\|/\gamma$. Since $u^\rho \le u$ for all $\rho \ge 0$, it remains to derive an upper bound of $u - u^\rho$. Note that

$$u - u^\rho \le u - u^n + u^n - u^{\rho,n} + u^{\rho,n} - u^\rho,$$

[4] Recall that $f \sim g$ as $\rho \to \infty$ if $g(\rho) \ne 0$ for all $\rho > 0$ and $\lim_{\rho\to\infty} f(\rho)/g(\rho) = 1$.

where u^n and $u^{\rho,n}$ are recursively defined by the equations (2.3.1) and (2.3.5), respectively. Proposition 2.3.6 implies that $u^{\rho,n} \leq u^\rho$ for all $\rho \geq 0$ and $n \in \mathbb{N} \cup \{0\}$. Hence we can deduce from Theorem 2.3.4 and Proposition 2.3.7 that for any $\kappa \in (0, c)$,

$$\|u - u^\rho\| \leq L_\kappa \frac{(1-\mu)^n}{\mu} + \left(\frac{C}{\tau\rho}\right)^\sigma n \quad \forall n \in \mathbb{N} \cup \{0\}, \tag{2.3.8}$$

where C and L_κ are defined as in (2.2.3) and (2.3.2) respectively, and $\mu = \min(1, \kappa/L_\kappa) < 1$ due to the assumption that $c \leq 2\|F(0)\|/\gamma$. Now we minimize the right-hand side of (2.3.8) over n by applying Lemma 2.3.8 with $\nu = L_\kappa/\mu$, $a = 1-\mu$ and $b = \big(C/(\tau\rho)\big)^\sigma$. If ρ is sufficiently large, then $-b/(\nu \ln a) < 1$, which implies that

$$\|u - u^\rho\| \leq -\frac{1}{\ln(1-\mu)}\left(\frac{C}{\tau\rho}\right)^\sigma \left[1 - \mu - \ln(1-\mu) - \ln\left(\frac{\mu}{L_\kappa \ln(1-\mu)}\left(\frac{C}{\tau\rho}\right)^\sigma\right)\right].$$

Thus by using the following identity:

$$-\ln\left(\frac{\mu}{L_\kappa \ln(1-\mu)}\left(\frac{C}{\tau\rho}\right)^\sigma\right) = -\ln\left(\frac{\mu}{L_\kappa \ln(1-\mu)}\left(\frac{C}{\tau}\right)^\sigma\right) + \sigma \ln \rho,$$

we deduce that $\|u - u^\rho\| \leq f(\rho)$, where f satisfies that $f \sim -\frac{\sigma(C/\tau)^\sigma}{\ln(1-\mu)}\rho^{-\sigma}\ln\rho$, as $\rho \to \infty$. □

Remark 2.3.4. Recall that $\mu = \mathcal{O}(c)$ as $c \to 0$, hence the upper bound behaves as $\rho^{-\sigma}\ln\rho/c$ for small switching cost c. Unfortunately, we are not sure whether this dependence on c is optimal since the possible blow-up of the penalization error for small enough c could be due to the fact that the iterated optimal stopping approximation does not provide an accurate approximation to Problem 2.1.1 with small switching cost. Our numerical experiments show that as the switching cost tends to zero, the penalization error with a fixed penalty parameter indeed grows at a rate $\mathcal{O}(c^{-1/2})$ for certain ranges of switching costs, but then stablizes to a limiting value (see Section 2.5). As we shall see in Section 2.4, Problem 2.1.1 degenerates into an HJB equation as the switching cost $c \to 0$, and Problem 2.2.1 with $c = 0$ provides a penalty approximation to such equation, with an asymptotic error $\mathcal{O}(1/\rho^\sigma)$ as the penalty parameter $\rho \to \infty$. Thus the penalization error with sufficiently small positive switching cost is dominated by this limiting error (see (2.4.2)).

We proceed to outline the key results of the convergence analysis by using the time-marching iteration, and demonstrate that even if the time-marching iteration could improve the regularization error for small switching cost by adjusting the pseudo-time parameter ε, it reduces the accuracy of penalty approximations at each iterate. Hence it leads to the same error estimate of Problem 2.2.1 as the iterated optimal stoping approximation for small switching cost.

For any given parameters $\rho \geq 0$ and $\varepsilon \geq 0$, we introduce the following sequence of auxiliary problems: find $u^{\rho,0} \in \mathbb{R}^{N\times d}$ satisfying $F_i(u^0) = 0$, $i \in \mathcal{I}$, and for each $n \geq 1$, given $u^{\rho,n-1} \in \mathbb{R}^{N\times d}$, find $u^{\rho,n} \in \mathbb{R}^{N\times d}$ such that $u^{\rho,n} = T^\rho u^{\rho,n-1}$, where for any given $u \in \mathbb{R}^{N\times d}$, we define $T^\rho u \coloneqq ((T^\rho u)^1, \dots, (T^\rho u)^d) \in \mathbb{R}^{N\times d}$ to be the quantity which satisfies the following penalized equation:

$$F_i(T^\rho u) - \rho \sum_{j\neq i} \pi\big((T^\rho u)^j - c - (T^\rho u)^i - \varepsilon((T^\rho u)^i - u^i)\big) = 0, \quad i \in \mathcal{I}.$$

One can establish analogue results of Proposition 2.3.6 for the operator T^ρ, and demonstrate that $T^\rho u^{n-1} + (C_1/\rho)^\sigma/\varepsilon$ is a supersolution to (2.3.4) under the same assumptions of Proposition 2.3.7. Therefore, following the proof of Theorem 2.3.9, we deduce from Theorem 2.3.5 that for any $\kappa \in (0, c)$,

$$\|u - u^\rho\| \leq L_\kappa \frac{(1-\mu)^n}{\mu} + \frac{1}{\varepsilon}\left(\frac{C}{\tau\rho}\right)^\sigma n \quad \forall n \in \mathbb{N},\, \varepsilon > 0,$$

where C and L_κ are defined as in (2.2.3) and (2.3.2) respectively, and $\mu = \kappa/(\kappa+\varepsilon L)$. Minimizing over n, we obtain $\|u - u^\rho\| = \mathcal{O}(-\frac{\rho^{-\sigma}\ln\rho}{\varepsilon\ln(1-\frac{c}{c+\varepsilon L_\kappa})})$ for all $\varepsilon > 0$ and large enough ρ. Therefore, by discussing the cases $\varepsilon = \mathcal{O}(c)$ and $\varepsilon/c \to \infty$ separately, we arrive at the same error estimate as that in Theorem 2.3.9.

Finally we end this section with an exact construction of the optimal switching regions

$$\Gamma_i \coloneqq \{l \in \mathcal{N} \mid u_l^i = (\mathcal{M}_i u)_l\}, \quad i \in \mathcal{I}, \tag{2.3.9}$$

of Problem 2.1.1 with a given switching cost $c > 0$ using the solution of Problem 2.2.1. Suppose the estimate $0 \leq u - u^\rho \leq C_0\rho^{-\sigma}\ln\rho$ holds for some constants $C_0, \sigma > 0$, where σ is the degree of the penalty function π and C_0 in practice can be estimated using numerical results. Then we shall define the sets

$$\Gamma_{\rho,i} \coloneqq \{l \in \mathcal{N} \mid |u_l^{\rho,i} - (\mathcal{M}_i u^\rho)_l| \leq C_0\rho^{-\sigma}\ln\rho\}, \quad i \in \mathcal{I}. \tag{2.3.10}$$

The next result demonstrates that $\Gamma_{\rho,i}$ in fact coincides with Γ_i for large enough ρ.

Theorem 2.3.10. *Suppose that there exist positive constants C_0 and σ such that the estimate $0 \leq u - u^\rho \leq C_0\rho^{-\sigma}\ln\rho$ holds for all $\rho > 0$. For each $i \in \mathcal{I}$, let Γ_i and $\Gamma_{\rho,i}$ be the sets defined as in (2.3.9) and (2.3.10), respectively. Then for a given switching cost $c > 0$, there exists $\rho_0 > 0$ such that $\Gamma_i = \Gamma_{\rho,i}$ for all $\rho \geq \rho_0$ and $i \in \mathcal{I}$.*

Proof. We first show that $\Gamma_i \subset \Gamma_{\rho,i}$ for all $\rho > 0$ and $i \in \mathcal{I}$. For any fixed i, we can deduce from the estimate $u \leq u^\rho + C_0\rho^{-\sigma} \ln \rho$ and the monotonicity of $\mathcal{M}_i$ that

$$\mathcal{M}_i u \leq \mathcal{M}_i(u^\rho + C_0\rho^{-\sigma} \ln \rho) = \mathcal{M}_i u^\rho + C_0\rho^{-\sigma} \ln \rho.$$

Now let l be an arbitrary element of Γ_i so that $u_l^i = (\mathcal{M}_i u)_l$. It follows that

$$\begin{aligned} u_l^{\rho,i} &\leq u_l^i = (\mathcal{M}_i u)_l \leq \mathcal{M}_i u^\rho + C_0\rho^{-\sigma} \ln \rho \\ u_l^{\rho,i} &\geq u_l^i - C_0\rho^{-\sigma} \ln \rho \geq (\mathcal{M}_i u^\rho)_l - C_0\rho^{-\sigma} \ln \rho, \end{aligned}$$

which implies that l is in $\Gamma_{\rho,i}$.

Suppose the statement of Theorem 2.3.10 does not hold, then by using the finiteness of $\mathcal{N}$ and the pigeonhole principle, there exists a sequence $\{\rho_n\}$ such that $\rho_n \to \infty$ as $n \to \infty$, and an index $l \in \Gamma_{\rho_n,i} \setminus \Gamma_i$ for all n. However, the definition of $\Gamma_{\rho_n,i}$ implies that

$$\begin{aligned} u_l^i - (\mathcal{M}_i u)_l &= u_l^i - u_l^{\rho_n,i} + u_l^{\rho_n,i} - (\mathcal{M}_i u^{\rho_n})_l + (\mathcal{M}_i u^{\rho_n})_l - (\mathcal{M}_i u)_l \\ &\leq C_0\rho_n^{-\sigma} \ln \rho_n + C_0\rho_n^{-\sigma} \ln \rho_n + 0 \to 0, \end{aligned}$$

which together with $u_l^i \geq (\mathcal{M}_i u)_l$ leads to $l \in \Gamma_i$, and hence a contradiction. □

2.4 Penalization errors for vanishing switching cost

In this section, we investigate the asymptotic behaviours of Problems 2.1.1 and 2.2.1 as the switching cost c tends to zero. We shall show that the system in Problem 2.1.1 degenerates into a single equation of HJB type, and establish that the penalty error of Problem 2.2.1 with zero switching cost is of the same order (up to a log factor) as that in Theorem 2.3.9.

Throughout this section, to emphasize the dependence on c, we shall denote by u^c and $u^{c,\rho}$ the solutions to Problems 2.1.1 and 2.2.1 with a given positive switching cost c, respectively, and by u^ρ the solution to Problem 2.2.1 with $c = 0$. Moreover, to identify the limiting behaviour of Problems 2.1.1 and 2.2.1, we introduce the following regularity condition on the monotone system F:

Assumption 2. *The function F in Problem 2.1.1 is locally Lipschitz continuous.*

We emphasize that even though Assumption 1 is a sufficient condition for Assumption 2, it will not be used in this section. In particular, Assumption 2 is general enough to cover non-convex/non-concave equations, such as Isaacs equations.

We first introduce the degenerate problem for zero switching cost.

Problem 2.4.1. Find $u \in \mathbb{R}^N$ such that the vector $\mathbf{u} = (u, \ldots, u) \in \mathbb{R}^{N\times d}$ satisfies

$$\min_{i\in\mathcal{I}} F_i(\mathbf{u}) = 0. \tag{2.4.1}$$

For the classical HJBQVIs where $F_i(u) = \inf_{\alpha\in\mathcal{A}_i} B(\alpha)u^i - b(\alpha)$ for all $i \in \mathcal{I}$, Problem 2.4.1 can be equivalently written as an HJB equation as studied in [101, 102]: find $u \in \mathbb{R}^N$ satisfying $\inf_{\alpha\in\mathcal{A}} B(\alpha)u - b(\alpha) = 0$ with $\mathcal{A} = \cup_{i\in\mathcal{I}}\mathcal{A}_i$. However, we reiterate that in this work, F_i can depend on all components of the system, and is not assumed to be concave in this section. Moreover, even for concave equations, both the penalty scheme, i.e., Problem 2.2.1 with $c = 0$, and its error analysis are essentially different from those in [101, 102].

By using the monotonicity condition (2.1.2), one can easily establish the following comparison principle for Problem 2.4.1, i.e., if $\mathbf{u} = (u, \ldots, u) \in \mathbb{R}^{N\times d}$ and $\mathbf{v} = (v, \ldots, v) \in \mathbb{R}^{N\times d}$ satisfy $\min_{i\in\mathcal{I}} F_i(\mathbf{u}) \le 0$ and $\min_{i\in\mathcal{I}} F_i(\mathbf{v}) \ge 0$ respectively, then $u \le v$, which subsequently implies the uniqueness of solutions to Problem 2.4.1. We shall now demonstrate that the solution to Problem 2.4.1 can be identified as the limit of the solutions to Problem 2.1.1 with vanishing switching cost.

Proposition 2.4.1. *Let u^c solve Problem 2.1.1 with a switching cost $c > 0$. Then $(u^c)_{c>0}$ converges monotonically from below to the solution $\mathbf{u}$ of Problem 2.4.1 as $c \to 0$.*

Proof. It is easy to check that if $c_1 > c_2 > 0$, then u^{c_2} is a supersolution to Problem 2.1.1 with a switching cost c_1. Hence the comparison principle and the a priori bound (see Remark 2.2.2) imply that for each $i \in \mathcal{I}$, $(u^{c,i})_{c>0}$ converges monotonically from below to some vector $\bar{u}^i \in \mathbb{R}^N$ as $c \to 0$. Moreover, since $u^{c,i} \ge \mathcal{M}_i u^c \ge u^{c,j} - c$ for all $j \neq i$, $c > 0$, we have $\bar{u}^i \equiv \bar{u}$ for all $i \in \mathcal{I}$.

We now show $\mathbf{u} := (\bar{u}, \ldots, \bar{u})$ solves Problem 2.4.1. For any given $i \in \mathcal{I}$, using the supersolution property of u^c, we have $F_i(u^c) \ge 0$ for all $c > 0$. Hence the continuity of F implies that $\min_{i\in\mathcal{I}} F_i(\mathbf{u}) \ge 0$. On the other hand, let $l \in \mathcal{N}$ be a fixed index. For any given $c > 0$, we consider the component $i_{l,c}$ where $u_l^{c,i_{l,c}} = \max_{j\in\mathcal{I}} u_l^{c,j} > (\mathcal{M}_{i_{l,c}} u^c)_l$, and consequently $F_{i_{l,c}}(u^c)_l = 0$. As $c \to 0$, since $\mathcal{I}$ is a finite set, by passing to a subsequence, we can assume there exists $\{c_n\} \to 0$ as $n \to \infty$, and a component $i_l \in \mathcal{I}$ such that $F_{i_l}(u^{c_n})_l = 0$ for all $n \in \mathbb{N}$. Thus letting $c_n \to 0$, we have $\min_{i\in\mathcal{I}} F_i(\mathbf{u})_l \le F_{i_l}(\mathbf{u})_l = 0$. Since l is an arbitrary index, we conclude $\min_{i\in\mathcal{I}} F_i(\mathbf{u}) = 0$. □

Because Problem 2.4.1 is the limiting equation of Problem 2.1.1 as $c \to 0$, we now analyze the approximation error of Problem 2.2.1 with $c = 0$ to Problem 2.4.1, which indicates the asymptotic behaviour of the penalization error of Problem 2.2.1 for small enough switching cost.

Theorem 2.4.2. *The solution u^ρ of Problem 2.2.1 (with $c = 0$) converges monotonically from below to the solution $\mathbf{u}$ of Problem 2.4.1 as $\rho \to \infty$. Moreover, if we further assume Assumption 2 holds and there exist positive constants τ and σ such that $\pi(y) \geq \tau y^{1/\sigma}$ for all $0 \leq y \leq 2\|F(0)\|/\gamma$, then we have the following error estimate:*

$$0 \leq \mathbf{u} - u^\rho \leq C_1/\rho^\sigma,$$

for some constant $C_1 > 0$, independent of the penalty parameter ρ.

Proof. By Theorem 2.2.6, $(u^\rho)_{\rho \geq 0}$ converge monotonically from below to some element $\mathbf{u} = (\bar{u}^1, \ldots, \bar{u}^d) \in \mathbb{R}^{N \times d}$. Since it holds that

$$0 = \lim_{\rho \to \infty} \left\{ F_i(u^\rho) - \rho \sum_{j \neq i} \pi(u^{\rho,j} - u^{\rho,i}) \right\} \leq \lim_{\rho \to \infty} F_i(u^\rho) = F_i(\mathbf{u}),$$

we know $\bar{u}^1 = \ldots = \bar{u}^d$ (otherwise the first limit above would blow up). It remains to establish that $\min_{i \in \mathcal{I}} F_i(\mathbf{u}) \leq 0$. To do so, we pick, for each $\rho > 0$ and $l \in \mathcal{N}$, a component $i_{l,\rho}$ such that $u_l^{\rho, i_{l,\rho}} = \max_{j \in \mathcal{I}} u_l^{\rho,j}$, so that $F_{i_{l,\rho}}(u^\rho)_l = 0$. The desired result is then established by passing to a subsequence as in the proof of Proposition 2.4.1.

The fact that $u^\rho \leq \mathbf{u}$ implies that it suffices to show there exists a constant C_1, such that for each $i \in \mathcal{I}$ and $\rho > 0$, $(u^{\rho,i}, \ldots, u^{\rho,i}) + C_1/\rho^\sigma \in \mathbb{R}^{N \times d}$ is a supersolution to (2.4.1). Note that $(u^\rho)_{\rho \geq 0}$ are bounded by $\|F(0)\|/\gamma$ (see Lemma 2.2.3), hence we have

$$\sum_{j \neq i} \tau [(u^{\rho,j} - u^{\rho,i})^+]^{1/\sigma} \leq \sum_{j \neq i} \pi(u^{\rho,j} - u^{\rho,i}) \leq C/\rho, \quad \forall i \in \mathcal{I},\, j \neq i,$$

where C is defined as in (2.2.3). Then using the local Lipschitz continuity of F, we obtain

$$\|F_j(u^\rho) - F_j(\hat{\mathbf{u}}^i)\| \leq L_{\text{lip}} \|u^\rho - \hat{\mathbf{u}}^i\| \leq L_{\text{lip}} \left(\frac{C}{\tau \rho} \right)^\sigma,$$

where $\hat{\mathbf{u}}^i := (u^{\rho,i}, \ldots, u^{\rho,i})$. Therefore, by using the inequality (2.3.6) and setting $\gamma C_1 = L_{\text{lip}} (C/\tau)^\sigma$, one can conclude $\hat{\mathbf{u}}^i + C_1/\rho^\sigma$ is a supersolution to (2.4.1) through the following estimate:

$$F_j(\hat{\mathbf{u}}^i + C_1/\rho^\sigma) \geq F_j(\hat{\mathbf{u}}^i) + \gamma C_1/\rho^\sigma \geq F_j(u^\rho) \geq 0 \quad \forall j \in \mathcal{I},$$

which implies that $\hat{\mathbf{u}}^i + C_1/\rho^\sigma \geq \mathbf{u}$ for all $\rho > 0$ and $i \in \mathcal{I}$. □

Remark 2.4.1. Under Assumption 2, one can also show the rate of convergence for Problem 2.1.1 to Problem 2.4.1 is of first order in the switching cost. In fact, let u^c solve Problem 2.1.1 with a switching cost $c > 0$. Then we have $\|u^{c,i} - u^{c,j}\| \leq c$ for all $i, j \in \mathcal{I}$. Then following the proof of Theorem 2.4.2, we see it holds for some constant $K > 0$ that $0 \leq \mathbf{u} - u^c \leq Kc$.

A similar convergence result holds for switching system approximations to second-order HJB PDEs, with an order 1/3 as the switching cost tends to zero (see e.g. [11, Theorem 2.3]). The convergence order is reduced in the continuous setting since one needs to estimate the derivatives of non-smooth viscosity solutions via a regularization procedure. In the present finite-dimensional setting, we can work with the sup-norm of solutions and obtain the optimal first-order convergence.

Summarizing the above discussions, we can derive another upper bound of the penalization error for Problem 2.2.1 with the penalty function $\pi(y) = y^+$ and positive switching cost, which enables us to explain the asymptotic behaviours of the penalty errors observed in Section 2.5.

Theorem 2.4.3. *For any given switching cost $c > 0$ and penalty parameter $\rho > 0$, let u^c and $u^{c,\rho}$ be the solutions to Problems 2.1.1 and 2.2.1, respectively. Suppose Assumption 2 holds and the penalty function is given by $\pi(y) = y^+$. Then we have $0 \leq u^c - u^{c,\rho} \leq C_1(1/\rho + c\rho)$, for some constant $C_1 > 0$, independent of c and ρ.*

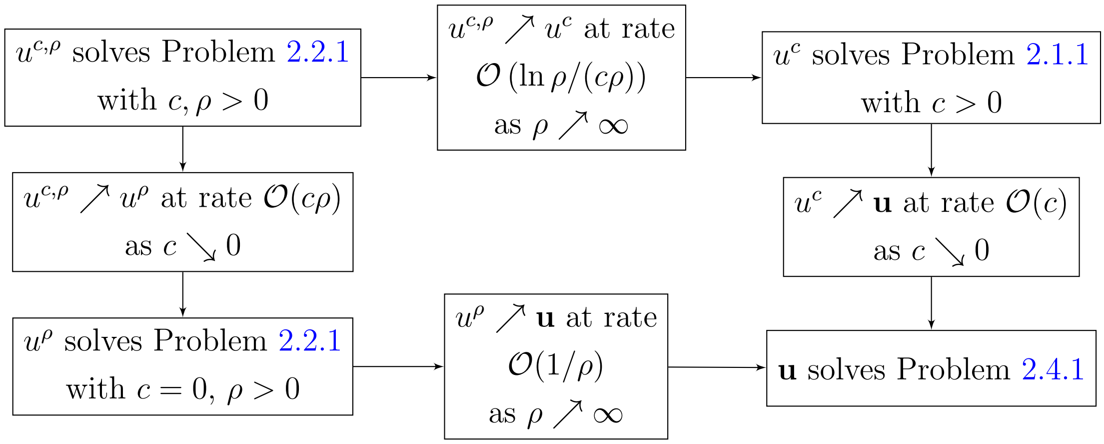
Proof. Note that for any $b \in \mathbb{R}$ and $c > 0$, we have

$$c + \pi(b - c) = c + \max(b - c, 0) = \max(b, c) \geq \max(b, 0) = \pi(b),$$

which, along with the inequality (2.3.6) (with $u = u^{c,\rho}$ and $L = (d-1)c\rho/\gamma$), implies that $u^{c,\rho} + (d-1)c\rho/\gamma$ is a supersolution to (2.2.1) with $c = 0$. Hence one can deduce from the comparison principle of the penalized equation (2.2.1) the estimate $0 \leq u^\rho - u^{c,\rho} \leq (d-1)c\rho/\gamma$. Then using Proposition 2.4.1 and Theorem 2.4.2, we conclude for some constant C_1, independent of ρ and c, that

$$u^c - u^{c,\rho} \leq u^c - \mathbf{u} + \mathbf{u} - u^\rho + u^\rho - u^{c,\rho} \leq C_1(1/\rho + c\rho). \qquad \square$$

We now summarize some of our main results (Theorems 2.3.9 and 2.4.3) in the following diagram. Suppose the function F in Problem 2.1.1 is concave (Assumption 1) and the penalty function in Problem 2.2.1 is given by $\pi(y) = y^+$. Then the following error estimates hold:



Hence for a fixed large enough penalty parameter, we have the following estimate of the penalization error as $c \to 0$:

$$0 \le u^c - u^{c,\rho} \le C_1 \min\left(\frac{\ln\rho}{c\rho}, \frac{1}{\rho} + c\rho\right), \tag{2.4.2}$$

for some constant C_1, independent of ρ and c. Thus for a sufficiently small switching cost, the penalty error is dominated by the term C_1/ρ, i.e., the penalization error with $c = 0$.

2.5 Numerical experiments

In this section, we illustrate the theoretical findings and demonstrate the effectiveness of the penalty schemes through numerical experiments. We present an infinite-horizon optimal switching problem and investigate the convergence of Problem 2.2.1 with respect to the penalty parameter. We shall also examine the dependence of the penalization errors on the switching cost.

Motivated by Remark 2.3.3, we shall focus on the penalty function with degree 1, i.e., $\pi(y) = y^+$. Due to the semismoothness of the chosen penalty function, one can easily construct convergent iterative methods for solving Problem 2.2.1 with a fixed penalty parameter (see e.g. [19, 101, 102, 87] for details). Roughly speaking, starting with an initial guess $u^{(0)}$ of the solution to Problem 2.2.1, for each $k \ge 0$, we compute the next iterate $u^{(k+1)}$ by solving

$$G^\rho[u^{(k)}] + \mathcal{L}^{(k+1)}[u^{(k)}](u^{(k+1)} - u^{(k)}) = 0,$$

where $\mathcal{L}^{(k+1)}[u^{(k)}]$ is a generalized derivative of G^ρ at the iterate $u^{(k)}$. In practice, such a generalized derivative can be computed by policy iteration if $F_i(u) =$

$\inf_{\alpha\in\mathcal{A}_i}\sup_{\beta\in\mathcal{B}_i} B(\alpha,\beta)u - b(\alpha,\beta)$ for some sets $\mathcal{A}_i, \mathcal{B}_i$ and some coefficients B and b, or more generally by a slanting function of $F = (F_i)_{i\in\mathcal{I}}$ if it exists. One can further show that the iterates $(u^k)_{k\geq 0}$ are locally superlinearly convergent to the solution of Problem 2.2.1 or even globally convergent if F is concave.

To motivate the discrete QVIs solved in our numerical experiments, we introduce the following infinite-horizon optimal switching problem (see e.g. [83]). Let $(\Omega, \mathcal{F}_t, \mathbb{P})$ be a filtered probability space and $\alpha = (\alpha_t)_{t\geq 0}$ be a control process such that $\alpha_t = \sum_{k\geq 0} i_k 1_{[\tau_k,\tau_{k+1})}(t)$, where $(\tau_k)_{k\geq 0}$ is a non-decreasing sequence of stopping times representing the decision on "when to switch", and for each $k \geq 0$, i_k is an $\mathcal{F}_{\tau_k}$-measurable random variable valued in the discrete space $\mathcal{I} = \{1, \dots, d\}$, $d \geq 2$, representing the decision on "where to switch". That is, the decision maker chooses regime i_k at the time τ_k for all $k \geq 0$.

For any given control strategy α, we consider the following controlled state equation:

$$dX_t^\alpha = (r + \nu(\alpha_t)(\mu - r))X_t^\alpha dt + \sigma\nu(\alpha_t)X_t^\alpha\, dB_t, \quad t > 0; \quad X_0^\alpha = x,$$

where $r, \mu, \sigma, x > 0$ are given constants, $(B_t)_{t>0}$ is a one-dimensional Brownian motion defined on $(\Omega, \mathcal{F}_t, \mathbb{P})$, and $\nu(i) = (i-1)/(d-1)$, $i \in \mathcal{I}$. Then the objective function associated with the control strategy α is given by:

$$J(x,\alpha) = \mathbb{E}\left[\int_0^\infty e^{-rt}\ell(X_t^\alpha)\, dt - \sum_{k\geq 0} e^{-r\tau_{k+1}} c^{i_k,i_{k+1}}\right],$$

where ℓ represents the running reward function and $c^{i,j}$ represents the switching cost from regime i to j, $\forall i,j \in \mathcal{I}$. For each $i \in \mathcal{I}$, let $\mathbf{A}^i$ be all control strategies starting with regime i, i.e., $i_0 = i$ and $\tau_0 = 0$. Then the decision maker has the following value functions:

$$V^i(x) = \sup_{\alpha\in\mathbf{A}^i} J(x,\alpha), \quad i \in \mathcal{I}.$$

Suppose the switching costs are positive, i.e., $c^{i,j} > 0$ for $i \neq j$, then one can show by using the dynamic programming principle (see [83]) that the value function $V = (V^i)_{i\in\mathcal{I}}$ satisfies the following system of quasi-variational inequalities: for each $i \in \mathcal{I}$,

$$\min\left[-\frac{1}{2}\sigma^2\nu(i)^2x^2V_{xx}^i - (r+\nu(i)(\mu-r))xV_x^i + rV^i - \ell(x), V^i - (\mathcal{M}_iV)\right] = 0, \quad x \in (0,\infty), \tag{2.5.1}$$

where $\mathcal{M}_iV = \max_{j\neq i}(V^i - c^{i,j})$. For our numerical tests, we assume $c^{i,j} \equiv c$ for $i \neq j$, and set other parameters as $\sigma = 0.2$, $\mu = 0.06$, $r = 0.02$.

Now we derive the finite-dimensional QVIs by discretizing (2.5.1). Note that in this work, we focus on examining the performance of penalty methods for solving discrete QVIs resulting from discretizing (2.5.1) with a fixed mesh size, instead of the convergence of the discretization to (2.5.1) as the mesh size tends to zero. Therefore, for simplicity, we shall localize (2.5.1) on the computational domain $(0, 2)$ with homogenous Dirichlet boundary condition $u = 0$ at $x = 2$, and solve the localized equation on a uniform grid $\{x_l\} = \{lh\}_{l=0}^{N-1}$ with $h = 2/N$. We further derive a monotone discretization of (2.5.1), which uses forward differences for the first derivates and central difference for all second derivatives. It is easy to verify that the resulting discrete system (2.1.1) satisfies the monotonicity condition (2.1.2) with $\gamma = 0.02$ and also Assumption 1.

We proceed to discuss implementation details for solving Problem 2.2.1 with semismooth Newton methods. The initial guess $u^{(0)}$ shall be taken as the solution to (2.2.1) with $\rho = 0$, i.e., $F_i(u) = 0$ for all $i \in \mathcal{I}$, and the iterations will be terminated once the desired tolerance is achieved:

$$\max_{l=0,\cdots,N-1} \left\{ \frac{|u_l^{(k)} - u_l^{(k-1)}|}{\max(|u_l^{(k)}|, \text{scale})} \right\} < \text{tol},$$

where the scale parameter is chosen to guarantee that no unrealistic level of accuracy will be imposed if the solution is close to zero. We take tol $= 10^{-9}$ and scale $= 1$ for all experiments. Computations are performed using Matlab R2016b on a laptop with 2.2 GHz Intel Core i7 and 16 GB memory.

We first study the performance of the penalty approximation for the discrete system corresponding to the two-regime case, i.e., $d = 2$, and the running reward $\ell(x) = 2(1 - x)1_{[0.75,1)}(x)$. Note that the discontinuity of ℓ at $x = 0.75$ should not affect our convergence analysis since we are solving a finite-dimensional nonlinear system resulting from a fixed discretization of (2.5.1).

Table 2.1 contains, for different switching costs and penalty parameters, the numerical solutions of Problem 2.2.1 with a fixed mesh size $h = 0.01$ (hence the total number of unknowns is $2N = 400$). Line (a) shows that for a fixed switching cost c, regardless of whether c is positive or not, the numerical solutions converge monotonically from below to the exact solution as the penalty parameter $\rho \to \infty$. The first-order convergence of the penalization error (in the sup-norm) with respect to the penalty parameter ρ can be deduced from line (b), which confirms the theoretical results (the log factor has not been observed, c.f. Theorems 2.3.9 and 2.4.2). Moreover, by fixing the penalty parameter ρ and comparing the increment $\|u^{c,\rho} - u^{c,\rho/2}\|$

ρ	10^3	2×10^3	4×10^3	8×10^3	16×10^3	32×10^3
$c = 1/2$						
(a)	3.38261	3.38633	3.38819	3.38913	3.38959	3.38983
(b)		0.00884	0.00444	0.00222	0.00111	0.00056
(c)	5	6	6	6	6	6
(d)	0.0021	0.0026	0.0027	0.0034	0.0031	0.0027
$c = 1/8$						
(a)	5.27999	5.28860	5.29292	5.29508	5.29617	5.29671
(b)		0.02039	0.01025	0.00514	0.00258	0.00129
(c)	6	5	5	5	5	5
(d)	0.0030	0.0026	0.0025	0.0025	0.0025	0.0024
$c = 1/32$						
(a)	6.01704	6.03478	6.04370	6.04817	6.05041	6.05153
(b)		0.04183	0.02114	0.01063	0.00533	0.00267
(c)	6	6	5	5	5	5
(d)	0.0038	0.0036	0.0032	0.0033	0.0029	0.0024
$c = 1/128$						
(a)	6.30708	6.34232	6.36011	6.36906	6.37354	6.37578
(b)		0.08234	0.04201	0.02122	0.01066	0.00534
(c)	5	5	4	4	4	4
(d)	0.0022	0.0026	0.0023	0.0019	0.0019	0.0018
$c = 1/512$						
(a)	6.42179	6.45776	6.47593	6.48506	6.48964	6.49193
(b)		0.08406	0.04288	0.02166	0.01089	0.00546
(c)	5	5	4	4	4	4
(d)	0.0021	0.0022	0.0018	0.0018	0.0019	0.0018
$c = 1/2048$						
(a)	6.45047	6.48662	6.50488	6.51406	6.51866	6.52097
(b)		0.08449	0.04310	0.02177	0.01094	0.00548
(c)	4	4	4	4	4	4
(d)	0.0018	0.0019	0.0019	0.0018	0.0019	0.0019
$c = 0$						
(a)	6.46003	6.49624	6.51454	6.52373	6.52834	6.53064
(b)		0.08464	0.04318	0.02181	0.01096	0.00549
(c)	4	4	3	3	3	3
(d)	0.0018	0.0019	0.0015	0.0015	0.0015	0.0015

Table 2.1: Numerical results for the two-regime optimal switching problem with different switching costs and penalty parameters. Shown are: (a) the numerical solutions $u^{c,\rho,1}$ at $x = 0.5$; (b) the increments $\|u^{c,\rho} - u^{c,\rho/2}\|$; (c) the number of iterations; (d) the overall runtime in seconds.

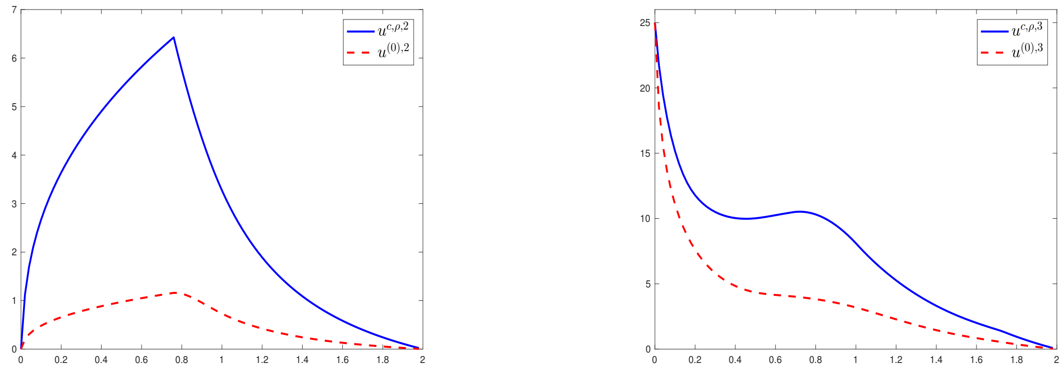


Figure 2.1: Differences between the last components of the initial guess $u^{(0)}$ and the solution $u^{c,\rho}$ for the penalized switching system. Shown are: two-regime problem with $\rho = 10^3$ and $c = 1/8$ (left), and three-regime problem with $\rho = 4\times10^3$ and $c = 1/1024$ (right).

columnwise, one can observe that the penalty errors first grow at a rate 1/2 when the switching cost decreases from 1/2 to 1/128, and then stablize to the penalty errors of the limiting case (with $c = 0$) when c tends to 0, as asserted by (2.4.2).

The lines (c) and (d) clearly indicate the efficiency of the iterative solver. We remark that compared with parabolic QVIs, elliptic QVIs are more challenging to solve due to the fact we cannot take the solution at the previous timestep as an accurate initial guess [41, 6]. In fact, Figure 2.1 (left) illustrates a large disagreement in the shape and magnitude between the initial guess $u^{(0)}$ and the final solution $u^{c,\rho}$ of Problem 2.2.1 with $\rho = 10^3$ and $c = 1/8$. However, we can see that the iterative method solves Problem 2.2.1 at the accuracy 10^{-9} using only a small number of iterations within several milliseconds, which seems to be independent of the size of the penalty parameter ρ.

We now turn to analyze the convergence of the penalty methods for the nonlinear system resulting from a three-regime problem, i.e., $d = 3$. The running reward function is chosen as

$$\ell(x) = \begin{cases} x - 0.5, & x \in (0.5, 1], \\ -(x - 1.5), & x \in (1, 1.5], \\ x - 1.5, & x \in (1.5, 1.75], \\ -5(x - 1.8), & x \in (1.75, 1.8], \\ 0, & \text{otherwise}, \end{cases}$$

which admits a mixed convexity. Table 2.2 presents the numerical solutions of the three-regime penalized equations with a fixed mesh $h = 0.01$ (hence the total number of unknowns is $3N = 600$), and different penalty parameter ρ and switching cost c. Lines (a) and (b) indicate that for a fixed switching cost, the numerical solutions converge monotonically with a first-order accuracy, as the penalty parameter tends to infinity. Moreover, similar to the two-regime problem, we can observe that as the switching cost tends to zero, the penalty errors corresponding to a fixed penalty parameter ρ first increase at a rate $\mathcal{O}(c^{-1/2})$ (for $c \in [1/4, 1/1024]$) and then approach to the penalization errors with $c = 0$. Lines (c) and (d) summarize the number of required iterations and the computational time, which illustrate the efficiency of the iterative solvers for the penalized problems. Despite the relatively poor initial guess (c.f. Figure 2.1 (right)), the desired accuracy 10^{-9} is in general obtained within 0.015 seconds using a reasonable amount of iterations, which does not depend on the magnitude of the penalty parameter.

ρ	4×10^3	8×10^3	16×10^3	32×10^3	64×10^3	128×10^3
$c = 1/4$						
(a)	6.849917	6.849942	6.849954	6.849960	6.849962	6.849964
(b)		0.000208	0.000104	0.000052	0.000026	0.000013
(c)	12	12	12	12	12	12
(d)	0.0112	0.0113	0.0111	0.0112	0.0113	0.0113
$c = 1/16$						
(a)	7.405239	7.405507	7.405641	7.405708	7.405742	7.405758
(b)		0.000451	0.000226	0.000113	0.000056	0.000028
(c)	12	12	12	12	12	12
(d)	0.0115	0.0115	0.0115	0.0111	0.0115	0.0117
$c = 1/64$						
(a)	7.791271	7.792091	7.792499	7.792703	7.792805	7.792856
(b)		0.001003	0.000501	0.000250	0.000125	0.000062
(c)	13	13	13	13	13	13
(d)	0.0121	0.0122	0.0151	0.0130	0.0127	0.0127
$c = 1/256$						
(a)	8.009477	8.011330	8.012258	8.012722	8.012955	8.013071
(b)		0.002016	0.001010	0.000505	0.000253	0.000126
(c)	14	14	14	14	14	14
(d)	0.0131	0.0130	0.0132	0.0129	0.0129	0.0130
$c = 1/1024$						
(a)	8.108554	8.112341	8.114262	8.115229	8.115715	8.115958
(b)		0.003980	0.002018	0.001017	0.000510	0.000256
(c)	15	15	14	15	15	15
(d)	0.0144	0.0145	0.0130	0.0145	0.0144	0.0143
$c = 1/4096$						
(a)	8.135298	8.138958	8.141012	8.142047	8.142567	8.142828
(b)		0.003854	0.002156	0.001087	0.000546	0.000273
(c)	14	14	14	14	14	14
(d)	0.0135	0.0130	0.0131	0.0132	0.0133	0.0132
$c = 1/16384$						
(a)	8.143553	8.146389	8.147826	8.148752	8.149280	8.149545
(b)		0.002975	0.001508	0.000974	0.000554	0.000278
(c)	12	12	14	14	14	14
(d)	0.0115	0.0115	0.0134	0.0134	0.0134	0.0133
$c = 0$						
(a)	8.146313	8.149164	8.150603	8.151326	8.151688	8.151869
(b)		0.002990	0.001509	0.000758	0.000380	0.000190
(c)	12	12	12	12	12	11
(d)	0.0112	0.0112	0.0111	0.0111	0.0111	0.0103

Table 2.2: Numerical results for the three-regime optimal switching problem with different switching costs and penalty parameters. Shown are: (a) the numerical solutions $u^{c,\rho,1}$ at $x = 1$; (b) the increments $\|u^{c,\rho} - u^{c,\rho/2}\|$; (c) the number of iterations; (d) the overall runtime in seconds.

Chapter 3

Penalty schemes for continuous quasi-variational inequalities

3.1 Introduction

In this chapter we study penalty schemes and their convergence for the following weakly coupled system of degenerate Hamilton-Jacobi-Bellman quasi-variational inequalities (HJBQVIs), which in particular can be applied to study optimal control of exchange rates introduced in Section 1.1: for all $i \in \mathcal{I} := \{1, \dots, M\}$,

$$\max\Big\{ \sup_{\alpha \in \mathcal{A}_i} \mathcal{L}_i^\alpha(x, u(x), Du_i(x), D^2u_i(x));\ (u_i - \mathcal{M}_i u)(x) \Big\} = 0, \quad x \in \mathbb{R}^d, \tag{3.1.1}$$

where $u = (u_i)_{i \in \mathcal{I}}$ with $u_i : \mathbb{R}^d \to \mathbb{R}$ denotes the unknown solution, $(\mathcal{L}_i^\alpha)_{i \in \mathcal{I}}$ is a family of second order differential operators, and $\mathcal{M}_i$ is an intervention operator of the following form:

$$(\mathcal{M}_i u)(x) = \min_{z \in Z_i(x)} \{ u_i(\Gamma_i(x, z)) + K_i(x, z) \}. \tag{3.1.2}$$

We shall specify the precise expression of the operators $(\mathcal{L}_i^\alpha)_{i \in \mathcal{I}}$ and the regularity of $\mathcal{A}_i, \mathcal{Z}_i, \Gamma_i, K_i$ in Section 3.2.

The above system extends the classical scalar HJBQVIs, and arises naturally from hybrid control problems of regime-switching models with both continuous and impulse controls (see [15, 100, 103, 97, 98]). For instance, let α be a càdlàg adapted stochastic control process, and let $\gamma = (\tau_1, \xi_1; \tau_2, \xi_2; \dots)$ be an impulse control strategy consisting of a sequence of impulse times $0 = \tau_0 \leq \tau_1 \leq \tau_2 \leq \dots$, and adapted impulse controls $(\xi_1, \xi_2, \dots)$. Between impulse times, we assume the state process X follows a controlled regime-switching process defined as follows: $X_0 = x \in \mathbb{R}^d$, $I_0 = i \in \mathcal{I}$, and for all $k \in \mathbb{N} \cup \{0\}$,

$$dX_t = b(\alpha_t, I_t, X_t)\, dt + \sigma(\alpha_t, I_t, X_t)\, dW_t, \quad \tau_k < t < \tau_{k+1},$$

where W is a standard Brownian motion, and I is a continuous-time Markov chain with values in the finite set $\mathcal{I}$, which represents the uncertainty in the environment and randomly switches among $M = |\mathcal{I}|$ states, governed by a controlled Markov transition matrix $(d_{ij}^{\alpha_t}(X_t))_{i,j\in\mathcal{I}}$. At an impulse time τ_k, the impulse control ξ_k is applied and instantaneously changes the state into $X_{\tau_k} = \Gamma(I_{\tau_k^-}, X_{\tau_k^-}, \xi_k)$. The aim is to minimize the expected cost over all admissible strategies (α, γ) by considering the following value function:

$$u_i(x) \coloneqq \inf_{\alpha,\gamma} \mathbb{E}\left[\int_0^\infty \ell(\alpha_t, I_t, X_t)e^{-c(I_t,X_t)t}\,dt + \sum_{k=1}^\infty K(I_{\tau_k^-}, X_{\tau_k^-}, \xi_k)e^{-c(I_{\tau_k^-},X_{\tau_k^-})\tau_k}\right] \tag{3.1.3}$$

for each $x \in \mathbb{R}^d$ and $i \in \mathcal{I}$, where ℓ and K are the running cost and the impulse cost, respectively.

Such hybrid control problems appear in mathematical finance, such as in the following applications: portfolio optimization with transaction costs [70, 81, 6], control of exchange rates [70, 81, 31, 6], credit securitization [92], inventory control and dividend control [70, 8]. It is well-known that under suitable assumptions, the value functions $(u_i)_{i\in\mathcal{I}}$ in (3.1.3) can be characterized by the viscosity solution to (3.1.1) (see e.g. [100, 98]). Note that due to the random switching process I, each operator $\mathcal{L}_i^\alpha$ involves all components of the solution u, which leads us to a weakly coupled system of HJBQVIs (see e.g. [103, 23]).

As the solution to (3.1.1) is in general not known analytically, several classes of numerical schemes have been proposed to solve such nonlinear equations. By writing the obstacle term $u_i - \mathcal{M}_i u$ as $\max_{z\in Z_i(x)}[u_i - u_i(\Gamma_i(x,z)) - K_i(x,z)]$, one can extend the "direct control" scheme of HJB equations to solve (3.1.1), which discretizes the operators in (3.1.1) and attempts to solve the resulting nonlinear discrete equations using policy iteration [27, 6]. However, due to the non-strict monotonicity of the term $u_i - \mathcal{M}_i u$, such a scheme in general requires a very accurate initial guess for the policy iteration to converge. In fact, as we shall show in Remark 3.6.1, even for some simple intervention operators, policy iteration in the direct control scheme may not be well-defined for an arbitrary initial guess due to the possible singularity of the matrix iterates.

An alternative approach to solving (3.1.1), referred to as iterated optimal stopping, approximates the QVI by a sequence of HJB variational-inequalities (see (3.4.1)–(3.4.2)), which can subsequently be solved by the direct control scheme [81, 92]. However, since this approach can be equivalently formulated as a fixed point algorithm for the QVI, one can show that this approach in general suffers from slow convergence (i.e. rate close to 1), especially for small impulse costs [88].

In this work, we shall extend the penalty schemes in [68, 5] for scalar equations (i.e. $M = 1$) to systems of HJBQVIs, and construct the solution of (3.1.1) from a sequence of penalized equations. The major advantage of the penalty approximation is that one can easily construct convergent monotone discretizations of the penalized equation with a fixed penalty parameter, and policy iteration applied to the discrete equation is monotonically convergent with any initial guess (see Section 3.6). Moreover, the Lagrange multipliers of the penalized equations enjoy better regularity than those of the unpenalized QVI (3.1.1). It is observed empirically that this improved regularity leads to mesh-independent behaviour of policy iteration for solving the penalized equations, i.e., the number of iterations for solving the discrete problem remains bounded as the mesh size tends to zero, in contrast to the direct scheme (see Figure 3.2 in Section 3.7, see also [86, 48]).

The main contributions of this chapter are:

- We design efficient penalty schemes for solving systems of HJBQVIs with general intervention operators. We further establish that as the penalty parameter ρ tends to infinity, the solution of the penalized equation converges monotonically from above to the solution of the HJBQVI. We shall also construct novel convergent approximations of the action regions and optimal impulse control strategies based on the penalized solutions.

- We quantify the convergence rate of penalty approximations for degenerate HJBQVIs, which is novel even in the scalar case (i.e. $M = 1$). Although the convergence of penalty schemes for QVIs has been proved in various works (e.g. [77, 68, 5]), to the best of our knowledge, there is no published work on the accuracy of the penalty approximation with a given penalty parameter (except for Chapter 2 where the penalty error for discrete QVIs has been analyzed). This is not only important for the choice of penalty parameters and the practical implementation of penalty schemes, but is also crucial for the construction of action regions (see Remark 3.4.3) and optimal impulse control strategies. In this work, we shall close the gap by giving a rigorous analysis of penalty errors.

- We further extend the penalty scheme and its error estimate to a class of HJBQVIs with possibly *negative impulse costs*. Note that signed costs are not only of mathematical interest, but are also important to model the situation where the controller can obtain a positive impulse benefit, for example, receive financial support for investing in renewable energy production (see [83, 79]). In this setting, we deduce error estimates for a different type of penalty schemes, which apply the penalty to each impulse control strategy, instead of the pointwise maximum over all impulse control strategies (Remark 3.5.1). These convergence results rely on a novel construction of a strict subsolution to HJBQVIs with general switching costs, for which we impose less restrictive conditions on the switching costs than those given in the literature (see the discussion after Assumption 9 for details).

Let us briefly comment on the two main difficulties encountered in deriving the error estimates of penalty schemes for HJBQVIs. In contrast to the results for finite-dimensional (discretized) QVIs in Chapter 2, the convergence rate of penalty approximations for HJBQVIs depends on the regularity of the solution. Since in this work we focus on degenerate HJBQVIs, including the fully degenerate case where $\mathcal{L}_i^\alpha$ reduces to a first-order differential operator, the solution of (3.1.1) is typically not differentiable due to the lack of regularization from the Laplacian operator. Therefore, we need to obtain suitable regularity of the solution to weakly coupled systems based on viscosity solution theory [30].

Moreover, the non-diagonal dominance of the obstacle term $u_i - \mathcal{M}_i u$ poses a significant challenge for estimating the penalization errors. In fact, a crucial step in estimating the penalty error for HJB variational-inequalities is to show that there exists a constant C, depending on the regularity of the obstacle, such that for any $\rho > 0$, if u^ρ solves the penalized equation with the parameter ρ, then $u^\rho - C/\rho$ satisfies the constraint of the variational inequality (see e.g. [64, 102]). However, this is in general false for the QVIs since the term $u_i - \mathcal{M}_i u$ remains invariant under any vertical shift of the solutions.

We shall overcome the above difficulty by combining the ideas of [20, 64] and precise regularity estimates (i.e., Lipschitz continuity and semiconcavity) of solutions to HJB variational inequalities. In particular, we shall construct a family of auxiliary approximations for our error analysis via iterated optimal stopping. This reduces the problem to estimating the solution regularity and penalty errors for a sequence of obstacle problems. We shall derive a more precise estimate for the semiconcavity constant of the solution to HJB variational inequalities with respect to the obstacle term than those in prior works (see the discussion above Proposition 3.4.3). This is crucial for us to be able to conclude that the penalty approximation is half-order accurate for HJBQVIs with Lipschitz coefficients, and first-order accurate for equations with more regular coefficients (see Theorems 3.4.10 and 3.5.2). These convergence rates of penalty schemes for HJBQVIs are optimal in the sense that they are of the same order (up to logarithmic terms) as those for conventional HJB variational inequalities.

Finally, we would like to point out a control-theoretic interpretation of our penalty schemes. As observed in [75, 76], the viscosity solution of the penalized equation with parameter ρ can be identified as the value function of a hybrid control problem where the controller is only allowed to perform impulse controls at a sequence of Poisson arrival times with intensity ρ, instead of any stopping times. Our error estimates give a convergence rate of these hybrid control problems with random intervention times in terms of the intensity ρ, which is of independent interest.

We organize this chapter as follows. Section 3.2 states the main assumptions and recalls basic results for the system of HJBQVIs with positive impulse costs. In Section 3.3 we shall propose a penalty approximation to the HJBQVIs and establish its monotone convergence. Then by exploiting the regularization introduced in Section 3.4.1, we estimate the convergence rates of the penalty schemes in Section 3.4.3, and construct convergent approximations to action regions and optimal impulse controls in Section 3.4.4. We extend the convergence results to HJBQVIs with signed costs in Section 3.5, and discuss the monotone convergence of policy iteration in Section 3.6. Numerical examples for infinite-horizon optimal switching problems are presented in Section 3.7 to illustrate the effectiveness of the penalty schemes. Appendix A.1 is devoted to the proofs of some technical results.

3.2 HJBQVIs with positive costs

In this section, we introduce the system of HJBQVIs of our interest, state the main assumptions on its coefficients, and recall the appropriate notion of solutions. We start with some useful notation which is needed frequently throughout this work.

For a function $\phi : \mathbb{R}^d \to \mathbb{R}$, we define the following (semi-)norms:

$$|\phi|_0 = \sup_{x\in\mathbb{R}^d} |\phi(x)|, \quad [\phi]_1 = \sup_{x,y\in\mathbb{R}^d} \frac{|\phi(x) - \phi(y)|}{|x - y|}, \quad |\phi|_1 = |\phi|_0 + [\phi]_1.$$

As usual, we denote by $C^0(\mathbb{R}^d)$ (resp. $C^n(\mathbb{R}^d)$) the space of bounded continuous functions (resp. n-times differentiable functions) in $\mathbb{R}^d$, and by $C^{0,1}(\mathbb{R}^d)$ the subset of functions in $C^0(\mathbb{R}^d)$ with finite $|\cdot|_1$ norm. Finally, we shall denote by $\mathbb{S}^d$ the set of $d \times d$ symmetric matrices, and by $X \geq Y$ in $\mathbb{S}^d$ the fact that $X - Y$ is positive semi-definite.

We shall consider the following weakly coupled system: for each $i \in \mathcal{I} := \{1, \dots, M\}$, and $x \in \mathbb{R}^d$,

$$F_i(x, u, Du_i, D^2 u_i) := \max\Big\{ \sup_{\alpha \in \mathcal{A}_i} \mathcal{L}_i^\alpha(x, u(x), Du_i(x), D^2 u_i(x));\ (u_i - \mathcal{M}_i u)(x) \Big\} = 0, \tag{3.2.1}$$

where $u = (u_i)_{i \in \mathcal{I}}$, $\mathcal{L}_i^\alpha : \mathbb{R}^d \times \mathbb{R}^M \times \mathbb{R}^d \times \mathbb{S}^d \to \mathbb{R}$ is the following linear operator:

$$\mathcal{L}_i^\alpha(x, s, p, X) = -\mathrm{tr}[a_i^\alpha(x) X] - b_i^\alpha(x) p + c_i^\alpha(x) s_i - \ell_i^\alpha(x) - \sum_{j \in \mathcal{I}^{-i}} d_{ij}^\alpha(x) s_j, \tag{3.2.2}$$

with $\mathcal{I}^{-i} := \{j \in \mathcal{I} \mid j \neq i\}$, $\mathcal{M}_i$ is the intervention operator (3.1.2), i.e.,

$$(\mathcal{M}_i u)(x) = \min_{z \in Z_i(x)} \{u_i(\Gamma_i(x, z)) + K_i(x, z)\}. \tag{3.2.3}$$

Before introducing the assumptions on the coefficients, let us recall the concept of semiconcavity of a continuous function [30, 9], which is crucial for the subsequent convergence analysis.

Definition 3.2.1 (Semiconcavity)**.** A continuous function ϕ is semiconcave around $x \in \mathbb{R}^d$ with constant $C \geq 0$, if it holds that

$$\phi(x+h) - 2\phi(x) + \phi(x-h) \leq C|h|^2, \quad \text{for all sufficiently small } h \in \mathbb{R}^d. \tag{3.2.4}$$

We say a continuous function ϕ is semiconcave with constant $C \geq 0$ if (3.2.4) holds for all $x \in \mathbb{R}^d$. For any given semiconcave function ϕ, we shall denote by $[\phi]_{2,+}$ its semiconcavity constant, i.e.,

$$[\phi]_{2,+} := \inf\{C \geq 0 \mid u(x+h) - 2u(x) + u(x-h) \leq C|h|^2,\ x, h \in \mathbb{R}^d\}.$$

A concave function is clearly semiconcave. Moreover, a C^1 function with locally Lipschitz gradient is semiconcave [9].

We now list the main assumptions on the coefficients.

Assumption 3. *For any $i, j \in \mathcal{I}$, $\mathcal{A}_i$ is a nonempty compact set, $a_i^\alpha = \frac{1}{2}\sigma_i^\alpha \sigma_i^{\alpha T}$ for some $\sigma_i^\alpha \in \mathbb{R}^{d \times d'}$, and σ_i^α, b_i^α, ℓ_i^α, c_i^α, d_{ij}^α are continuous functions. Moreover, there exist constants C and λ_0 such that it holds for any $j \neq i$, $\alpha \in \mathcal{A}_i$ that*

$$|\sigma_i^\alpha|_1 + |b_i^\alpha|_1 + |\ell_i^\alpha|_1 + |c_i^\alpha|_1 + |d_{ij}^\alpha|_1 \leq C, \tag{3.2.5}$$

$$d_{ij}^\alpha \geq 0, \quad c_i^\alpha - \sum_{j \in \mathcal{I}^{-i}} d_{ij}^\alpha \geq \lambda_0 > 0. \tag{3.2.6}$$

Assumption 4. *For any $i \in \mathcal{I}$ and $x \in \mathbb{R}^d$, $Z_i(x)$ is a nonempty compact set in a metric space $(\boldsymbol{Z}, d_{\boldsymbol{Z}})$, Γ_i and K_i are continuous functions, and the mapping $x \to Z_i(x)$ is continuous in the Hausdorff metric. Moreover, there exists a constant κ_0 such that for all $i \in \mathcal{I}$, $x \in \mathbb{R}^d$ and $z \in Z_i(x)$, we have $K_i(x,z) \geq \kappa_0 > 0$.*

The condition (3.2.5) in Assumption 3 is the standard regularity assumption for the coefficients in viscosity solution theory, while (3.2.6) in Assumption 3 implies the monotonicity of the HJB equations. Assumption 4 on the intervention operator is the same as that in [92, 5], which ensures the well-posedness of (3.2.1) in the class of bounded continuous functions. As we shall show in Section 3.3, they are sufficient for the monotone convergence of the penalty approximation, even for non-convex/non-concave systems involving Isaacs' equations.

The following additional assumptions are necessary to derive the regularity of the value functions and quantify the error estimates of the penalty schemes.

Assumption 5. *The constant λ_0 in Assumption 3 satisfies $\lambda_0 > \sup_{\alpha,i}([\sigma_i^\alpha]_1^2 + [b^\alpha]_1)$.*

Assumption 6. *There exists a constant $C > 0$ such that for any $i, j \in \mathcal{I}$, $\alpha \in \mathcal{A}_i$, we have $\sigma_i^\alpha, b_i^\alpha, c_i^\alpha, d_{ij}^\alpha \in C^1(\mathbb{R}^d)$ satisfying the estimate*

$$|D\sigma_i^\alpha|_1 + |Db_i^\alpha|_1 + |Dc_i^\alpha|_1 + |Dd_{ij}^\alpha|_1 \leq C,$$

and ℓ_i^α is semiconcave with constant C in $\mathbb{R}^d$.

Assumption 7. *For any $i \in \mathcal{I}$, the operator $\mathcal{M}_i$ preserves Lipschitz functions, i.e., there exists a constant $C > 0$ such that for any $u \in C^{0,1}(\mathbb{R}^d)$, $\mathcal{M}_i u$ is Lipschitz continuous with a constant satisfying $[\mathcal{M}_i u]_1 \leq [u]_1 + C$.*

Assumption 8. *For any $i \in \mathcal{I}$, the operator $\mathcal{M}_i$ preserves semiconcave functions, i.e., there exists a constant $C > 0$ such that for any given bounded semiconcave function u, $\mathcal{M}_i u$ is semiconcave with a constant satisfying $[\mathcal{M}_i u]_{2,+} \leq [u]_{2,+} + C$.*

Let us briefly discuss the importance of the above assumptions. Assumption 5 is the standard assumption for the Lipschitz continuity of the value functions (see [77]), while Assumption 6 will be used to establish the semiconcavity of the solutions in Section 3.4.1, which is the maximal regularity that one can expect for the solutions of degenerate HJB equations (see e.g. [9]).

Assumptions 7 and 8 are certain structural assumptions for the intervention operator $\mathcal{M}_i$, which play an essential role in our error estimates. In general these conditions need to be verified in a problem dependent way, as demonstrated in the following special cases.

Example 3.2.1. For the commonly studied intervention operator (see e.g. [54, 20, 31, 97, 6, 4]):

$$\mathcal{M}_i u(x) = \inf_{z\in\mathbb{R}^p}[u_i(x+\gamma(z)) + K(z)], \quad x \in \mathbb{R}^d, \tag{3.2.7}$$

with $|K(z)| \to \infty$ as $|z| \to \infty$, it is straightforward to show that Assumptions 7 and 8 hold with $C = 0$ (note the growth of K ensures the optimal impulse strategy is attained in a compact set). See Section 3.5 for examples with state-dependent impulse costs.

Example 3.2.2. For the intervention operator $\mathcal{M}u(x) = \inf_{z\in Z(x)}[u(x-z)+K(z)]$, with $x \in \mathbb{R}^d_+ := (0,\infty)^d$ and $Z(x) = \{z \in \mathbb{R}^d \mid 0 \le z_i \le x_i,\ i = 1,\dots,d\}$, which is a concave analogue of the maximum utility operator for multi-dimensional optimal dividend/inventory problems (e.g. [8]), one can show that Assumption 7 (resp. Assumption 8) holds if K is Lipschitz continuous (resp. semiconcave).

We shall only discuss Assumption 8, since Assumption 7 can be shown by a similar approach. Let $x \in \mathbb{R}^d_+$ and $\hat{z} \in Z(x)$ such that $\mathcal{M}u(x) = u(x-\hat{z}) + K(\hat{z})$. Define the set $\mathbb{I}_x = \{1 \le i \le d \mid \hat{z}_i = x_i\}$ and the constant $h_0 = \min(\min_{i\notin\mathbb{I}_x}(x_i - \hat{z}_i), \min_{i=1,\dots,d} x_i) > 0$. Then for any given $h \in \mathbb{R}^d$ such that $|h|_0 < h_0$, we can consider the vector $h^{\mathbb{I}_x} = (h_i^{\mathbb{I}_x})_{i=1}^d$ defined by $h_i^{\mathbb{I}_x} = h_i$ if $i \in \mathbb{I}_x$ and 0 otherwise, which clearly satisfies the following properties:

$$0 \le \hat{z}_i + h_i^{\mathbb{I}_x} \le x_i + h_i, \quad 0 \le \hat{z}_i - h_i^{\mathbb{I}_x} \le x_i - h_i, \quad \forall i \in \mathcal{I}.$$

In other words, we have $z^+ := \hat{z} + h^{\mathbb{I}_x} \in Z(x+h)$ and $z^- := \hat{z} - h^{\mathbb{I}_x} \in Z(x-h)$. Thus one can deduce from the semiconcavity of u and K that $\mathcal{M}u$ is semiconcave around x:

$$\begin{aligned}\mathcal{M}u(x+h)-&2\mathcal{M}u(x) + \mathcal{M}u(x-h)\\ &\le u(x+h-z^+) + K(z^+) - 2(u(x-\hat{z}) + K(\hat{z})) + u(x-h-z^-) + K(z^-)\\ &\le [u]_{2,+}|h-h^{\mathbb{I}_x}|^2 + [K]_{2,+}|h^{\mathbb{I}_x}|^2 \le ([u]_{2,+} + [K]_{2,+})|h|^2,\end{aligned}$$

which subsequently leads to the desired estimate $[\mathcal{M}u]_{2,+} \le [u]_{2,+} + [K]_{2,+}$.

Example 3.2.3. A general intervention operator (3.2.3) satisfies Assumption 7 under the following Lipschitz conditions on the data: there exist constants $C_1, C_2, C_3, C_4 \ge 0$ such that $C_2 + C_1C_3 \le 1$ and

$$\begin{aligned}&Z_i(y) \subseteq Z_i(x) + \bar{B}_{C_1|x-y|}, \quad |\Gamma_i(x,z) - \Gamma_i(y,z')| \le C_2|x-y| + C_3 d_{\boldsymbol{Z}}(z,z'),\\ &|K_i(x,z) - K_i(y,z')| \le C_4(|x-y| + d_{\boldsymbol{Z}}(z,z')), \qquad \forall i \in \mathcal{I}, x, y \in \mathbb{R}^d, z, z' \in \boldsymbol{Z},\end{aligned}$$

where for each $r \geq 0$, $\bar{B}_r$ denotes a closed ball of center 0 and radius r in the metric space $\boldsymbol{Z}$. In fact, let $i \in \mathcal{I}$, $u \in C_1^0(\mathbb{R}^d)$, $x, y \in \mathbb{R}^d$, $\hat{z}(y) \in Z(y)$ such that $(\mathcal{M}_i u)(y) = u(\Gamma_i(y, \hat{z}(y))) + K_i(y, \hat{z}(y))$. Then we can find $z(x) \in Z_i(x)$ such that $d_{\boldsymbol{Z}}(z(x), \hat{z}(y)) \leq C_1|x - y|$, which leads to the following estimate that

$$
\begin{aligned}
&(\mathcal{M}_i u)(x) - (\mathcal{M}_i u)(y) \\
&\leq \big[u(\Gamma_i(x, z(x))) + K_i(x, z(x))\big] - \big[u(\Gamma_i(y, \hat{z}(y))) + K_i(y, \hat{z}(y))\big] \\
&\leq [u]_1|\Gamma_i(x, z(x)) - \Gamma_i(y, \hat{z}(y))| + |K_i(x, z(x)) - K_i(y, \hat{z}(y))| \\
&\leq [u]_1\big(C_2|x - y| + C_3 d_{\boldsymbol{Z}}(z(x), \hat{z}(y))\big) + C_4(|x - y| + d_{\boldsymbol{Z}}(z(x), \hat{z}(y))) \\
&\leq \big([u]_1\big(C_2 + C_1C_3\big) + C_4(1 + C_1)\big)|x - y|.
\end{aligned}
$$

Hence we can conclude from the assumption $C_2 + C_1C_3 \leq 1$ that $\mathcal{M}_i$ satisfies Assumption 7 with $C = C_4(1 + C_1)$. A sufficient condition of Assumption 8 for the intervention operator (3.2.3) in general involves technical second-order conditions on the set-valued mapping $x \to Z(x)$, which will not be derived here for the sake of simplicity.

Note that we do not require any non-degeneracy condition on the diffusion coefficients, i.e., the coefficient a_i^α may vanish at certain points, hence our results apply to the fully degenerate case with $a^\alpha = 0$, where (3.2.1) reduces to QVIs of first order.

We now discuss the well-posedness of HJBVI (3.2.1). Due to the lack of regularization from a Laplacian operator, the solution of (3.2.1) is typically nonsmooth and we shall understand all equations in this work in the following viscosity sense.

Definition 3.2.2 (Viscosity solution)**.** A bounded upper-semicontinuous (resp. lower-semicontinuous) function $u = (u_i)_{i\in\mathcal{I}}$ is a viscosity subsolution (resp. supersolution) to (3.2.1), if for each $i \in \mathcal{I}$ and function $\phi \in C^2(\mathbb{R}^d)$, at each local maximum (resp. minimum) point x of $u_i - \phi$ we have $F_i(x, u(x), D\phi(x), D^2\phi(x)) \leq 0$ (resp. ≥ 0). A continuous function is a viscosity solution of (3.2.1) if it is both a subsolution and a supersolution.

Remark 3.2.1. Definition 3.2.2 formulates the notation of viscosity solution with suitable test functions. It is well-known that one can equivalently define the viscosity solution to (3.2.1) in terms of the superjet and subjet of u_i at $x \in \mathbb{R}^d$, denoted by $J^{2,+}u_i(x)$ and $J^{2,-}u_i(x)$ respectively, or their closures $\bar{J}^{2,+}u_i(x)$ and $\bar{J}^{2,-}u_i(x)$ (see e.g. [57, Proposition 2.3]).

More precisely, for a function $v : \mathbb{R}^d \to \mathbb{R}$ and $x \in \mathbb{R}^d$, let the semijects of v at x be defined by:

$$\begin{aligned} J^{2,+}v(x) = \{(p, X) \in \mathbb{R}^d \times \mathbb{S}^d \mid v(x+h) \le v(x) + \langle p, h\rangle \\ + \tfrac{1}{2}\langle Xh, h\rangle + o(|h|^2), \text{ as } h \to 0\}, \\ J^{2,-}v(x) = \{(p, X) \in \mathbb{R}^d \times \mathbb{S}^d \mid v(x+h) \ge v(x) + \langle p, h\rangle \\ + \tfrac{1}{2}\langle Xh, h\rangle + o(|h|^2), \text{ as } h \to 0\}, \end{aligned}$$

and their closures be defined by

$$\begin{aligned} \bar{J}^{2,+}v(x) = \{(p, X) \in \mathbb{R}^d \times \mathbb{S}^d \mid & \text{ there exists a sequence } (x_k, p_k, X_k) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{S}^d \\ & \text{such that } (p_k, X_k) \in J^{2,+}v(x_k) \text{ for all } k \text{ and} \\ & \lim_{k\to\infty} (x_k, v(x_k), p_k, X_k) = (x, v(x), p, X)\}, \\ \bar{J}^{2,-}v(x) = \{(p, X) \in \mathbb{R}^d \times \mathbb{S}^d \mid & \text{ there exists a sequence } (x_k, p_k, X_k) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{S}^d \\ & \text{such that } (p_k, X_k) \in J^{2,-}v(x_k) \text{ for all } k \text{ and} \\ & \lim_{k\to\infty} (x_k, v(x_k), p_k, X_k) = (x, v(x), p, X)\}. \end{aligned}$$

Then a bounded upper-semicontinuous (resp. lower-semicontinuous) function $u = (u_i)_{i\in\mathcal{I}}$ is a viscosity subsolution (resp. supersolution) to (3.2.1) if and only if

$$\begin{aligned} F_i(x, u(x), p, X) \le 0 \quad \forall i \in \mathcal{I},\ x \in \mathbb{R}^d,\ (p, X) \in J^{2,+}u_i(x), \\ \big(\text{resp. } F_i(x, u(x), p, X) \ge 0 \quad \forall i \in \mathcal{I},\ x \in \mathbb{R}^d,\ (p, X) \in J^{2,-}u_i(x).\big) \end{aligned}$$

Another equivalent definition of viscosity subsolution (resp. supersolution) for (3.2.1) is to replace $J^{2,+}u_i(x)$ (resp. $J^{2,-}u_i(x)$) by $\bar{J}^{2,+}u_i(x)$ (resp. $\bar{J}^{2,-}u_i(x)$) in the above formulation.

The fact that the impulse cost is strictly positive (see Assumption 4) implies $-C$ is a strict subsolution to (3.2.1) for a large enough constant $C > 0$. Therefore, one can establish a comparison principle of (3.2.1) by using similar arguments as in [54] (cf. the proof of Proposition 3.3.1; see also [92, Theorem 2.5.11] for a related result for solutions of polynomial growth). The comparison principle directly leads to the uniqueness of bounded viscosity solutions to (3.2.1), which can be explicitly constructed through penalty approximations (Theorem 3.3.2).

Proposition 3.2.1. *Suppose Assumptions 3 and 4 hold. If u (resp. v) is a bounded subsolution (resp. supersolution) of (3.2.1), then $u \le v$ in $\mathbb{R}^d$.*

We end this section by collecting several important properties of the intervention operator $\mathcal{M}_i$.

Lemma 3.2.2. *For any $i \in \mathcal{I}$, we have:*

(1) $\mathcal{M}_i$ is concave, i.e., it holds for any locally bounded functions $u, v : \mathbb{R}^d \to \mathbb{R}$ and constant $\lambda \in [0,1]$ that $\mathcal{M}_i[(1-\lambda)u + \lambda v] \geq (1-\lambda)\mathcal{M}_i u + \lambda \mathcal{M}_i v$.

(2) $\mathcal{M}_i$ is monotone, i.e., if $u \geq v$, then $\mathcal{M}_i u \geq \mathcal{M}_i v$.

(3) Suppose Assumption 4 holds, and let $(u^\rho)_{\rho\in\mathbb{N}}$ be a family of uniformly bounded functions on $\mathbb{R}^d$ with the following half-relaxed limits u^* and u_*:

$$u^*(x) \coloneqq \limsup_{\rho\to\infty, y\to x} u^\rho(y), \quad u_*(x) \coloneqq \liminf_{\rho\to\infty, y\to x} u^\rho(y), \quad x \in \mathbb{R}^d. \tag{3.2.8}$$

Then it holds for any given $x \in \mathbb{R}^d$ and sequence $(x^\rho)_{\rho\in\mathbb{N}}$ with $\lim_{\rho\to\infty} x^\rho = x$ that

$$(\mathcal{M}_i u_*)(x) \leq \liminf_{\rho\to\infty} (\mathcal{M}_i u^\rho)(x^\rho) \leq \limsup_{\rho\to\infty} (\mathcal{M}_i u^\rho)(x^\rho) \leq (\mathcal{M}_i u^*)(x). \tag{3.2.9}$$

Proof. Properties (1) and (2) follow directly from the structure of $\mathcal{M}_i$. Property (3) is an analogue of [5, Lemma 12] to the present concave intervention operator $\mathcal{M}_i$ and compact set Z_i, whose proof will be given in Appendix A.1 for completeness. □

3.3 Penalty approximations for HJBQVIs

In this section, we propose a penalty approximation for the system of HJBQVIs (3.2.1), which is an extension of the ideas used for scalar HJBQVIs in [14, 5]. We shall also establish the monotone convergence of the penalized solutions in terms of the penalty parameter.

For any given penalty parameter $\rho \geq 0$, we consider the following system of HJB equations: for all $i \in \mathcal{I}$ and $x \in \mathbb{R}^d$,

$$F_i^\rho(x, u^\rho, Du_i^\rho, D^2 u_i^\rho) \coloneqq \sup_{\alpha\in\mathcal{A}_i} \mathcal{L}_i^\alpha(x, u^\rho(x), Du_i^\rho(x), D^2 u_i^\rho(x)) + \rho(u_i^\rho - \mathcal{M}_i u^\rho)^+(x) = 0, \tag{3.3.1}$$

where the operators $\mathcal{L}_i^\alpha$ and $\mathcal{M}_i$ are defined as in (3.2.2) and (3.2.3), respectively.

The definitions of viscosity solution, sub- and supersolution for (3.3.1) extend naturally from Definition 3.2.2. The following result asserts the comparison principle and the well-posedness of (3.3.1) for any given penalty parameter.

Proposition 3.3.1. *Suppose Assumptions 3 and 4 hold, and let $\rho \geq 0$ be a given penalty parameter. If u^ρ (resp. v^ρ) is a bounded subsolution (resp. supersolution) of (3.3.1), then $u^\rho \leq v^\rho$ in $\mathbb{R}^d$. Consequently, (3.3.1) admits a unique viscosity solution, which is uniformly bounded in ρ.*

Proof. We postpone the proof of the comparison principle to Appendix A.1, which adapts the strict subsolution technique in [54] to the penalized equation, and reduces the problem to a HJB equation without the penalty part.

Since $K(x,z) \geq \kappa_0 > 0$, there exits a large enough constant C, independent of ρ, such that $-C$ and C are the viscosity sub- and supersolution of (3.3.1) with any parameter ρ, respectively. Thus by using the comparison principle and Perron's method (see [57, Theorem 3.3]), one can deduce the well-posedness of (3.3.1) in the viscosity sense. □

The next result demonstrates the monotone and locally uniform convergence of the solution $(u^\rho)_{\rho\geq 0}$ of (3.3.1) in terms of the penalty parameter ρ.

Theorem 3.3.2. *Suppose Assumptions 3 and 4 hold. Then as $\rho \to \infty$, the solution of (3.3.1) converges monotonically from above to the bounded viscosity solution of (3.2.1), uniformly on compact sets.*

Proof. It is clear that if $\rho_1 \leq \rho_2$ and u^{ρ_2} is a subsolution to (3.3.1) with the parameter ρ_2, then u^{ρ_2} is a subsolution to (3.3.1) with the parameter ρ_1. Hence the comparison principle leads to the fact that $u^0 \geq u^{\rho_1} \geq u^{\rho_2}$. Now we shall adopt the equivalent definition of viscosity solution in terms of semi-jets and prove that the component-wise half-relaxed limit u^* (resp. u_*) is a subsolution (resp. supersolution) to (3.2.1).

We start by showing u^* is a subsolution. Let $x \in \mathbb{R}^d$, $i \in \mathcal{I}$ and $(p,X) \in J^{2,+}u_i^*(x)$, then by applying [30, Lemma 6.1], there exist sequences $(x^\rho, p^\rho, X^\rho)_{\rho\in\mathbb{N}}$ such that $(p^\rho, X^\rho) \in J^{2,+}u_i^\rho(x^\rho)$ for each ρ and $(x^\rho, u_i^\rho(x^\rho), p^\rho, X^\rho) \to (x, u_i^*(x), p, X)$ as $\rho \to \infty$. Since u^ρ is a subsolution to (3.3.1), we have

$$\sup_{\alpha\in\mathcal{A}_i} \mathcal{L}_i^\alpha(x^\rho, u^\rho(x^\rho), p^\rho, X^\rho) + \rho(u_i^\rho - \mathcal{M}_i u^\rho)^+(x^\rho) \leq 0, \quad \forall \rho \in \mathbb{N}. \tag{3.3.2}$$

Then it follows from the boundedness of coefficients that there exists a constant $C > 0$ such that

$$u_i^\rho(x^\rho) - \mathcal{M}_i u^\rho(x^\rho) \leq (u_i^\rho - \mathcal{M}_i u^\rho)^+(x^\rho) \leq C/\rho,$$

hence by letting $\rho \to \infty$ and using Lemma 3.2.2 (3), we deduce that

$$u_i^*(x) = \lim_{\rho\to\infty} u_i^\rho(x^\rho) \leq \limsup_{\rho\to\infty}(\mathcal{M}_i u^\rho(x^\rho) + C/\rho) \leq \mathcal{M}_i u^*(x).$$

On the other hand, (3.3.2) yields for any $\alpha \in \mathcal{A}_i$, $\mathcal{L}_i^\alpha(x^\rho, u^\rho(x^\rho), p^\rho, X^\rho) \leq 0$, which implies

$$
\begin{aligned}
&- \operatorname{tr}[a_i^\alpha(x)X] - b_i^\alpha(x)p + c_i^\alpha(x)u_i^*(x) - \ell_i^\alpha(x) \\
&\quad \leq \sum_{j \in \mathcal{I}^{-i}} \limsup_{\rho\to\infty} d_{ij}^\alpha(x^\rho)u_j^\rho(x^\rho) \leq \sum_{j \in \mathcal{I}^{-i}} d_{ij}^\alpha(x)u_j^*(x),
\end{aligned}
\tag{3.3.3}
$$

where we have used the fact that $\lim_{\rho\to\infty} d_{ij}^\alpha(x^\rho) = d_{ij}^\alpha(x) \geq 0$. Then by taking the supremum over α, we have $\sup_{\alpha\in\mathcal{A}_i} \mathcal{L}_i^\alpha(x, u^*(x), p, X) \leq 0$, which shows u^* is a subsolution to (3.2.1).

Then we proceed to study u_* by fixing $x \in \mathbb{R}^d$, $i \in \mathcal{I}$ and $(p, X) \in J^{2,-}(u_*)_i(x)$. Let $(x^\rho, p^\rho, X^\rho)_{\rho\in\mathbb{N}}$ be a sequence such that $(p^\rho, X^\rho) \in J^{2,-}u_i^\rho(x^\rho)$ for each ρ and $(x^\rho, u_i^\rho(x^\rho), p^\rho, X^\rho) \to (x, (u_*)_i(x), p, X)$ as $\rho \to \infty$. Then the supersolution property of u^ρ implies for each $n \in \mathbb{N}$,

$$
\sup_{\alpha\in\mathcal{A}_i} \mathcal{L}_i^\alpha(x^\rho, u^\rho(x^\rho), p^\rho, X^\rho) + \rho(u_i^\rho - \mathcal{M}_i u^\rho)^+(x^\rho) \geq 0. \tag{3.3.4}
$$

Suppose that $\limsup_{\rho\to\infty} \rho(u_i^\rho - \mathcal{M}_i u^\rho)^+(x^\rho) > 0$, then, by possibly passing to a subsequence, we have $u_i^\rho(x^\rho) > (\mathcal{M}_i u^\rho)(x^\rho)$ for all ρ. Then, by using Lemma 3.2.2 (3), we obtain that $(u_*)_i(x) \geq \liminf_{\rho\to\infty}(\mathcal{M}_i u^\rho)(x^\rho) \geq (\mathcal{M}_i u_*)(x)$.

On the other hand, suppose that $\limsup_{\rho\to\infty} \rho(u_i^\rho - \mathcal{M}_i u^\rho)^+(x^\rho) = 0$, then for any $\delta > 0$, by passing to a subsequence, we deduce that for large enough $\rho \in \mathbb{N}$, there exists $\alpha^{\rho,\delta} \in \mathcal{A}_i$ such that

$$
-\operatorname{tr}[a_i^{\alpha^{\rho,\delta}}(x^\rho)X^\rho] - b_i^{\alpha^{\rho,\delta}}(x^\rho)p^\rho + c_i^{\alpha^{\rho,\delta}}(x^\rho)u_i^\rho(x^\rho) - \ell_i^{\alpha^{\rho,\delta}}(x^\rho) - \sum_{j\in\mathcal{I}^{-i}} d_{ij}^{\alpha^{\rho,\delta}}(x^\rho)u_j^\rho(x^\rho) \geq -\delta.
$$

Since $\mathcal{A}_i$ is compact, we can assume $\alpha^{\rho,\delta} \to \alpha^\delta \in \mathcal{A}_i$ as $\rho \to \infty$. Then, by taking the limit inferior and using the fact $\liminf_{\rho\to\infty} d_{ij}^{\alpha^{\rho,\delta}}(x^\rho)u_j^\rho(x^\rho) \geq d_{ij}^{\alpha^\delta}(x)(u_*)_j(x)$, we have

$$
-\operatorname{tr}[a_i^{\alpha^\delta}(x)X] - b_i^{\alpha^\delta}(x)p + c_i^{\alpha^\delta}(x)(u_*)_i(x) - \ell_i^{\alpha^\delta}(x) - \sum_{j\in\mathcal{I}^{-i}} d_{ij}^{\alpha^\delta}(x)(u_*)_j(x) \geq -\delta, \tag{3.3.5}
$$

from which by taking the supremum over α and sending $\delta \to 0$, we have that $\sup_{\alpha\in\mathcal{A}_i} \mathcal{L}_i^\alpha(x, u^*(x), p, X) \geq 0$, and conclude that u_* is a supersolution to (3.2.1).

Finally, by using Proposition 3.2.1, we have $u := u^* = u_*$ is the unique continuous viscosity solution of (3.2.1) and consequently (u^ρ) converges to u locally uniformly. □

Remark 3.3.1. Theorem 3.3.2 provides us with a constructive proof for the existence of solutions of (3.2.1) based on penalty approximations. Moreover, since the convergence analysis relies only on the comparison principle of (3.2.1) and the local boundedness of $(u^\rho)_{\rho\geq 0}$, it is possible to extend the results to nonlocal non-convex/non-concave systems with coefficients of polynomial growth.

3.4 Error estimates for penalty approximations

In this section, we shall proceed to analyze the convergence rate of the penalty approximation for (3.2.1). As pointed out in Section 3.1, unlike the variational inequalities [64], the non-strict monotonicity of the term $u_i - \mathcal{M}_i u$ prevents us from obtaining an upper bound of $u^\rho - u$ by constructing a subsolution of (3.2.1) directly from the penalized equations, which significantly complicates the error analysis. We shall overcome this difficulty by regularizing the HJBQVIs, and recover the same convergence rates (up to a logarithmic term) as those for conventional obstacle problems.

3.4.1 Regularization of HJBQVIs

In this section, we approximate (3.2.1) by a sequence of obstacle problems, through the iterated optimal stopping approximation (see e.g. [77, 92, 39] for its application to QVIs). We shall quantify the approximation errors of these obstacle problems depending on the regularity of the solution, which we also establish.

Let $u^0 = (u_i^0)_{i\in\mathcal{I}}$ be the viscosity solution of the following system of HJB equations:

$$\sup_{\alpha\in\mathcal{A}_i} \mathcal{L}_i^\alpha(x, u(x), Du_i(x), D^2u_i(x)) = 0, \quad x \in \mathbb{R}^d,\, i \in \mathcal{I}. \tag{3.4.1}$$

We then inductively define a sequence of functions $\{u^n\}_{n\in\mathbb{N}}$, where for each $n \in \mathbb{N}$, i.e., $n > 0$, given functions u^{n-1}, let $u^n = (u_i^n)_{i\in\mathcal{I}}$ be the viscosity solution to the following obstacle problem:

$$\max\Big\{ \sup_{\alpha\in\mathcal{A}_i} \mathcal{L}_i^\alpha(x, u^n(x), Du_i^n(x), D^2u_i^n(x));\ (u_i^n - \mathcal{M}_i u^{n-1})(x) \Big\} = 0, \quad x \in \mathbb{R}^d,\, i \in \mathcal{I}. \tag{3.4.2}$$

Under Assumptions 3, 4, 5 and 7, one can establish the comparison principles for (3.4.1) and (3.4.2), and then demonstrate the existence of $u^n \in [C^{0,1}(\mathbb{R}^d)]^M$ for each $n \geq 0$ (see Theorem 3.4.4 for the Lipschitz regularity). Moreover, by using the comparison principle of (3.4.2), we can further deduce from an inductive argument that $u^{n-1} \geq u^n$ for all $n \in \mathbb{N}$.

The following proposition estimates the approximation error $u^n - u$, which extends the results in [20, 39] to weakly coupled systems with (possibly) negative running cost $(\ell_i)_{i\in\mathcal{I}}$.

Proposition 3.4.1. *Suppose Assumptions 3, 4, 5 and 7 hold, and let u and u^n be the viscosity solution to (3.2.1) and (3.4.2), respectively. Then there exist constants $\mu \in (0, 1]$ and $C \geq 0$ such that*

$$0 \leq u^n - u \leq C(1-\mu)^n, \quad n \geq 0.$$

Consequently, the iterates $(u^n)_{n\geq 0}$ *are bounded uniformly in* n.

Proof. We adapt the arguments for [88, Theorem 3.4] to the current continuous setting, and present the main steps in Appendix A.1 for the reader's convenience. □

Now we turn to investigate the regularity of solutions to (3.4.2) based on different assumptions on the coefficients. We shall first focus on the following variational inequalities:

$$\max\left\{\sup_{\alpha\in\mathcal{A}_i}\mathcal{L}_i^\alpha(x,u(x),Du_i(x),D^2u_i(x));\ (u_i-\Psi_i)(x)\right\}=0,\quad i\in\mathcal{I},\qquad(3.4.3)$$

with given obstacles $(\Psi_i)_{i\in\mathcal{I}}$, which serves as a general form of the iterative equations (3.4.2).

The following result shows the Lipschitz continuity of the solution to the obstacle problem (3.4.3), which can be proved by using the standard doubling of variables technique (see e.g. [23, 39]).

Proposition 3.4.2. *Suppose Assumptions 3 and 5 hold, and* $\Psi\in[C^{0,1}(\mathbb{R}^d)]^M$. *Then the viscosity solution* u *to* (3.4.3) *is Lipschitz continuous with constant* $\sup_{i\in\mathcal{I}}[u_i]_1\leq\max(C,\sup_{i\in\mathcal{I}}[\Psi_i]_1)$, *where* C *is a constant independent of* $[\Psi_i]_1$.

We then proceed to study higher regularity of the solutions, which enables us to deduce a higher convergence rate of the penalty approximation. The next proposition extends the results in [55, 53] to weakly coupled systems, and asserts that if the coefficients are sufficiently regular, then the solution to the obstacle problem (3.4.3) is semiconcave.

Note that instead of viewing the obstacle problem (3.4.3) as a convex HJB equation as is studied in [55, 53], we shall separately analyze the obstacle part and the HJB part of (3.4.3), which leads to a sharper estimate for the semiconcavity constant of u in terms of Ψ. Moreover, instead of requiring $\Psi_i\in W^{2,\infty}(\mathbb{R}^d)$ as in [55, 53] (which essentially means Ψ_i is differentiable with bounded and Lipschitz continuous derivative), we only assume the obstacles to be semiconcave, which is crucial for the subsequent analysis of penalty errors.

Proposition 3.4.3. *Suppose Assumptions 3 and 6 hold. Assume further that the constant* λ_0 *in Assumption 3 is sufficiently large and the obstacle* $\Psi\in[C^{0,1}(\mathbb{R}^d)]^M$ *is semiconcave. Then the viscosity solution* $u\in[C^{0,1}(\mathbb{R}^d)]^M$ *to* (3.4.3) *is semiconcave with a constant satisfying*

$$\sup_{i\in\mathcal{I}}[u_i]_{2,+}\leq\max\left\{C\sup_{i\in\mathcal{I}}|u_i|_1,\sup_{i\in\mathcal{I}}[\Psi_i]_{2,+}\right\},$$

for some constant C, independent of $[\Psi_j]_{2,+}$, $[\Psi_j]_1$ and $[u_j]_1$ for all $j \in \mathcal{I}$.

Proof. For $\delta, \varepsilon, \gamma > 0$, we define for all $x, y, z \in \mathbb{R}^d$ that

$$\begin{aligned}\phi(x,y,z) &= \delta|x-y|^4 + \varepsilon|x+y-2z|^2 + \gamma|x|^2,\\ \Phi_i(x,y,z) &= u_i(x) + u_i(y) - 2u_i(z) - \phi(x,y,z),\end{aligned}$$

and let $m_{\delta,\varepsilon,\gamma} := \sup_{(x,y,z)\in\mathbb{R}^{3d}, i\in\mathcal{I}} \Phi_i(x,y,z)$. By the finiteness of $\mathcal{I}$, the boundedness and continuity of $(u_i)_{i\in\mathcal{I}}$, and the penalization term ϕ, there exists $i \in \mathcal{I}$, independent of $\delta, \varepsilon, \gamma$, and $(\bar{x}^{\delta,\varepsilon,\gamma}, \bar{y}^{\delta,\varepsilon,\gamma}, \bar{z}^{\delta,\varepsilon,\gamma}) \in \mathbb{R}^{3d}$ such that $m_{\delta,\varepsilon,\gamma} = \Phi_i(\bar{x}^{\delta,\varepsilon,\gamma}, \bar{y}^{\delta,\varepsilon,\gamma}, \bar{z}^{\delta,\varepsilon,\gamma})$. In the following we shall omit the dependence on $\delta, \varepsilon, \gamma$ for notational simplicity. Then we can deduce from [30, Theorem 3.2] that for any $\theta > 1$, there exist $X, Y, Z \in \mathbb{S}^d$ such that

$$(p_x, X) \in \bar{J}^{2,+}u(\bar{x}), \quad (p_y, Y) \in \bar{J}^{2,+}u(\bar{y}), \quad (-p_z/2, -Z/2) \in \bar{J}^{2,-}u(\bar{z}),$$

where $(p_x, p_y, p_z) = (D_x\phi(\bar{x},\bar{y},\bar{z}), D_y\phi(\bar{x},\bar{y},\bar{z}), D_z\phi(\bar{x},\bar{y},\bar{z}))$, and

$$\begin{pmatrix} X & 0 & 0\\ 0 & Y & 0\\ 0 & 0 & Z\end{pmatrix} \le \theta D^2\phi(\bar{x},\bar{y},\bar{z}).$$

Hence, by the definition of viscosity solution, we obtain that

$$\begin{aligned}\max\Big\{\sup_{\alpha\in\mathcal{A}_i}\mathcal{L}_i^\alpha(\bar{x}, u(\bar{x}), p_x, X);\ (u_i - \Psi_i)(\bar{x})\Big\} &\le 0,\\ \max\Big\{\sup_{\alpha\in\mathcal{A}_i}\mathcal{L}_i^\alpha(\bar{y}, u(\bar{y}), p_y, Y);\ (u_i - \Psi_i)(\bar{y})\Big\} &\le 0,\\ \max\Big\{\sup_{\alpha\in\mathcal{A}_i}\mathcal{L}_i^\alpha(\bar{z}, u(\bar{z}), -p_z/2, -Z/2);\ (u_i - \Psi_i)(\bar{z})\Big\} &\ge 0.\end{aligned} \tag{3.4.4}$$

Now we discuss two cases. Suppose the maximum in the third inequality of (3.4.4) is attained by its second argument, then we obtain from (3.4.4) that

$$\begin{aligned}u_i(\bar{x}) + u_i(\bar{y}) - 2u_i(\bar{z}) &\le \Psi_i(\bar{x}) + \Psi_i(\bar{y}) - 2\Psi_i(\bar{z})\\ &= \Psi_i(\bar{x}) + \Psi_i(\bar{y}) - 2\Psi_i(\frac{\bar{x}+\bar{y}}{2}) + 2\Psi_i(\frac{\bar{x}+\bar{y}}{2}) - 2\Psi_i(\bar{z})\\ &\le [\Psi_i]_{2,+}|\bar{x}-\bar{y}|^2/4 + [\Psi_i]_1|\bar{x}+\bar{y}-2\bar{z}|,\end{aligned}$$

where we have used the Lipschitz continuity and semiconcavity of Ψ_i. Thus, the definition of $m_{\delta,\varepsilon,\gamma}$ and the fact that $\sup_{r>0}(-\delta r^2 + Cr) = C^2/(4\delta)$ give us that

$$\begin{aligned}m_{\delta,\varepsilon,\gamma} &\le [\Psi_i]_{2,+}|\bar{x}-\bar{y}|^2/4 + [\Psi_i]_1|\bar{x}+\bar{y}-2\bar{z}| - \delta|\bar{x}-\bar{y}|^4 - \varepsilon|\bar{x}+\bar{y}-2\bar{z}|^2\\ &\le \frac{[\Psi_i]_{2,+}^2}{64\delta} + \frac{[\Psi_i]_1^2}{4\varepsilon}.\end{aligned}$$

Thus, by letting $\gamma \to \infty$, we have for all $x, y, z \in \mathbb{R}^d$ that

$$u_i(x) + u_i(y) - 2u_i(z) \le \frac{[\Psi_i]_{2,+}^2}{64\delta} + \frac{[\Psi_i]_1^2}{4\varepsilon} + \delta|x-y|^4 + \varepsilon|x+y-2z|^2, \quad \forall \delta, \varepsilon > 0,$$

from which by minimizing over δ, ε separately, and setting $x = z + h$, $y = z - h$, we get that

$$u_i(z+h) + u_i(z-h) - 2u_i(z) \le [\Psi_i]_{2,+}|2h|^2/4 = [\Psi_i]_{2,+}|h|^2. \tag{3.4.5}$$

On the other hand, suppose the maximum in the third inequality of (3.4.4) is attained by the first argument, then for any $\eta > 0$, there exists $\alpha^\eta \in \mathcal{A}$ such that the following inequality holds:

$$\mathcal{L}_i^{\alpha^\eta}(\bar{x}, u(\bar{x}), p_x, X) + \mathcal{L}_i^{\alpha^\eta}(\bar{y}, u(\bar{y}), p_y, Y) - 2(\mathcal{L}_i^{\alpha^\eta}(\bar{z}, u(\bar{z}), -p_z/2, -Z/2) + \eta) \le 0.$$

More precisely, we have

$$\begin{aligned}
&-\mathrm{tr}[a_i^{\alpha^\eta}(\bar{x})X + a_i^{\alpha^\eta}(\bar{y})Y + a_i^{\alpha^\eta}(\bar{z})Z] - [b_i^{\alpha^\eta}(\bar{x})p_x + b_i^{\alpha^\eta}(\bar{y})p_y + b_i^{\alpha^\eta}(\bar{z})p_z] \\
&\quad + c_i^{\alpha^\eta}(\bar{x})u_i(\bar{x}) + c_i^{\alpha^\eta}(\bar{y})u_i(\bar{y}) - 2c_i^{\alpha^\eta}(\bar{z})u_i(\bar{z}) - [\ell_i^{\alpha^\eta}(\bar{x}) + \ell_i^{\alpha^\eta}(\bar{y}) - 2\ell_i^{\alpha^\eta}(\bar{z})] - 2\eta \\
&\quad \le \sum_{j \in \mathcal{I}^{-i}} d_{ij}^{\alpha^\eta}(\bar{x})u_j(\bar{x}) + d_{ij}^{\alpha^\eta}(\bar{y})u_j(\bar{y}) - 2d_{ij}^{\alpha^\eta}(\bar{z})u_j(\bar{z}).
\end{aligned} \tag{3.4.6}$$

Comparing with [53, Theorem 5 (ii)], it remains to estimate the terms in the last line of the above inequality. Note for any given function $g \in C_1^0(\mathbb{R}^d)$ and $x, z \in \mathbb{R}^d$, we have that

$$\begin{aligned}
|g(x) - g(z)| &\le \left|g(x) - g(\frac{x+y}{2})\right| + \left|g(\frac{x+y}{2}) - g(z)\right| \\
&\le [g]_1 \frac{|x-y|}{2} + (|g|_0[g]_1|x+y-2z|)^{1/2} \\
&\le |g|_1(|x-y| + |x+y-2z|^{1/2}), \quad \forall y \in \mathbb{R}^d.
\end{aligned}$$

Therefore, we obtain for each $j \in \mathcal{I}^{-i}$ that

$$\begin{aligned}
&d_{ij}^{\alpha^\eta}(\bar{x})u_j(\bar{x}) + d_{ij}^{\alpha^\eta}(\bar{y})u_j(\bar{y}) - 2d_{ij}^{\alpha^\eta}(\bar{z})u_j(\bar{z}) \\
&= d_{ij}^{\alpha^\eta}(\bar{z})(u_j(\bar{x}) + u_j(\bar{y}) - 2u_j(\bar{z})) + \big(d_{ij}^{\alpha^\eta}(\bar{x}) + d_{ij}^{\alpha^\eta}(\bar{y}) - 2d_{ij}^{\alpha^\eta}(\bar{z})\big)u_j(\bar{z}) \\
&\quad + (d_{ij}^{\alpha^\eta}(\bar{x}) - d_{ij}^{\alpha^\eta}(\bar{z}))(u_j(\bar{x}) - u_j(\bar{z})) + (d_{ij}^{\alpha^\eta}(\bar{y}) - d_{ij}^{\alpha^\eta}(\bar{z}))(u_j(\bar{y}) - u_j(\bar{z})) \\
&\le d_{ij}^{\alpha^\eta}(\bar{z})(u_i(\bar{x}) + u_i(\bar{y}) - 2u_i(\bar{z})) + \big([Dd_{ij}^{\alpha^\eta}]_1|\bar{x} - \bar{y}|^2/4 + [d_{ij}^{\alpha^\eta}]_1|\bar{x} + \bar{y} - 2\bar{z}|\big)|u_j|_0 \\
&\quad + 4|d_{ij}^{\alpha^\eta}|_1|u_j|_1(|x-y|^2 + |x+y-2z|).
\end{aligned}$$

Then if λ_0 is sufficiently large, we can proceed along lines of the proof of [53, Theorem 5 (ii)], and deduce that there exists a constant $C \geq 0$, independent of $[u_i]_1$ for any $i \in \mathcal{I}$, such that it holds for all $z, h \in \mathbb{R}^d$ and $i \in \mathcal{I}$ that $u_i(z+h)+u_i(z-h)-2u_i(z) \leq C(\sup_{i\in\mathcal{I}} |u_i|_1)|h|^2$ (cf. equation (5.8) in [53]), which together with (3.4.5) completes our proof. □

With Propositions 3.4.1 and 3.4.3 in hand, we are ready to present the following upper bounds of the Lipschitz and semiconcavity constants of the iterates $(u^n)_{n\in\mathbb{N}}$ defined as in (3.4.1) and (3.4.2).

Theorem 3.4.4. *Suppose Assumptions 3, 4, 5 and 7 hold, then for any $n \in \mathbb{N}$, the iterate u^n is Lipschitz continuous with a constant satisfying $\sup_{i\in\mathcal{I}}[u_i^n]_1 \leq Cn$, where C is a constant independent of n. If we further assume Assumptions 6 and 8 hold, and the constant λ_0 in Assumption 3 is sufficiently large, then the iterate u^n is semiconcave with a constant satisfying $\sup_{i\in\mathcal{I}}[u_i^n]_{2,+} \leq Cn$.*

Proof. It is well understood that the solution u^0 to a weakly coupled system with convex Hamiltonians is Lipschitz continuous under Assumptions 3, 4 and 5 (see [23]), and is semiconcave if the coefficients enjoy higher regularity (see the second case in the proof of Proposition 3.4.3). We now use an inductive argument to estimate the regularity of the iterates $(u^n)_{n\in\mathbb{N}}$.

It has been shown in Proposition 3.4.1 that $(u^n)_{n\in\mathbb{N}}$ are bounded uniformly in n. Now suppose u^{n-1} is Lipschitz continuous, and Assumption 7 holds. Then we can deduce from Proposition 3.4.2 that

$$\sup_{i\in\mathcal{I}}[u_i^n]_1 \leq \max\left(C', \sup_{i\in\mathcal{I}}[\mathcal{M}_i u^{n-1}]_1\right) \leq \max\left(C', \sup_{i\in\mathcal{I}}[u_i^{n-1}]_1 + C\right), \tag{3.4.7}$$

where C' is a constant independent of n, and C is the constant in Assumption 7. An inductive argument enables us to conclude the desired estimate $[u_i^n]_1 = \mathcal{O}(n)$ for all i.

Moreover, by further assuming Assumption 8 and the assumptions of Proposition 3.4.3, we can obtain for all $i \in \mathcal{I}$ the following estimate:

$$\begin{aligned}
[u_i^n]_{2,+} &\leq \max\left(C' \sup_{i\in\mathcal{I}} |u_i^n|_1, \sup_{i\in\mathcal{I}}[\mathcal{M}_i u_i^{n-1}]_{2,+}\right) \\
&\leq \max\left(C' \sup_{i\in\mathcal{I}} |u_i^n|_1, \sup_{i\in\mathcal{I}}[u_i^{n-1}]_{2,+} + C\right), \quad i \in \mathcal{I},
\end{aligned} \tag{3.4.8}$$

where C' is a constant independent of n, and C is the constant in Assumption 8. Then, by using the previous Lipschitz estimates of $(u^n)_{n\in\mathbb{N}}$, we conclude from (3.4.8) that $[u_i^n]_{2,+} = \mathcal{O}(n)$ for all i. □

Remark 3.4.1. Suppose Assumptions 7 and 8 hold with $C = 0$ (e.g. the intervention operator $\mathcal{M}$ is of the form (3.2.7)), then the estimates (3.4.7) and (3.4.8) hold with $C = 0$. Thus one can show inductively that for any $n \geq 0$, the Lipschitz constant $[u^n]_1$ and the semiconcavity constant $[u^n]_{2,+}$ of the iterate u^n are uniformly bounded in terms of n, which along with Proposition 3.4.1, imply that the solution to HJBQVI (3.2.1) is Lipschitz continuous and semiconcave. As we shall see in Remark 3.4.2, this observation enables us to improve the convergence rate of the penalty approximation by a log factor.

3.4.2 Regularization of penalized equations

In this section, we shall propose a sequence of auxiliary problems to the penalized equation (3.3.1) with a fixed parameter $\rho > 0$, which is similar to the regularization of the QVI (3.2.1) discussed in Section 3.4.1. These auxiliary problems will serve as an important tool for quantifying convergence orders of the penalty approximations.

More precisely, for any given penalty parameter $\rho > 0$, we shall consider the following sequence of auxiliary problems: let $u^{\rho,0} = u^0$ be the solution to (3.4.1), and for each $n \geq 1$, given $u^{\rho,n-1}$, let $u^{\rho,n}$ be the solution to the following equations:

$$\sup_{\alpha \in \mathcal{A}_i} \mathcal{L}_i^\alpha(x, u^{\rho,n}(x), Du_i^{\rho,n}(x), D^2 u_i^{\rho,n}(x)) + \rho(u_i^{\rho,n} - \mathcal{M}_i u^{\rho,n-1})^+(x) = 0, \quad i \in \mathcal{I}. \tag{3.4.9}$$

The above iterates $(u^{\rho,n})_{n\geq 0}$ can be equivalently expressed as $u^{\rho,n} = Q^\rho u^{\rho,n-1}$ for all $n \in \mathbb{N}$ with an operator $Q^\rho : [C^{0,1}(\mathbb{R}^d)]^M \mapsto [C^{0,1}(\mathbb{R}^d)]^M$ defined as follows: for any given u, $Q^\rho u$ is defined as the unique solution to the following equations:

$$\sup_{\alpha \in \mathcal{A}_i} \mathcal{L}_i^\alpha(x, Q^\rho u(x), D(Q^\rho u)_i(x), D^2 (Q^\rho u)_i(x)) + \rho((Q^\rho u)_i - \mathcal{M}_i u)^+(x) = 0, \quad i \in \mathcal{I}. \tag{3.4.10}$$

We now present some important properties of the operator Q^ρ. Suppose Assumptions 3 and 7 hold, then for any given $u \in [C^{0,1}(\mathbb{R}^d)]^M$, we see $\mathcal{M}u \in [C^{0,1}(\mathbb{R}^d)]^M$, from which one can establish the comparison principle of (3.4.10) and the well-posedness of (3.4.10) in the class of bounded continuous functions.

The following lemma strengthens the comparison principle by indicating that Q^ρ is monotone and Lipschitz continuous with constant 1, which is essential for the error estimates in Section 3.4.3. The proof is included in Appendix A.1.

Lemma 3.4.5. *Suppose Assumptions 3 and 7 hold. Then for any $u, v \in [C^{0,1}(\mathbb{R}^d)]^M$, we have*

$$\sup_{i \in \mathcal{I}} \left| \big((Q^\rho u)_i - (Q^\rho v)_i\big)^+ \right|_0 \leq \sup_{i \in \mathcal{I}} \left| (u_i - v_i)^+ \right|_0 .$$

Consequently, if $u \le v$, then $Q^\rho u \le Q^\rho v$.

Finally, a straightforward modification of the doubling arguments for Lemma 3.4.5 enables us to show that under Assumptions 3, 4, 5 and 7, the solution $Q^\rho u$ to (3.4.10) is in fact Lipschitz continuous provided that u is Lipschitz continuous, which subsequently implies the iterates $(u^{\rho,n})_{n\ge 0}$ are well-defined functions in $[C^{0,1}(\mathbb{R}^d)]^M$. We omit the proof of these Lipschitz estimates, by pointing out that the analysis for the obstacle part is exactly the same as those for Lemma 3.4.5, and referring the reader to [23] for a discussion on the HJB part.

We then proceed to study the convergence of the iterates $(u^{\rho,n})_{n\ge 0}$. The next lemma shows the sequence $(u^{\rho,n})_{n\in\mathbb{N}}$ is monotone and uniformly bounded.

Lemma 3.4.6. *Suppose Assumptions 3, 4, 5 and 7 hold, and $\rho > 0$ is a fixed penalty parameter. Then the iterates $(u^{\rho,n})_{n\ge 0}$ are monotonically decreasing and uniformly bounded in terms of n.*

Proof. Note that the comparison principle of (3.4.10) yields $u^{\rho,0} \ge u^{\rho,1}$, which together with the monotonicity of Q^ρ leads to $u^{\rho,n-1} \ge u^{\rho,n}$ for all $n \ge 1$. Now we show by induction that $u^{\rho,n} \ge u^\rho$ for all $n \ge 0$, where u^ρ is the solution to (3.3.1). The statement holds clearly for $n = 0$. Suppose for some $n \in \mathbb{N}$, we have $u^{\rho,n-1} \ge u^\rho$, then Lemma 3.2.2 (2) implies $\mathcal{M}_i u^{\rho,n-1} \ge \mathcal{M}_i u^\rho$ for all $i \in \mathcal{I}$, and hence $\rho(u^\rho - \mathcal{M}_i u^\rho)^+ \ge \rho(u^\rho - \mathcal{M}_i u^{\rho,n-1})^+$, which implies u^ρ is a subsolution of the equation for $u^{\rho,n}$. Consequently, we obtain from the comparison principle that $u^{\rho,n} \ge u^\rho$. □

The following theorem presents the convergence of $(u^{\rho,n})_{n\ge 0}$ to the solution of (3.3.1).

Theorem 3.4.7. *Suppose Assumptions 3, 4, 5 and 7 hold. Then for any given $\rho \ge 0$, the iterates $(u^{\rho,n})_{n\ge 0}$ converge monotonically from above to the solution u^ρ of (3.3.1) as $n \to \infty$.*

Proof. With the comparison principle of (3.3.1) (Proposition 3.3.1) in mind, it remains to show the component-wise relaxed half-limit $u^{*,\rho}$ (resp. u^ρ_*) of $(u^{\rho,n})_{n\ge 0}$ is a subsolution (resp. supersolution) to (3.3.1). For notational simplicity, we shall omit the dependence on ρ in the subsequent analysis if no confusion can occur.

Let $x \in \mathbb{R}^d$, $i \in \mathcal{I}$ and $(p, X) \in J^{2,+}u_i^*(x)$. It follows from [30, Lemma 6.1] that there exist $(x^n, p^n, X^n)_{n\in\mathbb{N}}$ such that $(p^n, X^n) \in J^{2,+}u_i^{\rho,n}(x^n)$ for each n and $(x^n, u_i^{\rho,n}(x^n), p^n, X^n) \to (x, u_i^*(x), p, X)$ as $n \to \infty$. Then we have for all $n \in \mathbb{N}$ that

$$\sup_{\alpha\in\mathcal{A}_i} \mathcal{L}_i^\alpha(x^n, u^{\rho,n}(x^n), p^n, X^n) + \rho(u_i^{\rho,n} - \mathcal{M}_i u^{\rho,n-1})^+(x^n) \le 0. \tag{3.4.11}$$

Note that by Lemma 3.2.2 (3), we have

$$\liminf_{n\to\infty}(u_i^{\rho,n} - \mathcal{M}_i u^{\rho,n-1})^+(x^n) \geq \liminf_{n\to\infty}(u_i^{\rho,n} - \mathcal{M}_i u^{\rho,n-1})(x^n) \geq (u_i^* - \mathcal{M}_i u^*)(x),$$

which implies $\liminf_{n\to\infty}(u_i^{\rho,n} - \mathcal{M}_i u^{\rho,n-1})^+(x^n) \geq (u_i^* - \mathcal{M}_i u^*)^+(x)$. Thus, for any given $\alpha \in \mathcal{A}_i$, we can take limit inferior in (3.4.11) and obtain from the inequality $\liminf_{n\to\infty} -d_{ij}^\alpha(x^\rho)u_j^\rho(x^\rho) \geq -d_{ij}^\alpha(x)u_j^*(x)$ (see (3.3.3)) that

$$\begin{aligned}
&\mathcal{L}_i^\alpha(x, u^*(x), p, X) + \rho(u_i^* - \mathcal{M}_i u^*)^+(x) \\
&\leq \liminf_{n\to\infty} \mathcal{L}_i^\alpha(x^n, u^{\rho,n}(x^n), p^n, X^n) + \rho(u_i^* - \mathcal{M}_i u^*)^+(x) \leq 0.
\end{aligned}$$

Then taking the supremum over $\alpha \in \mathcal{A}_i$ gives us the desired result.

We then turn to study u_* by fixing $x \in \mathbb{R}^d$, $i \in \mathcal{I}$ and $(p, X) \in J^{2,-}(u_*)_i(x)$. Let $(x^n, p^n, X^n)_{n\in\mathbb{N}}$ be a sequence such that $(p^n, X^n) \in J^{2,-}u_i^{\rho,n}(x^n)$ for each n and $(x^n, u_i^{\rho,n}(x^n), p^n, X^n) \to (x, (u_*)_i(x), p, X)$ as $n \to \infty$. Then for all $n \in \mathbb{N}$, the supersolution property of $u^{\rho,n}$ implies

$$\mathcal{L}_i^{\alpha^n}(x^n, u^{\rho,n}(x^n), p^n, X^n) = \sup_{\alpha\in\mathcal{A}_i} \mathcal{L}_i^\alpha(x^n, u^{\rho,n}(x^n), p^n, X^n) \geq -\rho(u_i^{\rho,n} - \mathcal{M}_i u^{\rho,n-1})^+(x^n),$$

for some $\alpha^n \in \mathcal{A}_i$. Then by taking limit superior as $n \to \infty$ on both sides of the above inequality and using similar arguments as (3.3.5), we obtain that

$$\begin{aligned}
\sup_{\alpha\in\mathcal{A}_i} \mathcal{L}_i^\alpha(x, (u_*)_i(x), p, X) &\geq \limsup_{n\to\infty} -\rho(u_i^{\rho,n} - \mathcal{M}_i u^{\rho,n-1})^+(x^n) \\
&\geq -\rho \limsup_{n\to\infty}(u_i^{\rho,n} - \mathcal{M}_i u^{\rho,n-1})^+(x^n).
\end{aligned}$$

Now it remains to show

$$m := \limsup_{n\to\infty}(u_i^{\rho,n} - \mathcal{M}_i u^{\rho,n-1})^+(x^n) \leq ((u_*)_i - \mathcal{M}_i u_*)^+(x).$$

We shall assume without loss of generality that $m > 0$. Then by extracting a subsequence, we can further assume $u_i^{\rho,n}(x^n) > \mathcal{M}_i u^{\rho,n-1}(x^n)$ for n, and $\lim_{n\to\infty}(u_i^{\rho,n}(x^n) - \mathcal{M}_i u^{\rho,n-1}(x^n)) = m$. These properties along with Lemma 3.2.2 (3) yield

$$m = \limsup_{n\to\infty}(u_i^{\rho,n}(x^n) - \mathcal{M}_i u^{\rho,n-1}(x^n)) \leq (u_*)_i(x) - \mathcal{M}_i(u_*)(x) \leq ((u_*)_i - \mathcal{M}_i u_*)^+(x),$$

which finishes the proof of the statement that u_* is a supersolution of (3.3.1). □

3.4.3 Convergence rates of value functions

In this section, we shall exploit the regularization procedures discussed in Sections 3.4.1 and 3.4.2 to estimate the convergence rates of the penalty approximation, depending on the regularity of the coefficients.

Let us first recall the penalty errors for the classical obstacle problem, which have been analyzed in [64, 87] and play an important role in our error estimates. To avoid confusion with the solutions to (3.2.1) and (3.3.1), we shall denote by v the solution to the obstacle problem (3.4.3), and by v^ρ the solution to the following penalized equation with a given parameter $\rho \geq 0$:

$$\sup_{\alpha\in\mathcal{A}_i} \mathcal{L}_i^\alpha(x, v^\rho(x), Dv_i^\rho(x), D^2v_i^\rho(x)) + \rho(v_i^\rho - \Psi_i)^+(x) = 0, \quad i \in \mathcal{I}. \tag{3.4.12}$$

Proposition 3.4.8. *For any given penalty parameter $\rho > 0$, let v and v^ρ be the solution to the obstacle problem (3.4.3) and the penalized equation (3.4.12), respectively. Suppose Assumption 3 holds and $\Psi_i \in C^{0,1}(\mathbb{R}^d)$ for all $i \in \mathcal{I}$. Then there exists a constant C, independent of $(\Psi_i)_{i\in\mathcal{I}}$, such that*

$$0 \leq v_i^\rho(x) - v_i(x) \leq C\left(\sup_{j\in\mathcal{I}} |\Psi_j|_1\right)\rho^{-1/2}, \quad x \in \mathbb{R}^d, i \in \mathcal{I}. \tag{3.4.13}$$

If, in addition, Ψ_i is semiconcave for all $i \in \mathcal{I}$, then we have

$$0 \leq v_i^\rho(x) - v_i(x) \leq C \sup_{j\in\mathcal{I}}\Big(|\Psi_j|_1 + [\Psi_j]_{2,+}\Big)/\rho, \quad x \in \mathbb{R}^d, i \in \mathcal{I}. \tag{3.4.14}$$

Proof. The statement extends the results for scalar HJB equations studied in [64], and can be established by using similar arguments. The main step is to observe that for any given constant C_ρ satisfying $C_\rho \geq \rho(v_i^\rho - \Psi_i)^+$ for all $i \in \mathcal{I}$, $v^\rho - C_\rho/\rho$ is a subsolution to (3.4.3), which implies $v^\rho - v \leq C_\rho/\rho$. Then if we suppose that $\Psi \in [C^2(\mathbb{R}^d)]^M$, one can deduce that there exists a constant $C > 0$, independent of ρ and $(\Psi_i)_{i\in\mathcal{I}}$, such that the upper bound $C_\rho \leq C \sup_i \big(|\Psi_i|_1 + |(D^2\Psi_i)^+|_0\big)$ holds for all $\rho \geq 0$, which enables us to conclude (3.4.14) for smooth obstacles. Finally, we can regularize a general nonsmooth obstacle with mollifiers, and balance the approximation errors to obtain the desired error estimates (3.4.13) and (3.4.14). □

We then present the following elementary lemma, which extends [20, Lemma 6.1] to polynomials with higher degrees. The proof follows from a straightforward computation, which is included in Appendix A.1 for completeness.

Lemma 3.4.9. *For any given $\alpha > 0$, $\mu \in (0,1)$ and $\gamma \in \mathbb{N}$, consider the function $\phi^\alpha : (0,\infty) \to \mathbb{R}$, $\phi^\alpha(x) = \alpha x^\gamma + \mu^x$. Then there exists a constant $C > 0$, depending only on γ and μ, such that*

$$m^\alpha := \min_{n\in\mathbb{N}} \phi^\alpha(n) \le C\alpha(-\log\alpha)^\gamma, \quad \text{as } \alpha \to 0.$$

Now we are ready to state the main result of this chapter, which gives an upper bound of the penalization error $u^\rho - u$.

Theorem 3.4.10. *Let u and u^ρ solve the QVI (3.2.1) and the penalized problem (3.3.1), respectively. If Assumptions 3, 4, 5 and 7 hold, then for all large enough penalty parameter ρ, we have*

$$0 \le u_i^\rho(x) - u_i(x) \le C(\log\rho)^2\rho^{-1/2}, \quad x \in \mathbb{R}^d,\, i \in \mathcal{I}. \tag{3.4.15}$$

If we further assume Assumptions 6 and 8 hold, and the constant λ_0 in Assumption 3 is sufficiently large, then

$$0 \le u_i^\rho(x) - u_i(x) \le C(\log\rho)^2/\rho, \quad x \in \mathbb{R}^d,\, i \in \mathcal{I}, \tag{3.4.16}$$

for some constant C independent of the parameter ρ.

Proof. For notational simplicity, in the subsequent analysis, we shall denote by C a generic constant, which is independent of the iterate index n and the penalty parameter ρ, and may take a different value at each occurrence.

The monotone convergence (see Theorem 3.3.2) of $(u^\rho)_{\rho\ge 0}$ implies that $u_i^\rho - u_i \ge 0$ for any given $\rho \ge 0$, hence it remains to establish an upper bound of $u^\rho - u$. Note that we have

$$u_i^\rho - u_i = (u_i^\rho - u_i^{\rho,n}) + (u_i^{\rho,n} - u_i^n) + (u_i^n - u_i), \quad i \in \mathcal{I}, n \ge 0,$$

where $u_i^{\rho,n}$ and u_i^n solve (3.4.9) and (3.4.2), respectively. Since $u^{\rho,n} \ge u^\rho$ for any $n \in \mathbb{N}$ (see Theorem 3.4.7), we obtain from Proposition 3.4.1 that

$$u_i^\rho - u_i \le (u_i^{\rho,n} - u_i^n) + (u_i^n - u_i) \le |u^{\rho,n} - u^n|_0 + C\mu^n, \quad n \in \mathbb{N}, \tag{3.4.17}$$

for some constants $\mu \in (0,1)$ and $C > 0$.

We now estimate the term $|u^{\rho,n} - u^n|_0$. Since the operator Q^ρ is Lipschitz continuous with constant 1 (see Lemma 3.4.5), it holds for all $n \in \mathbb{N}$ that

$$|u^{\rho,n} - u^n|_0 \le |Q^\rho u^{\rho,n-1} - Q^\rho u^{n-1}|_0 + |Q^\rho u^{n-1} - u^n|_0 \le |u^{\rho,n-1} - u^{n-1}|_0 + |Q^\rho u^{n-1} - u^n|_0. \tag{3.4.18}$$

Then by letting the obstacle $\Psi_i = \mathcal{M}_i u^{n-1}$ for all $i \in \mathcal{I}$ in (3.4.3) and using Proposition 3.4.8, we can bound the last term in (3.4.18) depending on the regularity of the iterates $(u^n)_{n\geq 0}$.

In particular, under Assumptions 3, 4, 5 and 7, we know from Theorem 3.4.4 that u^n is Lipschitz continuous with constant $|u_i^n|_1 \leq Cn$ for all $i \in \mathcal{I}$ and $n \in \mathbb{N}$. Then by using the estimates (3.4.13), (3.4.17) and (3.4.18), we get $u_i^\rho - u_i \leq C(n^2\rho^{-1/2} + \mu^n)$ for all $n \in \mathbb{N}$, from which we can conclude (3.4.15) by applying Lemma 3.4.9 with $\alpha = \rho^{-1/2}$ and $\gamma = 2$.

Similarly, by further assuming Assumptions 6 and 8, and the constant λ_0 in Assumption 3 is sufficiently large, we obtain from Theorem 3.4.4 that $|u_i^n|_1 + [u_i^n]_{2,+} \leq Cn$ for all n, which implies $u_i^\rho - u_i \leq C(n^2/\rho + \mu^n)$ for all $n \in \mathbb{N}$, and subsequently leads to the estimate (3.4.16). □

Remark 3.4.2. As pointed out in Remark 3.4.1, in the case where the intervention operator satisfies Assumptions 7 and 8 with $C = 0$ (e.g. $\mathcal{M}_i$ is of the form (3.2.7)), we know the iterates $(u^n)_{n\in\mathbb{N}}$ are uniformly Lipschitz continuous and uniformly semiconcave with respect to n. Therefore, by following the above arguments, we can improve the estimates (3.4.15) and (3.4.16) to $\mathcal{O}((\log \rho)\rho^{-1/2})$ and $\mathcal{O}((\log \rho)\rho^{-1})$, respectively.

3.4.4 Approximation of action regions and optimal impulse controls

In this section, we propose convergent approximations to the action regions and optimal control strategies of the HJBQVI (3.2.1) based on the penalized equations. Since in general an optimal continuous control strategy may not exist due to the nonsmoothness of value functions, we shall focus on the approximation of optimal impulse control strategies.

Throughout this section, instead of specifying the precise convergence rates of the penalty schemes, which depend on the regularity of coefficients (see Theorem 3.4.10 and Remark 3.4.2), we shall assume there exists a function $\omega : (0, \infty) \to (0, \infty)$ such that $\omega(\rho) \to 0$ as $\rho \to \infty$ and

$$0 \leq u_i^\rho - u_i \leq \omega(\rho), \quad i \in \mathcal{I},\ \rho > 0. \tag{3.4.19}$$

For each $i \in \mathcal{I}$, we shall approximate the action region of the i-th component $\mathcal{S}_i = \{x \in \mathbb{R}^d \mid u_i(x) - \mathcal{M}_i u(x) = 0\}$ of (3.2.1) by the following sets:

$$\mathcal{S}_i^\rho = \{x \in \mathbb{R}^d \mid |u_i^\rho(x) - \mathcal{M}_i u^\rho(x)| \leq \omega(\rho)\}, \quad \rho > 0. \tag{3.4.20}$$

The next result shows that $\mathcal{S}_i^\rho$ converges to $\mathcal{S}_i$ in the Hausdorff metric.

Proposition 3.4.11. *Suppose Assumption 3, Assumption 4 and the error estimate (3.4.19) hold, and let* $(\mathcal{S}_i^\rho)_{i\in\mathcal{I}}$ *be the sets defined in (3.4.20) for each* $\rho > 0$*. Then* $\mathcal{S}_i \subset \mathcal{S}_i^\rho$ *for all* $i \in \mathcal{I}$ *and* $\rho > 0$*. Moreover, it holds for any given compact set* $\mathcal{K} \subset \mathbb{R}^d$ *that* $\mathcal{S}_i^\rho \cap \mathcal{K}$ *converges to* $\mathcal{S}_i \cap \mathcal{K}$ *in the Hausdorff metric as* $\rho \to \infty$*.*

Proof. The fact that $\mathcal{S}_i \subset \mathcal{S}_i^\rho$ follows directly from the estimate (3.4.19) and the monotonicity of $\mathcal{M}_i$ (see Lemma 3.2.2 (2)). Hence it remains to show that for any given compact set $\mathcal{K} \subset \mathbb{R}^d$, we have $\lim_{\rho\to\infty} \sup_{y\in\mathcal{S}_i^\rho\cap\mathcal{K}} \inf_{x\in\mathcal{S}_i\cap\mathcal{K}} |y-x| = 0$. Suppose it does not hold, then by passing to a subsequence, we know there exists $\varepsilon > 0$ and sequences $y_n \in \mathcal{S}_i^{\rho_n} \cap \mathcal{K}$, $\rho_n \to \infty$, such that $y_n \to y^* \in \mathcal{K}$ and $\inf_{x\in\mathcal{S}_i\cap\mathcal{K}} |y^*-x| \geq \varepsilon$, i.e., $y^* \notin \mathcal{S}_i$. However, by using the continuity of the functions u_i and $\mathcal{M}_i u$ (see Lemma 3.2.2 (3)), the definition of $\mathcal{S}_i^\rho$, and the estimate (3.4.19), we can obtain:

$$\begin{aligned}(u_i - \mathcal{M}_i u)(y^*) =& (u_i - \mathcal{M}_i u)(y^*) - (u_i - \mathcal{M}_i u)(y_n) \\ &+ (u_i - \mathcal{M}_i u)(y_n) - (u_i^{\rho_n} - \mathcal{M}_i u^{\rho_n})(y_n) + (u_i^{\rho_n} - \mathcal{M}_i u^{\rho_n})(y_n) \\ \geq& (u_i - \mathcal{M}_i u)(y^*) - (u_i - \mathcal{M}_i u)(y_n) - \omega(\rho_n) - \omega(\rho_n) \to 0\end{aligned}$$

as $n \to \infty$, which along with the fact $u_i \leq \mathcal{M}_i u$ on $\mathbb{R}^d$ implies $y^* \in \mathcal{S}_i$, and hence a contradiction. □

Remark 3.4.3. It is essential to include the modulus of convergence ω in the definition of $\mathcal{S}_i^\rho$, since in general the naive approximation $\tilde{\mathcal{S}}_i^\rho = \{x \in \mathbb{R}^d \mid u_i^\rho(x) - \mathcal{M}_i u^\rho(x) = 0\}$ does not give a convergent approximation to the action region $\mathcal{S}_i$. For example, let the vector $v = (v_l)_{l\in\{1,2\}}$ solve the following discrete QVI:

$$\max(v_1 - b,\, v_1 - (v_2 + c)) = 0, \quad \text{and} \quad \max(v_2 - 2b,\, v_2 - (v_1 + c)) = 0,$$

where $b > c > 0$. It is clear that the solution is given by $v_1 = b$ and $v_2 = b + c$, and the action region is the second index, i.e., $\mathcal{S} = \{2\}$. However, for each $\rho > 0$, one can directly verify that $v_1^\rho = b$ and $v_2^\rho = b + c + \frac{b-c}{1+\rho}$ solve the penalized equation:

$$v_1^\rho - b + \rho(v_1^\rho - v_2^\rho - c)^+ = 0, \quad \text{and} \quad v_2^\rho - 2b + \rho(v_2^\rho - v_1^\rho - c)^+ = 0,$$

which implies that $\tilde{\mathcal{S}}^\rho = \emptyset$ for all ρ.

Now we proceed to study optimal impulse control strategies. For any given $u \in [C^0(\mathbb{R}^d)]^M$, we denote by $\mathcal{Z}_i^u(x) := \arg\min_{z\in Z(x)}[u_i(\Gamma_i(x,z)) + K_i(x,z)]$ the set of optimal impulse control strategies for all $i \in \mathcal{I}$ and $x \in \mathcal{S}_i$. The following result constructs a convergent approximation of $\mathcal{Z}_i^u$, based on the set of impulse controls $\mathcal{Z}_i^{u^\rho}(x)$, $x \in \mathcal{S}_i^\rho$, obtained by the penalized solution u^ρ.

Theorem 3.4.12. *Suppose the assumptions of Proposition 3.4.11 hold. Then for any $i \in \mathcal{I}$, $x \in \mathcal{S}_i$, and sequence of impulse controls $(z^\rho)_{\rho>0}$ satisfying $z^\rho \in \mathcal{Z}_i^{u^\rho}(x)$ for all ρ, we have*

$$\lim_{\rho\to\infty} \bar{d}_Z(z^\rho, \mathcal{Z}_i^u(x)) \coloneqq \lim_{\rho\to\infty} \inf\{d_{\boldsymbol{Z}}(z^\rho, z) \mid z \in \mathcal{Z}_i^u(x)\} = 0,$$

where $(\boldsymbol{Z}, d_{\boldsymbol{Z}})$ is the metric space in Assumption 4. Consequently, if $\mathcal{Z}_i^u(x)$ is a singleton, then $\mathcal{Z}_i^{u^\rho}(x)$ converges to $\mathcal{Z}_i^u(x)$ in the Hausdorff metric as $\rho \to \infty$.

Proof. Suppose there exists $i \in \mathcal{I}$, $x \in \mathcal{S}_i$, and a sequence $(z^\rho)_{\rho>0}$ satisfying $z^\rho \in \mathcal{Z}_i^{u^\rho}(x)$ and $\bar{d}_Z(z^\rho, \mathcal{Z}_i^u(x)) \geq \varepsilon > 0$. Now let us consider the compact set $Z_\varepsilon(x) = \{z \in Z(x) \mid \bar{d}_Z(z, \mathcal{Z}_i^u(x)) \geq \varepsilon\}$, and pick $z_\varepsilon \in Z_\varepsilon(x)$ such that

$$u_i(\Gamma_i(x, z_\varepsilon)) + K_i(x, z_\varepsilon) = \min_{z\in Z_\varepsilon(x)} [u_i(\Gamma_i(x, z)) + K_i(x, z)].$$

Since $Z_\varepsilon(x) \cap \mathcal{Z}_i^u(x) = \emptyset$, we can deduce from the facts $z_\varepsilon \notin \mathcal{Z}_i^u(x)$ and $z^\rho \in Z_\varepsilon(x)$ that

$$\begin{aligned} &[u_i(\Gamma_i(x, z^\rho)) + K_i(x, z^\rho)] - [u_i(\Gamma_i(x, \hat{z})) + K_i(x, \hat{z})] \\ &\geq [u_i(\Gamma_i(x, z_\varepsilon)) + K_i(x, z_\varepsilon)] - [u_i(\Gamma_i(x, \hat{z})) + K_i(x, \hat{z})] \coloneqq c_0 > 0, \end{aligned} \tag{3.4.21}$$

for some $\hat{z} \in \mathcal{Z}_i^u(x)$. On the other hand, we obtain from the estimate (3.4.19) that

$$\begin{aligned} &[u_i(\Gamma_i(x, z^\rho)) + K_i(x, z^\rho)] - [u_i(\Gamma_i(x, \hat{z})) + K_i(x, \hat{z})] \\ &= [u_i(\Gamma_i(x, z^\rho)) - u_i^\rho(\Gamma_i(x, z^\rho))] + [u_i^\rho(\Gamma_i(x, z^\rho)) + K_i(x, z^\rho)] - [u_i(\Gamma_i(x, \hat{z})) + K_i(x, \hat{z})] \\ &\leq \mathcal{M}_i u^\rho(x) - \mathcal{M}_i u(x) \leq |u_i^\rho - u_i|_0 \leq \omega(\rho), \end{aligned}$$

which contradicts (3.4.21) by passing $\rho \to \infty$, and finishes the proof. □

Remark 3.4.4. We refer the reader to [45, 83] and references therein, where the uniqueness of a pointwise optimal impulse strategy has been established for various practical impulse control problems by exploiting the regularity of the value functions and the structure of the intervention operator. Then in Theorem 3.4.12 the convergence of the approximate controls follows.

3.5 Extension to some HJBQVIs with signed costs

In this section, we extend the penalty schemes to a class of QVIs with possibly negative impulse costs which arise from optimal switching problems. In this setting, the controller has two mechanisms of affecting the *regime switching process* I in (3.1.3), namely through their continuous control process α acting on its Markov transition matrix, as well as directly and immediately by exercising an impulse control to change the regime, the latter at the expense of a positive impulse cost or benefitting from a negative impulse cost. We shall propose an efficient, alternative penalty scheme by taking advantage of the finiteness of the set of switching controls, and extend the convergence analysis in Section 3.4 to estimate the penalization error.

More precisely, we consider the following system of HJBQVIs: for each $i \in \mathcal{I} = \{1, \dots, M\}$ and $x \in \mathbb{R}^d$,

$$F_i(x, u, Du_i, D^2 u_i) := \max \Big\{ \sup_{\alpha \in \mathcal{A}_i} \mathcal{L}_i^\alpha(x, u(x), Du_i(x), D^2 u_i(x));\ (u_i - \mathcal{M}_i u)(x) \Big\} = 0, \tag{3.5.1}$$

where the linear operator $\mathcal{L}_i^\alpha$ is defined as in (3.2.2), and the intervention operator $\mathcal{M}_i$ is given by

$$(\mathcal{M}_i u)(x) = \min_{j \in \mathcal{I}^{-i}} \{u_j(x) + k_{ij}(x)\}, \quad x \in \mathbb{R}^d. \tag{3.5.2}$$

By enlarging the state space $\mathbb{R}^d$ into the product space $\mathbb{R}^d \times \mathcal{I}$, we can treat (3.5.2) as a special case of (3.2.3) with $Z_i(x) = \mathcal{I}^{-i}$, $\Gamma_i(x, z) = (x, z)$ and $K_i(x, z) = k_{iz}(x)$ for all $(x, i) \in \mathbb{R}^d \times \mathcal{I}$. Consequently, if the switching costs k_{ij} are strictly positive, i.e., $k_{ij} \geq \kappa_0 > 0$, we can directly apply the penalty scheme (3.3.1) to solve (3.5.1), and deduce from Theorem 3.4.10 the rate of convergence in terms of the penalty parameter ρ.

As we shall see shortly, the structure of the operator $\mathcal{M}_i$ and the finiteness of the set of impulse controls allow us to consider signed switching costs $(k_{ij})_j$ taking both positive and negative values.

Now we introduce an alternative penalty scheme for solving (3.5.1). For any given penalty parameter $\rho \geq 0$, we consider the following system of penalized equations: for all $i \in \mathcal{I}$, $x \in \mathbb{R}^d$,

$$\begin{aligned} &F_i^\rho(x, u, Du_i, D^2 u_i) \\ &:= \sup_{\alpha \in \mathcal{A}_i} \mathcal{L}_i^\alpha(x, u^\rho(x), Du_i^\rho(x), D^2 u_i^\rho(x)) + \rho \sum_{j \in \mathcal{I}^{-i}} (u_i^\rho - u_j^\rho - k_{ij})^+(x) = 0. \end{aligned} \tag{3.5.3}$$

Unlike (3.3.1), the above penalty scheme makes use of the finiteness of the set $\mathcal{I}^{-i}$, and performs penalization on each component of the system, which leads to easily implementable and efficient iterative schemes for the penalized equations without taking the pointwise maximum over all switching components (see [88]).

Remark 3.5.1. The penalty scheme (3.5.3) can be extended to the general intervention operator (3.2.3), for which we introduce the following penalty term:

$$\rho \int_{Z_i(x)} \Big(u_i(x) - u_i(\Gamma_i(x,z)) - K_i(x,z) \Big)^+ \nu(dz), \quad \rho \geq 0,$$

where ν is a given finite measure supported on the set $\cup_{i,x} Z_i(x)$ (see [68]).

In the remaining part of this section, we shall discuss how to extend the convergence analysis in the previous sections to study penalty schemes for (3.5.1) with possibly negative switching costs. We shall focus on the scheme (3.5.3), but the same analysis extends naturally to the scheme (3.3.1). More precisely, we shall replace Assumption 4 by the following condition on the switching costs:

Assumption 9. *There exist constants $C \geq 0$ and $\kappa_0 > 0$ such that for all $j \neq i, l \in \mathcal{I}$, we have $k_{ii} \equiv 0$,*

$$k_{ij}(x) + k_{jl}(x) - k_{il}(x) \geq \kappa_0 > 0, \quad x \in \mathbb{R}^d, \tag{3.5.4}$$

and the following regularity estimates: $|k_{ij}|_1 \leq C$, and k_{ij} is semiconcave with constant C around any point $x \in \mathbb{R}^d$ with $k_{ij}(x) < \kappa_0$.

The allowance of negative switching costs clearly complicates the assumptions on the switching costs, which is worth a detailed discussion. The triangular condition (3.5.4) is similar to the assumption used in [83, 79], which means that it is less expensive to switch directly from regime i to l than in two steps via an intermediate regime j. It also implies $k_{ij} + k_{ji} \geq \kappa_0 > 0$ for all $j \neq i$, which prevents arbitrage opportunities that one can gain a positive profit by instantaneously switching back and forth. This further leads to the "no loop condition" introduced by [56], which together with Assumption 3 enables us to conclude a comparison principle of (3.5.1) by using similar arguments as those for [79, Theorem 2.1], and consequently the uniqueness of viscosity solutions to (3.5.1) in the class of bounded continuous functions.

The Lipschitz continuity and semiconcavity assumptions in Assumption 9 are similar to those in [79], which ensure the existence of a strict subsolution to (3.5.1) (see Proposition 3.5.1). However, we remark that, instead of requiring the switching costs to be semiconcave on $\mathbb{R}^d$ as in [79], we only impose the semiconcavity condition around the points at which the costs are close to or less than zero, hence no additional regularity is required if we are in the classical context of strictly positive switching costs.

The following proposition explicitly constructs a strict subsolution to (3.5.1), which is crucial to the well-posedness of (3.5.1) and (3.5.3), but also the error estimates of the penalty approximations (cf. Propositions 3.3.1 and 3.4.1).

Proposition 3.5.1. *Suppose Assumptions 3 and 9 hold. Then there exists a constant $C > 0$, such that for any $\varepsilon \in (0, \kappa_0)$, the function $w \in [C^{0,1}(\mathbb{R}^d)]^M$ defined as*

$$w_i = -\tilde{k}_i - C, \quad \tilde{k}_i = \min\left\{\min_{j\in\mathcal{I}^{-i}}(k_{ji} - \varepsilon), 0\right\}, \quad i \in \mathcal{I}, \tag{3.5.5}$$

is a strict subsolution to (3.5.1) and (3.5.3) for any $\rho \geq 0$, i.e., $F_i(x, u, Du_i, D^2u_i) \leq -\min(\varepsilon, \kappa_0 - \varepsilon)$ and $F_i^\rho(x, u, Du_i, D^2u_i) \leq -\min(\varepsilon, \kappa_0 - \varepsilon)$ in the viscosity sense.

Proof. For any given $\varepsilon \in (0, \kappa_0)$, we first verify $w_i - \mathcal{M}_i w \leq -\min(\varepsilon, \kappa_0 - \varepsilon)$. Note that

$$w_i - w_j - k_{ij} = -\tilde{k}_i + \tilde{k}_j - k_{ij} = -\tilde{k}_i + \min\left\{\min_{l\in\mathcal{I}^{-j}}(k_{lj} - \varepsilon), 0\right\} - k_{ij}, \quad \forall j \in \mathcal{I}^{-i}.$$

Now if $\tilde{k}_i = 0$, we can pick $l = i$ and deduce that $w_i - w_j - k_{ij} \leq k_{ij} - \varepsilon - k_{ij} = -\varepsilon$. Otherwise, if $\tilde{k}_i = k_{mi} - \varepsilon$ for some $m \in \mathcal{I}^{-i}$, then we shall separate the discussions into two cases. If $m = j$, then (3.5.4) and the facts that $k_{ii} = 0$, $\tilde{k}_j \leq 0$ imply that $w_i - w_j - k_{ij} \leq -(k_{ji} - \varepsilon) - k_{ij} \leq -(\kappa_0 - \varepsilon)$. On the other hand, if $m \neq j$, by setting $l = m$, we obtain that

$$w_i - w_j - k_{ij} \leq -(k_{mi} - \varepsilon) + (k_{mj} - \varepsilon) - k_{ij} \leq -\kappa_0,$$

which completes the proof of the desired statement by taking the maximum over all $j \in \mathcal{I}^{-i}$.

Note that for any given $i \in \mathcal{I}$ and $x \in \mathbb{R}^d$, $\tilde{k}_i(x)$ is defined by taking the minimum over the indices $j \in \mathcal{I}^{-i}$ such that $k_{ji}(x) \leq \varepsilon < \kappa_0$, hence by using Assumption 9 one can show $\tilde{k}_i(x)$ is semiconcave with some constant $C \geq 0$ around x. Therefore, we can infer for each $i \in \mathcal{I}$ that $\tilde{k}_i$ is Lipschitz continuous and semiconcave in $\mathbb{R}^d$. Hence there exists a sequence of smooth functions $(\tilde{k}^\varepsilon)_{\varepsilon>0}$ such that $D(-\tilde{k}^\varepsilon)$ is bounded and $D^2(-\tilde{k}^\varepsilon)$ is bounded below uniformly in terms of ε, and $\tilde{k}^\varepsilon$ uniformly converges to $\tilde{k}$ as $\varepsilon \to 0$. Then by using the boundedness of coefficients and the stability of subsolutions, we deduce that there exists a constant C' such that for all $i \in \mathcal{I}$ and $x \in \mathbb{R}^d$, we have $\sup_{\alpha\in\mathcal{A}_i} \mathcal{L}_i^\alpha(x, -\tilde{k}, D(-\tilde{k}_i), D^2(-\tilde{k}_i)) \leq C'$ in the viscosity sense. Hence, for any constant C such that $C \geq (C' + \min(\varepsilon, \kappa_0 - \varepsilon))/\lambda_0$, we can conclude that $w \in [C^{0,1}(\mathbb{R}^d)]^M$ is a strict subsolution to (3.5.1) and (3.5.3) for any $\rho \geq 0$. □

With the strict subsolution in hand, we can establish the existence of solutions to (3.5.3) (cf. Proposition 3.3.1), the monotone convergence of (3.5.3) (cf. Theorem 3.3.2), and also the error estimate of the iterated optimal stopping approximation of (3.5.1) (cf. Proposition 3.4.1).

Moreover, we can easily see that Assumptions 7 and 8 hold provided that the switching costs enjoy sufficient regularity. In fact, it is clear that if $u \in [C^{0,1}(\mathbb{R}^d)]^M$ and $[k_{ij}]_1 \leq C$ for all $j \in \mathcal{I}^{-i}$, then $\mathcal{M}_i u \in C^{0,1}(\mathbb{R}^d)$ satisfies $[\mathcal{M}_i u]_1 \leq \sup_{j\in\mathcal{I}^{-i}}([u_j]_1 + [k_{ij}]_1)$. If u_i and k_{ij} are semiconcave in $\mathbb{R}^d$ for all $i, j \in \mathcal{I}^{-i}$, then $\mathcal{M}_i u$ is semiconcave in $\mathbb{R}^d$ with constant $[\mathcal{M}_i u]_{2,+} \leq \sup_{j\in\mathcal{I}^{-i}}([u_j]_{2,+} + [k_{ij}]_{2,+})$. Therefore, we can obtain as a direct consequence of Theorem 3.4.4 that the iterates $(u^n)_{n\in\mathbb{N}}$ are Lipschitz continuous with constant $\mathcal{O}(n)$ if the switching costs are Lipschitz continuous, and they are semiconcave with constant $\mathcal{O}(n)$ if the switching costs are semiconcave.

Finally, by assuming the obstacles $(\Psi_i)_{i\in\mathcal{I}}$ in (3.4.3) are of the form $\Psi_i = \min_{j\in\mathcal{I}^{-i}} \Psi_{ij}$ for all $i \in \mathcal{I}$, we can generalize Proposition 3.4.8 to study the following penalty approximation to the classical obstacle problem (3.4.3):

$$\sup_{\alpha\in\mathcal{A}_i} \mathcal{L}_i^\alpha(x, v^\rho(x), Dv_i^\rho(x), D^2 v_i^\rho(x)) + \rho \sum_{j\in\mathcal{I}^{-i}} (v_i^\rho - \Psi_{ij})^+(x) = 0, \quad i \in \mathcal{I}.$$

and obtain exactly the same error estimates (3.4.13) and (3.4.14).

Now we are ready to conclude the following analogue of Theorem 3.4.10, which gives the convergence rate of (3.5.3) to (3.5.1) with respect to the penalty parameter.

Theorem 3.5.2. *Let u and u^ρ solve the QVI (3.5.1) and the penalized problem (3.5.3), respectively. If Assumptions 3), 5 and 9 hold, then for all large enough penalty parameter ρ, we have*

$$0 \leq u_i^\rho(x) - u_i(x) \leq C(\log \rho)^2 \rho^{-1/2}, \quad x \in \mathbb{R}^d,\, i \in \mathcal{I}.$$

If we further assume Assumption 6 holds, the constant λ_0 in Assumption 3 is sufficiently large, and $(k_{ij})_{i,j\in\mathcal{I}}$ are semiconcave in $\mathbb{R}^d$, then we have

$$0 \leq u_i^\rho(x) - u_i(x) \leq C(\log \rho)^2/\rho, \quad x \in \mathbb{R}^d,\, i \in \mathcal{I},$$

for some constant C, independent of the parameter ρ and the number of switching components M.

3.6 Discretization and policy iteration for penalized equations

In this section, we shall discuss briefly how to construct convergent discretizations for the penalized equations, and propose a globally convergent iterative method to solve the discretized equation based on policy iteration.

Let us start with the discretization of the penalized equation (3.3.1) with a fixed penalty parameter $\rho > 0$. We shall denote by $\{x_l\}_l = h\mathbb{Z}^d$ a uniform spatial grid on $\mathbb{R}^d$ with mesh size h, by $u^\rho_{i,l}$ the discrete approximation to u^ρ_i at the point x_l, and by $Z_{i,l}$ the set of impulse controls at the point x_l.

It is standard to show that, by using monotone discretizations (e.g., the standard finite difference schemes in [73] for the case with diagonally dominant diffusion coefficient a^α_i, or the semi-Lagrangian scheme in [32] for the general case) for the differential operators and multilinear interpolations for the intervention operator (see [5, 87]), one can derive the following approximation to (3.3.1): for all $i \in \mathcal{I}$,

$$\begin{aligned}\sup_{\alpha\in\mathcal{A}_i}\bigg[&\sum_{m\in\mathbb{Z}^d}\theta^\alpha_{i,l,m}(u^\rho_{i,l}-u^\rho_{i,m})+c^\alpha_{i,l}u^\rho_{i,l}-\sum_{j\in\mathcal{I}^{-i}}d^\alpha_{ij,l}u^\rho_{j,l}-\ell^\alpha_{i,l}\\&+\rho\bigg(u^\rho_{i,l}-\inf_{z\in Z_{i,l}}\bigg[\sum_{m\in\mathbb{Z}^d}\gamma^z_{i,l,m}u^\rho_{i,m}+K^z_{i,l}\bigg]\bigg)^+\bigg]=0,\quad l\in\mathbb{Z}^d,\end{aligned}\tag{3.6.1}$$

with some coefficients $\theta^\alpha_{i,l,m} \geq 0$, $0 \leq \gamma^z_{i,l,m} \leq 1$ and $\sum_{m\in\mathbb{Z}^d}\gamma^z_{i,l,m} = 1$ for all $l, m \in \mathbb{Z}^d$, $\alpha \in \mathcal{A}_i$ and $z \in Z_{i,l}$. Under Assumptions 3 and 4, it is straightforward to show that the above scheme is monotone and consistent with the penalized equation (3.3.1) as h tends to zero, which enables us to conclude from [23, Proposition 3.3] that the numerical solution of (3.6.1) converges to the solution of (3.3.1) as $h \to 0$. Moreover, one can deduce by similar arguments as those in [5] that the numerical solution converges to the solution of the QVI (3.2.1) when $1/\rho$ and h tend to zero simultaneously.

Now we proceed to demonstrate the global convergence of policy iteration for solving (3.6.1). We shall first enlarge the control space and reformulate (3.6.1) into an HJB equation in a countably infinite space. Note that by introducing the set $\mathcal{B} = \{0, 1\}$ and using the fact $\sum_{m\in\mathbb{Z}^d}\gamma^z_{i,l,m} = 1$, one can rearrange the terms of (3.6.1) and obtain that: for all $(i,l) \in \mathcal{I} \times \mathbb{Z}^d$,

$$\sup_{(\alpha,\beta,z)\in\mathcal{A}_i\times\mathcal{B}\times Z_{i,l}} \Bigg[\Bigg(\sum_{m\neq l}(\theta^{\alpha}_{i,l,m} + \beta\rho\gamma^{z}_{i,l,m}) + c^{\alpha}_{i,l}\Bigg)u^{\rho}_{i,l} \\ - \sum_{m\neq l}(\theta^{\alpha}_{i,l,m} + \beta\rho\gamma^{z}_{i,l,m})u^{\rho}_{i,m} - \sum_{j\in\mathcal{I}^{-i}} d^{\alpha}_{ij,l}u^{\rho}_{j,l} - \ell^{\alpha}_{i,l} - \beta\rho K^{z}_{i,l}\Bigg] = 0,$$

which can be equivalently expressed in the following compact form:

$$\sup_{\omega\in\mathcal{A}} \Big(\tilde{A}(\omega)\mathbf{u}^{\rho} - \tilde{b}(\omega)\Big) = 0, \tag{3.6.2}$$

where $\mathbf{u}^{\rho} = (u^{\rho}_{i,l})_{(i,l)\in\mathcal{I}\times\mathbb{Z}^d}$, $\mathcal{A} = (\mathcal{A}_i\times\mathcal{B}\times Z_{i,l})^{\mathcal{I}\times\mathbb{Z}^d}$, and $\tilde{A}(\omega) = (\tilde{a}_{(i,l),(i',l')}(\omega))_{(i,l),(i',l')\in\mathcal{I}\times\mathbb{Z}^d}$, $\omega \in \mathcal{A}$, is the following "infinite" matrix (see [19]): for any $\omega = (\alpha_{i,l}, \beta_{i,l}, z_{i,l})_{(i,l)\in\mathcal{I}\times\mathbb{Z}^d} \in \mathcal{A}$, we have

$$\tilde{a}_{(i,l),(i',l')}(\omega) = \begin{cases} \sum_{m\neq l}(\theta^{\alpha_{i,l}}_{i,l,m} + \beta_{i,l}\rho\gamma^{z_{i,l}}_{i,l,m}) + c^{\alpha_{i,l}}_{i,l}, & i' = i, l' = l, \\ -(\theta^{\alpha_{i,l}}_{i,l,l'} + \beta_{i,l}\rho\gamma^{z_{i,l}}_{i,l,l'}), & i' = i, l' \neq l, \\ -d^{\alpha_{i,l}}_{ii',l}, & i' \neq i, l = l. \end{cases} \tag{3.6.3}$$

Now we can apply the classical policy iteration to solve (3.6.2), or equivalently (3.6.1): let $\omega^{(0)}$ be a given initial control value, for all $k \geq 0$, define $(\mathbf{u}^{\rho,(k)}, \omega^{(k+1)})$ as follows:

$$\tilde{A}(\omega^{(k)})\mathbf{u}^{\rho,(k)} - \tilde{b}(\omega^{(k)}) = 0, \qquad \omega^{(k+1)} \in \arg\max_{\omega\in\mathcal{A}}(\tilde{A}(\omega)\mathbf{u}^{\rho,(k)} - \tilde{b}(\omega)), \tag{3.6.4}$$

where the maximization is performed component-wise. Such maximization operation is well-defined under Assumptions 3 and 4, due to the fact that the control set $\mathcal{A}_i \times \mathcal{B} \times Z_{i,l}$ is compact and the coefficients $\tilde{A}$ and $\tilde{b}$ are continuous in ω.

The next theorem establishes the monotone convergence of $(\mathbf{u}^{\rho,(k)})_{k\geq 0}$ for any initial guess $\omega^{(0)}$, which extends the result in [91, 19, 6] to weakly coupled systems in an infinite dimensional setting.

Theorem 3.6.1. *Suppose Assumptions 3 and 4 hold. Then for any initial control value $\omega^{(0)}$, the iterates $(\mathbf{u}^{\rho,(k)})_{k\geq 0}$ are well-defined, and converge pointwise to the unique solution of (3.6.2), or equivalently (3.6.1), as $k \to \infty$. Moreover, we have $\mathbf{u}^{\rho,(k)} \geq \mathbf{u}^{\rho,(k+1)}$ for all $k \geq 0$.*

Proof. The statement is an analogue of Proposition B.1 in [19], where the monotone convergence of policy iteration has been proved for concave HJB equations. Note (3.2.6) implies that for each $\omega \in \mathcal{A}$ and $(i,l) \in \mathcal{I} \times \mathbb{Z}^d$,

$$\tilde{a}_{(i,l),(i,l)}(\omega) \geq \sum_{(i',l')\neq(i,l))} |\tilde{a}_{(i,l),(i',l')}(\omega)| + \lambda_0,$$

which gives the monotonicity of $\tilde{A}$, i.e., for any given $\omega \in \mathcal{A}$, if $\tilde{A}(\omega)\mathbf{u} \geq 0$ and $\mathbf{u}$ is bounded, then $\mathbf{u} \geq 0$. Moreover, the boundedness of coefficients leads to the uniform boundedness of the iterates $(u^{(k)})_{k\geq 0}$ and the fact that $\sup_{\omega\in\mathcal{A}}(\text{Card}\{(i',l') \mid \tilde{a}_{(i,l),(i',l')}(\omega) \neq 0\}) < \infty$ for each $(i,l) \in \mathcal{I}\times\mathbb{Z}^d$. Therefore, even though the control set in (3.6.2) varies for each component (i,l), it is straightforward to adapt the arguments for [19, Proposition B.1] and establish the desired convergence result. □

Remark 3.6.1. Theorem 3.6.1 presents one of the major advantages of penalty schemes over the direct control scheme studied in [6, 27], which applies policy iteration to solve a direct discretization of QVI (3.2.1). Such a scheme in general is not well-defined due to the possible singularity of the matrix iterates caused by the non-strict monotonicity of $u_i - \mathcal{M}_i u$ in u. In fact, consider the simple QVI $\max(u-g, u-\mathcal{M}u) = 0$ with $\mathcal{M}u := u + c$ and $c > 0$, whose solution is given by $u = g$ due to the fact that $u - \mathcal{M}u = -c < 0$. Suppose that we initialize policy iteration with the impulse control, then we need to solve $u-(u+c) = 0$, which clearly admits no solution. More complicated examples can be constructed to show that the direct control scheme can fail at any intermediate iterate (see [6]).

3.7 Numerical experiments

In this section, we illustrate the theoretical findings and demonstrate the efficiency improvement of the penalty schemes over the direct control scheme through numerical experiments. We shall present an infinite-horizon optimal switching problem and examine the performance of penalty schemes with respect to the spatial mesh size and the penalty parameter.

We first introduce the following two-regime infinite-horizon optimal switching problem (see e.g. [83, 88]). Let $(\Omega, \mathcal{F}_t, \mathbb{P})$ be a filtered probability space and $\gamma = (\gamma_t)_{t\geq 0}$ be a control process such that $\gamma_t = \sum_{k\geq 0} i_k 1_{[\tau_k,\tau_{k+1})}(t)$, where $(\tau_k)_{k\geq 0}$ is a non-decreasing sequence of stopping times representing the decision on "when to switch", and for each $k \geq 0$, i_k is an $\mathcal{F}_{\tau_k}$-measurable random variable valued in the discrete space $\mathcal{I} = \{1, 2\}$, representing the decision on "where to switch". That is, the decision maker chooses regime i_k at the time τ_k for all $k \geq 0$.

For any given switching control strategy γ, we consider the following controlled state equation:

$$dX_t^\gamma = (r + \nu(\gamma_t)(\mu - r))X_t^\gamma dt + \sigma\nu(\gamma_t)X_t^\gamma\, dW_t, \quad t > 0; \quad X_0^\gamma = x,$$

where $r, \mu, \sigma, x > 0$ are given constants, $(W_t)_{t>0}$ is a one-dimensional Brownian motion defined on $(\Omega, \mathcal{F}_t, \mathbb{P})$, and $\nu(i) = i - 1$, $i \in \mathcal{I}$. Then the objective function associated with the control strategy γ is given by:

$$J(x, \gamma) = \mathbb{E}\bigg[\int_0^\infty e^{-rt} \ell(X_t^\gamma)\, dt - \sum_{k \geq 0} e^{-r\tau_{k+1}} c_{i_k, i_{k+1}} \bigg],$$

where ℓ represents the running reward function and $c_{i,j}$ represents the switching cost from regime i to j, $\forall i, j \in \mathcal{I}$. For each $i \in \mathcal{I}$, let $\mathbf{A}^i$ be all control strategies starting with regime i, i.e., $i_0 = i$ and $\tau_0 = 0$. Then the decision maker has the following value functions:

$$u_i(x) = \sup_{\gamma \in \mathbf{A}^i} J(x, \gamma), \quad i \in \mathcal{I} = \{1, 2\}.$$

Suppose that the switching costs $c_{i,j} \equiv c > 0$, $i \neq j$, then we can deduce from the dynamic programming principle (see [83]) that the value functions (u_1, u_2) satisfy the following system of quasi-variational inequalities: for all $i \in \mathcal{I}$, $j \neq i$, $x \in (0, \infty)$,

$$\min\bigg[-\frac{1}{2}\sigma^2 (i-1)^2 x^2 D^2 u_i(x) - (r + (i-1)(\mu - r)) x D u_i(x) + r u_i(x) - \ell(x), \\ (u_i - u_j + c)(x) \bigg] = 0. \quad (3.7.1)$$

Moreover, even though (3.7.1) involves a pointwise minimization instead of a pointwise maximization as in (3.5.1), for any given penalty parameter $\rho > 0$, one can easily extend the scheme (3.5.3) and derive the corresponding penalized equation for (3.7.1): for all $i \in \mathcal{I}$, $j \neq i$, $x \in (0, \infty)$,

$$-\frac{1}{2}\sigma^2 (i-1)^2 x^2 D^2 u_i^\rho(x) - (r + (i-1)(\mu - r)) x D u_i^\rho(x) + r u_i^\rho(x) - \ell(x) \\ - \rho(u_j^\rho - c - u_i^\rho)^+(x) = 0. \quad (3.7.2)$$

For our numerical experiments, we set the parameters as $c = 1/8$, $\sigma = 0.2$, $\mu = 0.06$, $r = 0.02$ and choose a nonsmooth running reward function: $\ell(x) = 0.5 - |x - 1|$ for $x \in [0.5, 1.5]$ and $\ell(x) = 0$ otherwise.

Now let $\rho > 0$, $n \in \mathbb{N}$, and $\{x_l\} = \{lh\}_{l \in \mathbb{N} \cup \{0\}}$ be a uniform grid of $(0, \infty)$ with the mesh size $h = 2^{-n}$. We shall derive a monotone discretization of the penalized equation (3.7.2) by employing the standard (two-point) forward difference for the first derivates and (three-point) central difference for all second derivatives; see Section 3.6 and [5] for the convergence of the discretization as $n, \rho \to \infty$. We shall also localize the equation on the computational domain $(0, 2)$ with homogenous Dirichlet boundary condition $u = 0$ at $x = 2$, which leads to the following discrete equation for (3.7.2): find $\mathbf{u}_N^\rho = (\mathbf{u}_{1,N}^\rho, \mathbf{u}_{2,N}^\rho) \in \mathbb{R}^N$ satisfying

$$\begin{aligned} &A\mathbf{u}_N^\rho - \vec{\ell} - \rho(b - M\mathbf{u}_N^\rho)^+ \\ &:= \begin{pmatrix} B_1 + rI_{N/2} & 0 \\ 0 & B_2 + rI_{N/2} \end{pmatrix} \mathbf{u}_N^\rho - \begin{pmatrix} \ell \\ \ell \end{pmatrix} - \rho \max(b - M\mathbf{u}_N^\rho, 0) = 0, \end{aligned} \tag{3.7.3}$$

where $N = 4/h = 2^{n+2}$ is the total number of unknowns, $B_1, B_2 \in \mathbb{R}^{N/2\times N/2}$ are matrices resulting from discretization of the differential operators, $\ell \in \mathbb{R}^{N/2}$ is a vector such that $\ell_k = \ell(x_{k-1})$ for all $k = 1, \dots, N/2$,

$$M = \begin{pmatrix} I_{N/2} & -I_{N/2} \\ -I_{N/2} & I_{N/2} \end{pmatrix}$$

is a matrix representation of the switching operator, and $b \in \mathbb{R}^N$ is a constant vector with value $-c$. Similarly, we discretize the following localized (3.7.1) by the same finite difference approximations: $u_1(x) = u_2(x) = 0$ at $x = 2$, and for all $x \in (0, 2)$,

$$\begin{aligned} \min\big(- rxDu_1(x) + ru_1(x) - \ell(x), u_1(x) - u_2(x) + c\big) &= 0, \\ \min\big(- \tfrac{1}{2}\sigma^2 x^2 D^2 u_2(x) - \mu x D u_2(x) + ru_2(x) - \ell(x), u_2(x) - u_1(x) + c\big) &= 0, \end{aligned}$$

which leads to the following direct control scheme: find $\mathbf{u}_N = (\mathbf{u}_{1,N}, \mathbf{u}_{2,N}) \in \mathbb{R}^N$ such that

$$\min(A\mathbf{u}_N - \vec{\ell}, M\mathbf{u}_N - b) = 0. \tag{3.7.4}$$

In the following, we shall discuss the implementation details for solving (3.7.3) and (3.7.4) with policy iteration. The direct control scheme, which will serve as a benchmark for our penalized schemes, applies policy iteration to the discrete equation (3.7.4) directly (see [27, 6]). More precisely, let $\omega^{(0)} \in \{0, 1\}^N$ be a given initial control value. Then, for all $k \geq 0$, we find $(\mathbf{u}^{(k)}, \omega^{(k+1)}) \in \mathbb{R}^N \times \{0, 1\}^N$ such that

$$A^{(k)}\mathbf{u}^{(k)} - b^{(k)} = 0, \qquad \omega^{(k+1)} \in \arg\min_{\omega\in\{0,1\}} \Big[(1-\omega)(A\mathbf{u}_N^{(k)} - \vec{\ell}) + \omega(M\mathbf{u}_N^{(k)} - b)\Big], \tag{3.7.5}$$

where the ith row of the matrix $A^{(k)}$ and the i-th component of the vector $b^{(k)}$ are determined by:

$$A_i^{(k)} = (1 - \omega_i^{(k)})A_i + \omega_i^{(k)} M_i, \quad b_i^{(k)} = (1 - \omega_i^{(k)})\vec{\ell}_i + \omega_i^{(k)} b, \quad i = 1, \dots, N.$$

The iteration will be terminated once a desired tolerance is achieved, i.e.,

$$\frac{\|\mathbf{u}_N^{(k)} - \mathbf{u}_N^{(k-1)}\|}{\max(\|\mathbf{u}_N^{(k)}\|, \text{scale})} < \text{tol}, \tag{3.7.6}$$

where $\|\cdot\|$ denotes the sup-norm, and the scale parameter is chosen to guarantee that no unrealistic level of accuracy will be imposed if the solution is close to zero. On the other hand, the penalized scheme views (3.7.3) with a given penalty parameter ρ as a discrete HJB equation, and applies policy iteration (3.6.4) to solve it, which will be terminated by the same criterion (3.7.6), with $(\mathbf{u}_N^{(k)})_{k\geq 0}$ replaced by $(\mathbf{u}_N^{\rho,(k)})_{k\geq 0}$. We take tol $= 10^{-9}$ and scale $= 1$ for all the experiments, and perform computations using MATLAB R2018a on a 2.70GHz Intel Xeon E5-2680 processor.

We reiterate that, compared with the global convergence of policy iteration (3.6.4) applied to the penalized equation (3.7.3), policy iteration (3.7.5) applied to (3.7.4) in general is not well-defined for an arbitrary initial guess $\omega^{(0)}$, as already observed in Remark 3.6.1 and [6]. In fact, if we initialize (3.7.5) with $\omega^{(0)} = \{1\}^N$, then we need to solve $M\mathbf{u}_N^{(0)} - b = 0$, which has no solution due to the structure of the matrix M and the fact $b = -c < 0$. Therefore, we shall initialize policy iteration for (3.7.3) and (3.7.4) with the continuation value, i.e., $\mathbf{u}_N^{(0)} = \mathbf{u}_N^{\rho,(0)}$ satisfying $A\mathbf{u}_N^{(0)} = \vec{\ell}$, which admits a solution since A is a monotone matrix.

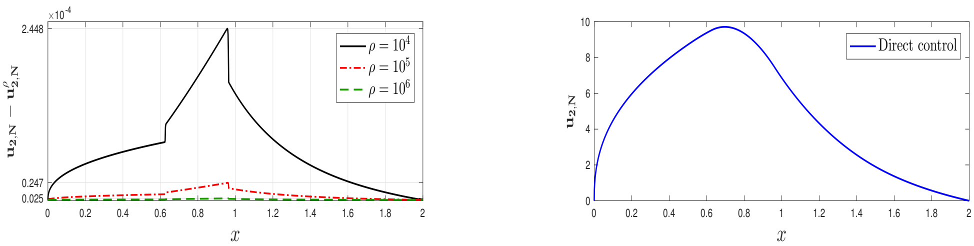


Figure 3.1: Numerical solutions of the value function u_2 obtained by the direct control scheme and the penalty schemes with different penalty parameters ($N = 65536$). Shown are: the difference $\mathbf{u}_{2,N} - \mathbf{u}^{\rho}_{2,N}$ of numerical solutions (left), and the numerical solution $\mathbf{u}_{2,N}$ of the direct control scheme (right).

We start by examining the convergence of the penalized schemes with respect to the penalty parameter and the mesh size. Figure 3.1 presents, for a fixed mesh size $h = 2^{-14}$ (the total number of unknowns is $N = 65536$), the difference between the numerical solutions obtained by the direct control scheme and the penalty scheme with different penalty parameters. It clearly indicates that, as the penalty parameter $\rho \to \infty$, the penalized solutions converge monotonically from below to the solution of the direct control scheme. Since the value functions are sufficiently smooth, we can also observe first order convergence of the penalization error (in the sup-norm) with respect to the penalty parameter ρ; see Figure 3.1 (right) for the graph of $\mathbf{u}_{2,N}$, while the graph of $\mathbf{u}_{1,N}$ is omitted due to a similar behaviour.

Table 3.1 summarizes, for different mesh sizes, the numerical solutions of the direct control scheme and the penalty scheme with a fixed parameter $\rho = 10^5$. It is interesting to observe that, for a fixed mesh size, the spatial discretization errors of both the direct control scheme and the penalty scheme are of the same magnitude and converge to zero with first order as the mesh size tends to 0. Moreover, the penalty parameter $\rho = 10^5$ already leads to a negligible penalization error (compared to the discretization error), which seems to be stable with respect to different mesh sizes.

Table 3.1: Results for the direct control scheme and the penalty scheme ($\rho = 10^5$) with different mesh sizes.

N	16384	32768	65536
Direct control scheme			
$\mathbf{u}_{1,N}(x=1)$	6.9339733	6.9330192	6.9325423
$\lvert\mathbf{u}_{1,N} - \mathbf{u}_{1,N/2}\rvert(x=1)$		9.54×10^{-4}	4.77×10^{-4}
Penalty scheme ($\rho = 10^5$)			
$\mathbf{u}^{\rho}_{1,N}(x=1)$	6.9339645	6.9330100	6.9325330
$\lvert\mathbf{u}^{\rho}_{1,N} - \mathbf{u}^{\rho}_{1,N/2}\rvert(x=1)$		9.54×10^{-4}	4.77×10^{-4}
$\lVert\mathbf{u}_N - \mathbf{u}^{\rho}_N\rVert$	2.42×10^{-5}	2.47×10^{-5}	2.47×10^{-5}

We proceed to analyze the computational efficiency of the direct control scheme and the penalty scheme. Figure 3.2 compares, for different mesh sizes and penalty parameters, the number of required policy iterations and the computational time of both schemes. One can observe clearly from Figure 3.2, left, that the number of required iterations for the direct control scheme (the blue line) exhibits a linear growth in the size of the discrete system. Moreover, our experiments show that policy iteration applied to (3.7.4) with fine meshes, i.e., $N \in \{131072, 262144\}$, is not able to meet the desired accuracy within 10^5 iterations, which suggests that the direct control scheme may diverge for sufficiently fine meshes. On the other hand, for penalty schemes with fixed penalty parameters (the green and black lines in Figure 3.2, left), the number of required iterations eventually stabilizes to a finite value for all fine meshes, which is significantly less than the number of iterations for the direct control scheme.

One can further compare the overall runtime of the direct control scheme and the penalty scheme for solving discrete systems with different sizes N (Figure 3.2, right). Note that for both methods, the computational time per iteration grows at a rate $O(N)$ due to the linear system solver. Hence, the total runtime of the direct control scheme increases at a rate $O(N^2)$ due to the linear growth of the required iterations (the blue line), while the penalized scheme (with a fixed penalty parameter) achieves a linear complexity in the computational time (the green and black lines), benefiting from a mesh-independence property of policy iteration for penalized equations. This suggests that the penalty schemes are significantly more efficient than the direct control scheme for solving large-scale discrete QVIs, as pointed out in [6].

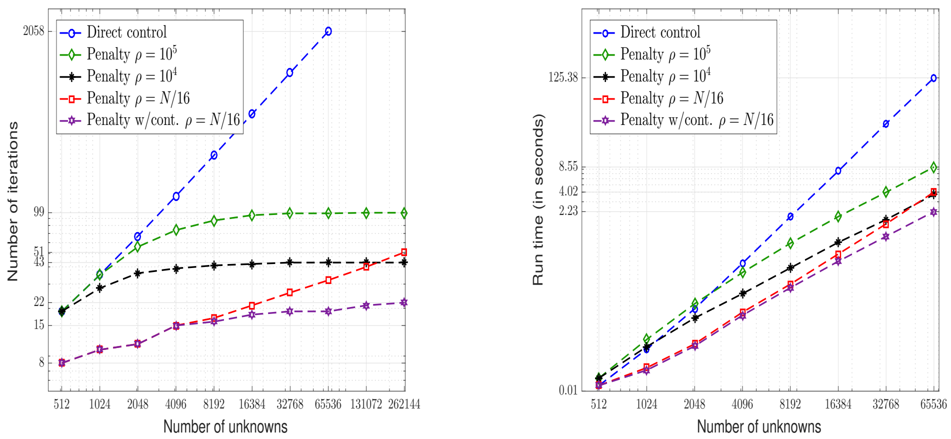


Figure 3.2: Comparison of the number of iterations and the runtime for the direct control scheme and the penalty method with different mesh sizes and penalty parameters (plotted in a log-log scale).

In practice, instead of solving the penalized equation (3.7.2) with a fixed penalty parameter ρ, we shall construct a convergent approximation to the solution of the QVI (3.7.1) based on the penalized solutions, by letting $1/\rho$ and the mesh size h tend to zero simultaneously (see also [6, 5]). The first order convergence of both the penalization error and the discretization error (see Figure 3.1 and Table 3.1) suggests us to take $\rho = CN$, where the constant $C = 1/16$ was found to achieve the optimal balance between the penalization error and the discretization error. Moreover, as suggested in [58], we can combine the penalty method with a continuation procedure in ρ to further improve the algorithm's efficiency. In particular, given a discrete penalized equation (3.7.3) of size N, if the corresponding penalty parameter $\rho = N/16 > 200$, we shall first solve a penalized equation (3.7.3) with the parameter $\rho = 100$ by using the initialization $\mathbf{u}^{(0)} = A^{-1}\ell$, and then use the solution as the initialization for the algorithm with the desired parameter ρ.

Figure 3.2 depicts the performance of the penalty scheme with the parameter $\rho = N/16$ (the red line) and the penalty scheme with the parameter $\rho = N/16$ and a continuation procedure (the purple line). The increasing penalty parameter results in an increasing number of iterations, but the growth rate is much lower than that of the discrete control scheme. A linear regression of the data without continuation procedure shows that the number of iterations is of the magnitude $O(N^{0.3})$. Moreover, the continuation strategy effectively enhances the efficiency of the algorithm, and the number of iterations has only a mild dependence on the size of the system.

We finally remark that one can choose $\Delta t = O(h)$ and $1/\rho = O(h)$ to construct a convergent penalty approximation to solutions of parabolic HJBQVIs. It has been observed in practice (see Table 6.6 in [6]) that the number of iterations for the penalty scheme remains stable with respect to the mesh refinement, due to the fact that refining the mesh size in general produces a more accurate initial guess for policy iteration, while the direct control scheme requires an increasing number of policy iterations per timestep as the mesh size tends to zero, which leads to significantly more policy iterations for high levels of refinement.

Chapter 4

Neural network based policy iteration for Hamilton-Jacobi-Bellman-Isaacs equations

4.1 Introduction

In this chapter, we propose a class of numerical schemes for solving Hamilton-Jacobi-Bellman-Isaacs (HJBI) boundary value problems of the following form:

$$-a^{ij}(x)\partial_{ij}u + G(x,u,\nabla u) = 0, \quad \text{in } \Omega \subset \mathbb{R}^n; \quad Bu = g, \quad \text{on } \partial\Omega, \tag{4.1.1}$$

where Ω is an open bounded domain, G is the (nonconvex) Hamiltonian defined as

$$G(x,u,\nabla u) = \max_{\alpha\in\mathbf{A}}\min_{\beta\in\mathbf{B}}\big(b^i(x,\alpha,\beta)\partial_i u(x) + c(x,\alpha,\beta)u(x) - f(x,\alpha,\beta)\big), \tag{4.1.2}$$

with given nonempty compact sets $\mathbf{A},\mathbf{B}$, and B is a boundary operator, i.e., if B is the identity operator, (4.1.1) is an HJBI Dirichlet problem, while if $Bu = \gamma^i\partial_i u + \gamma^0 u$ with some functions $\{\gamma^i\}_{i=0}^n$, (4.1.1) is an HJBI oblique derivative problem. Above and hereafter, when there is no ambiguity, we shall adopt the summation convention as in [43], i.e., repeated equal dummy indices indicate summation from 1 to n.

As we have seen in Chapter 3, such a nonconvex HJBI equation can arise naturally from a penalty approximation of hybrid control problems involving continuous controls, stopping times and impulse controls, where an HJB (quasi-)variational inequality can be reduced to an HJBI equation by penalizing the difference between the value function and the obstacles (see [58, 102] for penalty approximations of optimal stopping problems). Moreover, it is well-known that the value function of zero-sum stochastic differential games in domains satisfies the HJBI equation (4.1.1), and the optimal feedback controls can be constructed from the derivatives of the solutions (see e.g. [71] and references within; see also Section 4.6 for a concrete example). In particular, the HJBI Dirichlet problem corresponds to exit time problems of diffusion processes with controlled drift (see e.g. [71, 24, 80]), while the HJBI oblique derivative problem corresponds to state constraints (see e.g. [78, 73]). As (4.1.1) in general cannot be solved analytically, it is important to construct effective numerical schemes to find the solution of (4.1.1) and its derivatives.

The standard approach to solving (4.1.1) is to first discretize the operators in (4.1.1) by finite difference or finite element methods, and then solve the resulting nonlinear discretized equations by using policy iteration, also known as Howard's algorithm, or generally (finite-dimensional) semismooth Newton methods (see e.g. [40, 19, 96, 87]). However, this approach has the following drawbacks, as do most mesh-based methods: (1) it can be difficult to generate meshes and to construct consistent numerical schemes for problems in domains with complicated geometries; (2) the number of unknowns in general grows exponentially with the dimension n, i.e., it suffers from Bellman's *curse of dimensionality*, and hence this approach is infeasible for solving high-dimensional control problems. Moreover, since policy iteration is applied to a fixed finite-dimensional equation resulting from a particular discretization, it is difficult to infer whether the same convergence rate of policy iteration remains valid as the mesh size tends to zero ([91, 19]). We further remark that, for a given discrete HJBI equation, it can be difficult to determine a good initialization of policy iteration to ensure fast convergence of the algorithm; see [2] and references therein on possible accelerated methods.

Recently, numerical methods based on deep neural networks have been designed to solve high-dimensional partial differential equations (PDEs) (see e.g. [74, 36, 16, 37, 52, 95]). Most of these methods reformulate (4.1.1) into a nonlinear least-squares problem:

$$\inf_{u\in\mathcal{F}} \| - a^{ij}\partial_{ij}u + G(\cdot, u, \nabla u)\|^2_{L^2(\Omega)} + \|Bu - g\|^2_{L^2(\partial\Omega)}, \tag{4.1.3}$$

where $\mathcal{F}$ is a collection of neural networks with a smooth activation function. Based on collocation points chosen randomly from the domain, (4.1.3) is then reduced into an empirical risk minimization problem, which is subsequently solved by using stochastic optimization algorithms, in particular the Stochastic Gradient Descent (SGD) algorithm or its variants. Since these methods avoid mesh generation, they can be adapted to solve PDEs in high-dimensional domains with complex geometries. Moreover, the choice of smooth activation functions leads to smooth numerical solutions, whose values can be evaluated everywhere without interpolations. In the following, we shall refer to these methods as the Direct Method, due to the fact that there is no policy iteration involved.

We observe, however, that the Direct Method also has several serious drawbacks, especially for solving nonlinear nonsmooth equations including (4.1.1). Firstly, the nonconvexity of both the deep neural networks and the Hamiltonian G leads to a nonconvex empirical minimization problem, for which there is no theoretical guarantee on the convergence of SGD to a minimizer (see e.g. [94]). In practice, training a network with a desired accuracy could take hours or days (with hundreds of thousands of iterations) due to the slow convergence of SGD. Secondly, each SGD iteration requires the evaluation of ∇G (with respect to u and ∇u) on sample points, but ∇G is not necessarily defined everywhere due to the nonsmoothness of G. Moreover, evaluating the function G (again on a large set of sample points) can be expensive, especially when the sets $\mathbf{A}$ and $\mathbf{B}$ are of high dimensions, as we do not require more regularity than continuity of the coefficients with respect to the controls, so that approximate optimization may only be achieved by exhaustive search over a discrete coverage of the compact control set. Finally, as we shall see in Remark 4.4.2, merely including an $L^2(\partial\Omega)$-norm of the boundary data in the loss function (4.1.3) does not generally lead to convergence of the derivatives of numerical solutions or the corresponding feedback control laws.

In this work, we propose an efficient neural network based policy iteration algorithm for solving (4.1.1). At the $(k+1)$th iteration, $k \geq 0$, we shall update the control laws (α^k, β^k) by performing pointwise maximization/minimization of the Hamiltonian G based on the previous iterate u^k, and obtain the next iterate u^{k+1} by solving a linear boundary value problem, whose coefficients involve the control laws (α^k, β^k). This reduces the (nonconvex) semilinear problem into a sequence of *linear* boundary value problems, which are subsequently approximated by a multilayer neural network ansatz. Note that compared to Algorithm Ho-3 in [19] for discrete HJBI equations, which requires to solve a nonlinear HJB subproblem (involving minimization over the set $\mathbf{B}$) for each iteration, our algorithm only requires to solve a linear subproblem for each iteration, hence it is in general more efficient, especially when the dimension of $\mathbf{B}$ is high.

Policy iteration (or Successive Galerkin Approximation) was employed in [13, 12, 65, 67] to solve *convex HJB equations* on the whole space $\mathbb{R}^n$. Specifically, [13, 12, 65] approximate the solution to each linear equation via a separable polynomial ansatz (without concluding any convergence rate), while [67] assumes each linear equation is solved sufficiently accurately (without specifying a numerical method), and deduces pointwise *linear* convergence. The continuous policy iteration in [65] has also been applied to solve HJBI equations on $\mathbb{R}^n$ in [66], which is a direct extension of Algorithm Ho-3 in [19] and still requires to solve a nonlinear HJB subproblem at each iteration. In this chapter, we propose an easily implementable accuracy criterion for the numerical solutions of the *linear* PDEs which ensures the numerical solutions converge superlinearly in a suitable function space for nonconvex HJBI equations from an arbitrary initial guess.

Our algorithm enjoys the main advantage of the Direct Method, i.e., it is a mesh-free method and can be applied to solve high-dimensional stochastic games. Moreover, by utilizing the superlinear convergence of policy iteration, our algorithm effectively reduces the number of pointwise maximization/minimization over the sets $\mathbf{A}$ and $\mathbf{B}$, and significantly reduces the computational cost of the Direct Method, especially for high dimensional control sets. The superlinear convergence of policy iteration also helps eliminate the oscillation caused by SGD, which leads to smoother and more rapidly decaying loss curves in both the training and validation processes (see Figure 4.7). Our algorithm further allows training of the feedback controls on a separate network architecture from that representing the value function, or adaptively adjusting the architecture of networks for each policy iteration.

A major theoretical contribution of this work is the proof of global superlinear convergence of the policy iteration algorithm for the HJBI equation (4.1.1) in $H^2(\Omega)$, which is novel even for HJB equations (i.e., one of the sets $\mathbf{A}$ and $\mathbf{B}$ is singleton). Although the convergence of policy iteration for discrete equations has been proved in various works (see e.g. [84, 91, 40, 102, 96, 87] for a finite-dimensional setting, and [19, 89] for the setting of infinite-dimensional vector equations), to the best of our knowledge, there is no published work on the superlinear convergence of policy iteration for HJB PDEs in a function space, nor on the global convergence of policy iteration for solving nonconvex HJBI equations.

Moreover, this is the first work which demonstrates the convergence of neural network based methods for the solutions and their (first and second order) derivatives of nonlinear PDEs with merely measurable coefficients (cf. [46, 47, 52, 95]). We will also prove the pointwise convergence of the numerical solutions and their derivatives, which subsequently enables us to construct the optimal feedback controls from the numerical value functions and deduce convergence.

Let us briefly comment on the main difficulties encountered in studying the convergence of policy iteration for HJBI equations. Recall that at the $(k+1)$th iteration, we need to solve a linear boundary value problem, whose coefficients involve the control laws (α^k, β^k), obtained by performing pointwise maximization/minimization of the Hamiltonian G. The uncountability of the state space Ω and the nonconvexity of the Hamiltonian require us to exploit several technical measurable selection arguments to ensure the measurability of the controls (α^k, β^k), which is essential for the well-definedness of the linear boundary value problems and the algorithm.

Moreover, the nonconvexity of the Hamiltonian prevents us from following the arguments in [91, 40, 19] for discrete HJB equations to establish the *global* convergence of our inexact policy iteration algorithm for HJBI equations. In fact, a crucial step in the arguments for discrete HJB equations is to use the discrete maximum principle and show the iterates generated by policy iteration converge monotonically with an arbitrary initial guess, which subsequently implies the global convergence of the iterates. However, this monotone convergence is in general false for the iterates generated by the inexact policy iteration algorithm, due to the nonconvexity of the Hamiltonian and the fact that each linear equation is only solved approximately. We shall present a novel analysis technique for establishing the global convergence of our inexact policy iteration algorithm, by interpreting it as a fixed point iteration in $H^2(\Omega)$.

Finally, we remark that the proof of *superlinear* convergence of our algorithm is significantly different from the arguments for discrete equations. Instead of working with the sup-norm for (finite-dimensional) discrete equations as in [91, 40, 19, 102, 87], we employ a two-norm framework to establish the generalized differentiability of HJBI operators, where the norm gap is essential as has already been pointed out in [49, 99, 96]. Moreover, by taking advantage of the fact that the Hamiltonian only involves low order terms, we further demonstrate that the inverse of the generalized derivative is uniformly bounded. Furthermore, we include a suitable fractional Sobolev norm of the boundary data in the loss functions used in the training process, which is crucial for the $H^2(\Omega)$-superlinear convergence of the neural network based policy iteration algorithm.

We organize this chapter as follows. Section 4.2 states the main assumptions and recalls basic results for HJBI Dirichlet problems. In Section 4.3 we propose a policy iteration scheme for HJBI Dirichlet problems and establish its global superlinear convergence. Then in Section 4.4, we shall introduce the neural network based policy iteration algorithm, establish its various convergence properties, and construct convergent approximations to optimal feedback controls. We extend the algorithm and convergence results to HJBI oblique derivative problems in Section 4.5. Numerical examples for two-dimensional stochastic Zermelo navigation problems are presented in Section 4.6 to confirm the theoretical findings and to illustrate the effectiveness of our algorithms. The Appendix collects some basic results which are used in this chapter, and gives a proof for the main result on the HJBI oblique derivative problem.

4.2 HJBI Dirichlet problems

In this section, we introduce the HJBI Dirichlet boundary value problems of our interest, recall the appropriate notion of solutions, and state the main assumptions on its coefficients. We start with several important spaces used frequently throughout this work.

Let $n \in \mathbb{N}$ and Ω be a bounded $C^{1,1}$ domain in $\mathbb{R}^n$, i.e., a bounded open connected subset of $\mathbb{R}^n$ with a $C^{1,1}$ boundary. For each integer $k \geq 0$ and real p with $1 \leq p < \infty$, we denote by $W^{k,p}(\Omega)$ the standard Sobolev space of real functions with their weak derivatives of order up to k in the Lebesgue space $L^p(\Omega)$. When $p = 2$, we use $H^k(\Omega)$ to denote $W^{k,2}(\Omega)$. We further denote by $H^{1/2}(\partial\Omega)$ and $H^{3/2}(\partial\Omega)$ the spaces of traces from $H^1(\Omega)$ and $H^2(\Omega)$, respectively (see [44, Proposition 1.1.17]), which can be equivalently defined by using the surface measure σ on the boundaries $\partial\Omega$ as follows (see e.g. [42]):

$$\|g\|_{H^{1/2}(\partial\Omega)} = \Big[\int_{\partial\Omega} |g|^2 \, d\sigma + \iint_{\partial\Omega\times\partial\Omega} \tfrac{|g(x)-g(y)|^2}{|x-y|^n} \, d\sigma(x) d\sigma(y) \Big]^{\frac{1}{2}}, \tag{4.2.1}$$

$$\|g\|_{H^{3/2}(\partial\Omega)} = \Big[\int_{\partial\Omega} \big(|g|^2 + \textstyle\sum_{i=1}^n |\partial_i g|^2\big) \, d\sigma + \sum_{i=1}^n \iint_{\partial\Omega\times\partial\Omega} \tfrac{|\partial_i g(x)-\partial_i g(y)|^2}{|x-y|^n} \, d\sigma(x) d\sigma(y) \Big]^{\frac{1}{2}}.$$

We shall consider the following HJBI equation with nonhomogeneous Dirichlet boundary data:

$$F(u) := -a^{ij}(x)\partial_{ij} u + G(x, u, \nabla u) = 0, \quad \text{a.e. } \Omega, \tag{4.2.2a}$$

$$\tau u = g, \quad \text{on } \partial\Omega. \tag{4.2.2b}$$

where the nonlinear Hamiltonian is given as in (4.1.1):

$$G(x, u, \nabla u) = \max_{\alpha\in\mathbf{A}} \min_{\beta\in\mathbf{B}} \big(b^i(x, \alpha, \beta)\partial_i u(x) + c(x, \alpha, \beta)u(x) - f(x, \alpha, \beta)\big). \tag{4.2.3}$$

Throughout this chapter, we shall focus on the strong solution to (4.2.2), i.e., a twice weakly differentiable function $u \in H^2(\Omega)$ satisfying the HJBI equation (4.2.2a) almost everywhere in Ω, and the boundary values on $\partial\Omega$ will be interpreted as traces of the corresponding Sobolev space. For instance, $\tau u = g$ on $\partial\Omega$ in (4.2.2b) means that the trace of u is equal to g in $H^{3/2}(\partial\Omega)$, where $\tau \in \mathcal{L}(H^2(\Omega), H^{3/2}(\partial\Omega))$ denotes the trace operator (see [42, Proposition 1.1.17]). See Section 4.5 for boundary conditions involving the derivatives of solutions.

We now list the main assumptions on the coefficients of (4.2.2).

Assumption 10. *Let $n \in \mathbb{N}$, $\Omega \subset \mathbb{R}^n$ be a bounded $C^{1,1}$ domain, $\boldsymbol{A}$ be a nonempty finite set, and $\boldsymbol{B}$ be a nonempty compact metric space. Let $g \in H^{3/2}(\partial\Omega)$, $\{a^{ij}\}_{i,j=1}^n \subseteq C(\bar{\Omega})$ satisfy the following ellipticity condition with a constant $\lambda > 0$:*

$$\sum_{i,j=1}^n a^{ij}(x)\xi_i\xi_j \geq \lambda \sum_{i=1}^n \xi_i^2, \quad \text{for all } \xi \in \mathbb{R}^n \text{ and } x \in \Omega,$$

and $\{b^i\}_{i=1}^n, c, f \in L^\infty(\Omega \times \boldsymbol{A} \times \boldsymbol{B})$ satisfy that $c \geq 0$ on $\Omega \times \boldsymbol{A} \times \boldsymbol{B}$, and that $\phi(x, \alpha, \cdot) : \boldsymbol{B} \to \mathbb{R}$ is continuous, for all $\phi = b^i, c, f$ and $(x, \alpha) \in \Omega \times \boldsymbol{A}$.

As we shall see in Theorem 4.3.3 and Corollary 4.3.5, the finiteness of the set **A** enables us to establish the semismoothness of the HJBI operator (4.2.2a), whose coefficients involve a general nonlinear dependence on the parameters α and β. If all coefficients of (4.2.2a) are in a separable form, i.e., it holds for all $\phi = b^i, c, f$ that $\phi(x, \alpha, \beta) = \phi_1(x, \alpha) + \phi_2(x, \beta)$ for some functions ϕ_1, ϕ_2 (e.g. the penalized equation for variational inequalities with bilateral obstacles in [58]), then we can relax the finiteness of **A** to the same conditions on **B**.

Finally, in this work we focus on boundary value problems in a $C^{1,1}$ domain to simplify the presentation, but the numerical schemes and their convergence analysis can be extended to problems in nonsmooth convex domains with sufficiently regular coefficients (see e.g. [44, 96]).

We end this section by proving the uniqueness of solutions to the Dirichlet problem (4.2.2) in $H^2(\Omega)$. The existence of strong solutions shall be established constructively via policy iteration below (see Theorem 4.3.7).

Proposition 4.2.1. *Suppose Assumption 10 holds. Then the Dirichlet problem (4.2.2) admits at most one strong solution $u^* \in H^2(\Omega)$.*

Proof. Let $u, v \in H^2(\Omega)$ be two strong solutions to (4.2.2), we consider the following linear homogeneous Dirichlet problem:

$$-a^{ij}(x)\partial_{ij}w + \tilde{b}^i(x)\partial_i w + \tilde{c}(x)w = 0, \quad \text{a.e. in } \Omega; \quad \tau w = 0, \quad \text{on } \partial\Omega, \tag{4.2.4}$$

where we define the following measurable functions: for each $i = 1, \dots, n$,

$$\tilde{b}^i(x) = \begin{cases} \frac{G\left(x,v,((\partial_j v)_{1\le j<i},\partial_i u,(\partial_j u)_{i<j\le n})\right) - G\left(x,v,((\partial_j v)_{1\le j<i},\partial_i v,(\partial_j u)_{i<j\le n})\right)}{(\partial_i u - \partial_i v)(x)}, & \text{on } \Omega_1, \\ 0, & \text{otherwise,} \end{cases}$$

$$\tilde{c}(x) = \begin{cases} \frac{G(x,u,\nabla u) - G(x,v,\nabla u)}{(u-v)(x)}, & \text{on } \Omega_2, \\ 0, & \text{otherwise,} \end{cases}$$

with $\Omega_1 := \{x \in \Omega \mid \partial_i(u-v)(x) \neq 0\}$, $\Omega_2 := \{x \in \Omega \mid (u-v)(x) \neq 0\}$ and the Hamiltonian G defined as in (4.2.3). Note that the boundedness of coefficients implies that $\{\tilde{b}^i\}_{i=1}^n \subseteq L^\infty(\Omega)$, and $\tilde{c} \in L^\infty(\Omega)$. Moreover, one can directly verify that the following inequality holds for all parametrized functions $(f^{\alpha,\beta}, g^{\alpha,\beta})_{\alpha\in\mathbf{A},\beta\in\mathbf{B}}$: for all $x \in \mathbb{R}^n$,

$$\begin{aligned} \inf_{(\alpha,\beta)\in\mathbf{A}\times\mathbf{B}} f^{\alpha,\beta}(x) - g^{\alpha,\beta}(x) &\le \inf_{\alpha\in\mathbf{A}} \sup_{\beta\in\mathbf{B}} f^{\alpha,\beta}(x) - \inf_{\alpha\in\mathbf{A}} \sup_{\beta\in\mathbf{B}} g^{\alpha,\beta}(x) \\ &\le \sup_{(\alpha,\beta)\in\mathbf{A}\times\mathbf{B}} f^{\alpha,\beta}(x) - g^{\alpha,\beta}(x), \end{aligned}$$

which together with Assumption 10 leads to the estimate that $\tilde{c}(x) \ge \inf_{(\alpha,\beta)\in\mathbf{A}\times\mathbf{B}} c(x,\alpha,\beta) \ge 0$ on the set $\{x \in \Omega \mid (u-v)(x) \neq 0\}$, and hence we have $\tilde{c} \ge 0$ a.e. Ω. Then, we can deduce from Theorem B.1.1 that the Dirichlet problem (4.2.4) admits a unique strong solution $w^* \in H^2(\Omega)$ and $w^* = 0$. Since $w = u - v \in H^2(\Omega)$ satisfies (4.2.4) a.e. in Ω and $\tau w = 0$, we see that $w = u - v$ is a strong solution to (4.2.4) and hence $u - v = w^* = 0$, which subsequently implies the uniqueness of strong solutions to the Dirichlet problem (4.2.2). □

4.3 Policy iteration for HJBI Dirichlet problems

In this section, we propose a policy iteration algorithm for solving the Dirichlet problem (4.2.2). We shall also establish the global superlinear convergence of the algorithm, which subsequently gives a constructive proof for the existence of a strong solution to the Dirichlet problem (4.2.2).

We start by presenting the policy iteration scheme for the HJBI equations in Algorithm 1, which extends the policy iteration algorithm (or Howard's algorithm) for discrete HJB equations (see e.g. [40, 19, 87]) to the continuous setting.

The remaining part of this section is devoted to the convergence analysis of Algorithm 1. For notational simplicity, we first introduce two auxiliary functions: for

Algorithm 1 Policy iteration algorithm for Dirichlet problems.

1. Choose an initial guess u^0 in $H^2(\Omega)$, and set $k = 0$.
2. Given the iterate $u^k \in H^2(\Omega)$, update the following control laws: for all $\alpha \in \mathbf{A}, x \in \Omega$,
$$\alpha^k(x) \in \arg\max_{\alpha\in\mathbf{A}} \left[\min_{\beta\in\mathbf{B}} \big(b^i(x,\alpha,\beta)\partial_i u^k(x) + c(x,\alpha,\beta)u^k(x) - f(x,\alpha,\beta)\big)\right], \tag{4.3.1}$$
$$\beta^k(x) \in \arg\min_{\beta\in\mathbf{B}} \big(b^i(x,\alpha^k(x),\beta)\partial_i u^k(x) + c(x,\alpha^k(x),\beta)u^k(x) - f(x,\alpha^k(x),\beta)\big). \tag{4.3.2}$$
3. Solve the *linear* Dirichlet problem for $u^{k+1} \in H^2(\Omega)$:
$$-a^{ij}\partial_{ij}u + b^i_k \partial_i u + c_k u - f_k = 0, \quad \text{in } \Omega; \quad \tau u = g, \quad \text{on } \partial\Omega, \tag{4.3.3}$$
where $\phi_k(x) \coloneqq \phi(x, \alpha^k(x), \beta^k(x))$ for $\phi = b^i, c, f$.
4. If $\|u^{k+1} - u^k\|_{H^2(\Omega)} = 0$, then terminate with outputs u^{k+1}, α^k and β^k, otherwise increment k by one and go to step 2.

each $(x, \mathbf{u}, \alpha, \beta) \in \Omega \times \mathbb{R}^{n+1} \times \mathbf{A} \times \mathbf{B}$ with $\mathbf{u} = (z, p) \in \mathbb{R} \times \mathbb{R}^n$, we shall define the following functions

$$\ell(x, \mathbf{u}, \alpha, \beta) \coloneqq b^i(x,\alpha,\beta)p_i + c(x,\alpha,\beta)z - f(x,\alpha,\beta), \tag{4.3.4}$$
$$h(x, \mathbf{u}, \alpha) \coloneqq \min_{\beta\in\mathbf{B}} \ell(x, \mathbf{u}, \alpha, \beta). \tag{4.3.5}$$

Note that for all $k \geq 0$ and $x \in \Omega$, by setting $\mathbf{u}^k(x) = (u^k(x), \nabla u^k(x))$, we can see from (4.3.1) and (4.3.2) that

$$\ell(x, \mathbf{u}^k(x), \alpha^k(x), \beta^k(x)) = \min_{\beta\in\mathbf{B}} \ell(x, \mathbf{u}^k(x), \alpha^k(x), \beta) = \max_{\alpha\in\mathbf{A}} \min_{\beta\in\mathbf{B}} \ell(x, \mathbf{u}^k(x), \alpha, \beta). \tag{4.3.6}$$

We then recall several important concepts, which play a pivotal role in our subsequent analysis. The first concept ensures the existence of measurable feedback controls and the well-posedness of Algorithm 1.

Definition 4.3.1. Let (S, Σ) be a measurable space, and let X and Y be topological spaces. A function $\psi : S \times X \to Y$ is a Carathéodory function if:

1. for each $x \in X$, the function $\psi_x = \psi(\cdot, x) : S \to Y$ is $(\Sigma, \mathcal{B}_Y)$-measurable, where $\mathcal{B}_Y$ is the Borel σ-algebra of the topological space Y; and

2. for each $s \in S$, the function $\psi_s = \psi(s, \cdot) : X \to Y$ is continuous.

Remark 4.3.1. It is well-known that if X, Y are two complete separable metric spaces, and $\psi : S \times X \to Y$ is a Carathéodory function, then for any given measurable function $f : S \to X$, the composition function $s \to \psi(s, f(s))$ is measurable (see e.g. [3, Lemma 8.2.3]). Since any compact metric space is complete and separable, it is clear that Assumption 10 implies that the coefficients b^i, c, f are Carathéodory functions (with $S = \Omega$ and $X = \mathbf{A} \times \mathbf{B}$). Moreover, one can easily check that both ℓ and h are Carathéodory functions, i.e., ℓ (resp. h) is continuous in $(\mathbf{u}, \alpha, \beta)$ (resp. $(\mathbf{u}, \alpha)$) and measurable in x (see Theorem B.1.3 for the measurability of h in x).

We now recall a generalized differentiability concept for nonsmooth operators between Banach spaces, which is referred as semismoothness in [99] and slant differentiability in [28, 49]. It is well-known (see e.g. [19, 96, 87]) that the HJBI operator in (4.2.2a) is in general non-Fréchet-differentiable, and this generalized differentiability is essential for showing the superlinear convergence of policy iteration applied to HJBI equations.

Definition 4.3.2. Let $F : V \subset Y \mapsto Z$ be defined on an open subset V of the Banach space Y with images in the Banach space Z. In addition, let $\partial^* F : V \rightrightarrows \mathcal{L}(Y, Z)$ be a given set-valued mapping with nonempty images, i.e., $\partial^* F(y) \neq \emptyset$ for all $y \in V$. We say F is $\partial^* F$-semismooth in V if for any given $y \in V$, we have that F is continuous near y, and

$$\sup_{M \in \partial^* F(y+s)} \|F(y + s) - F(y) - Ms\|_Z = o(\|s\|_Y), \quad \text{as } \|s\|_Y \to 0.$$

The set-valued mapping $\partial^* F$ is called a generalized differential of F in V.

Remark 4.3.2. As in [99], we always require that $\partial^* F$ has a nonempty image, and hence the $\partial^* F$-semismooth of F in V shall automatically imply that the image of $\partial^* F$ is nonempty on V.

Now we are ready to analyze Algorithm 1. We first prove the semismoothness of the Hamiltonian G defined as in (4.2.3), by viewing it as the composition of a pointwise maximum operator and a family of HJB operators parameterized by the control α. Moreover, we shall simultaneously establish that, for each iteration, one can select measurable control laws α^k, β^k to ensure the measurability of the controlled coefficients b^i_k, c_k, f_k in the linear problem (4.3.3), which is essential for the well-posedness of strong solutions to (4.3.3), and the well-definedness of Algorithm 1.

The following proposition establishes the semismoothness of a parameterized family of first-order HJB operators, which extends the result for scalar-valued HJB operators in [96]. Moreover, by taking advantage of the fact that the operators involve only first-order derivatives, we are able to establish that they are semismooth from $H^2(\Omega)$ to $L^p(\Omega)$ for some $p > 2$ (cf. [96, Theorem 13]), which is essential for the superlinear convergence of Algorithm 1.

Proposition 4.3.1. *Suppose Assumption 10 holds. Let p be a given constant satisfying $p \geq 1$ if $n \leq 2$ and $p \in [1, \frac{2n}{n-2})$ if $n > 2$, and let $F_1 : H^2(\Omega) \to (L^p(\Omega))^{|\boldsymbol{A}|}$ be the HJB operator defined by*

$$F_1(u) := \left(\min_{\beta \in \boldsymbol{B}} \left(b^i(x, \alpha, \beta) \partial_i u + c(x, \alpha, \beta) u - f(x, \alpha, \beta) \right) \right)_{\alpha \in \boldsymbol{A}}, \quad \forall u \in H^2(\Omega).$$

Then F_1 is Lipschitz continuous and $\partial^* F_1$-semismooth in $H^2(\Omega)$ with a generalized differential

$$\partial^* F_1 : H^2(\Omega) \to \mathcal{L}(H^2(\Omega), (L^p(\Omega))^{|\boldsymbol{A}|})$$

defined as follows: for any $u \in H^2(\Omega)$, we have

$$\partial^* F_1(u) := \Big(b^i(\cdot, \alpha, \beta^u(\cdot, \alpha)) \partial_i + c(\cdot, \alpha, \beta^u(\cdot, \alpha)) \Big)_{\alpha \in \boldsymbol{A}}, \tag{4.3.7}$$

where $\beta^u : \Omega \times \boldsymbol{A} \to \boldsymbol{B}$ is any jointly measurable function such that for all $\alpha \in \boldsymbol{A}$ and $x \in \Omega$,

$$\beta^u(x, \alpha) \in \arg\min_{\beta \in \boldsymbol{B}} \left(b^i(x, \alpha, \beta) \partial_i u(x) + c(x, \alpha, \beta) u(x) - f(x, \alpha, \beta) \right). \tag{4.3.8}$$

Proof. Since $\mathbf{A}$ is a finite set, we shall assume without loss of generality that, the Banach space $(L^p(\Omega))^{|\mathbf{A}|}$ is endowed with the usual product norm $\|\cdot\|_{p,\mathbf{A}}$, i.e., for all $u \in (L^p(\Omega))^{|\mathbf{A}|}$, $\|u\|_{p,\mathbf{A}} = \sum_{\alpha \in \mathbf{A}} \|u(\cdot, \alpha)\|_{L^p(\Omega)}$. Note that the Sobolev embedding theorem shows that the following injections are continuous: $H^2(\Omega) \hookrightarrow W^{1,q}(\Omega)$, for all $q \geq 2, n \leq 2$, and $H^2(\Omega) \hookrightarrow W^{1,2n/(n-2)}(\Omega)$, for all $n > 2$. Thus for any given p satisfying the conditions in Proposition 4.3.1, we can find $r \in (p, \infty)$ such that the injection $H^2(\Omega) \hookrightarrow W^{1,r}(\Omega)$ is continuous. Then, the boundedness of b^i, c, f implies that the mappings F_1 and $\partial^* F_1$ are well-defined, and $F_1 : H^2(\Omega) \to (L^p(\Omega))^{|\mathbf{A}|}$ is Lipschitz continuous.

Now we show that the mapping $\partial^* F_1$ has a nonempty image from $W^{1,r}(\Omega)$ to $(L^p(\Omega))^{|\mathbf{A}|}$, where we choose $r \in (p, \infty)$ such that the injection $H^2(\Omega) \hookrightarrow W^{1,r}(\Omega)$ is continuous, and naturally extend the operators F_1 and $\partial^* F_1$ from $H^2(\Omega)$ to $W^{1,r}(\Omega)$.

For each $u \in W^{1,r}(\Omega)$, we consider the Carathéodory function $g : \Omega \times \mathbf{A} \times \mathbf{B} \to \mathbb{R}$ such that $g(x, \alpha, \beta) := \ell(x, (u, \nabla u)(x), \alpha, \beta)$ for all $(x, \alpha, \beta) \in \Omega \times \mathbf{A} \times \mathbf{B}$, where ℓ is defined by (4.3.5). Theorem B.1.3 shows there exists a function $\beta^u : \Omega \times \mathbf{A} \to \mathbf{B}$ satisfying (4.3.8), i.e.,

$$\beta^u(x, \alpha) \in \arg\min_{\beta \in \mathbf{B}} \ell\big(x, \big(u(x), \nabla u(x)\big), \alpha, \beta\big), \quad \forall (x, \alpha) \in \Omega \times \mathbf{A},$$

and β^u is jointly measurable with respect to the product σ-algebra on $\Omega \times \mathbf{A}$. Hence $\partial^* F_1(u)$ is nonempty for all $u \in W^{1,r}(\Omega)$.

We proceed to show that the operator F_1 is in fact $\partial^* F_1$-semismooth from $W^{1,r}(\Omega)$ to $(L^p(\Omega))^{|\mathbf{A}|}$, which implies the desired conclusion due to the continuous embedding $H^2(\Omega) \hookrightarrow W^{1,r}(\Omega)$. For each $\alpha \in \mathbf{A}$, we denote by $F_{1,\alpha} : W^{1,r}(\Omega) \to L^p(\Omega)$ the α-th component of F_1, and by $\partial^* F_{1,\alpha}$ the α-th component of $\partial^* F_1$. Theorem B.1.4 and the continuity of ℓ in $\mathbf{u}$ show that for each $(x, \alpha) \in \Omega \times \mathbf{A}$, the set-valued mapping

$$\mathbf{u} \in \mathbb{R}^{n+1} \rightrightarrows \arg\min_{\beta \in \mathbf{B}} \ell(x, \mathbf{u}, \alpha, \beta) \subseteq \mathbf{B},$$

is upper hemicontinuous, from which, by following precisely the steps in the arguments for [96, Theorem 13], we can prove that $F_{1,\alpha} : W^{1,r}(\Omega) \to L^p(\Omega)$ is $\partial^* F_{1,\alpha}$-semismooth. Then, by using the fact that a direct product of semismooth operators is again semismooth with respect to the direct product of the generalized differentials of the components (see [99, Proposition 3.6]), we can deduce that $F_1 : W^{1,r}(\Omega) \to (L^p(\Omega))^{|\mathbf{A}|}$ is semismooth with respect to the generalized differential $\partial^* F_1$ and finishes the proof. □

We then establish the semismoothness of a general pointwise maximum operator, by extending the result in [49] for the max-function $f : x \in \mathbb{R} \to \max(x, 0)$.

Proposition 4.3.2. *Let $p \in (2, \infty)$ be a given constant, $\boldsymbol{A}$ be a finite set, and Ω be a bounded subset of $\mathbb{R}^n$. Let $F_2 : (L^p(\Omega))^{|\boldsymbol{A}|} \to L^2(\Omega)$ be the pointwise maximum operator such that for each $u = (u(\cdot, \alpha))_{\alpha \in \boldsymbol{A}} \in (L^p(\Omega))^{|\boldsymbol{A}|}$,*

$$F_2(u)(x) := \max_{\alpha \in \boldsymbol{A}} u(x, \alpha), \quad \text{for a.e. } x \in \Omega. \tag{4.3.9}$$

Then F_2 is $\partial^* F_2$-semismooth in $(L^p(\Omega))^{|\boldsymbol{A}|}$ with a generalized differential

$$\partial^* F_2 : (L^p(\Omega))^{|\boldsymbol{A}|} \to \mathcal{L}((L^p(\Omega))^{|\boldsymbol{A}|}, L^2(\Omega))$$

defined as follows: for any $u = (u(\cdot, \alpha))_{\alpha \in \boldsymbol{A}}, v = (v(\cdot, \alpha))_{\alpha \in \boldsymbol{A}} \in (L^p(\Omega))^{|\boldsymbol{A}|}$, we have

$$\big(\partial^* F_2(u)v\big)(x) := v(x, \alpha^u(x)), \quad \text{for } x \in \Omega,$$

where $\alpha^u : \Omega \to \boldsymbol{A}$ is any measurable function such that

$$\alpha^u(x) \in \arg\max_{\alpha \in \boldsymbol{A}} (u(x,\alpha)), \quad \text{for } x \in \Omega. \tag{4.3.10}$$

Moreover, $\partial^* F_2(u)$ is uniformly bounded (in the operator norm) for all $u \in (L^p(\Omega))^{|\boldsymbol{A}|}$.

Proof. Let the Banach space $(L^p(\Omega))^{|\mathbf{A}|}$ be endowed with the product norm $\|\cdot\|_{p,\mathbf{A}}$ defined as in the proof of Proposition 4.3.1. We first show the mappings F_2 and $\partial^* F_2$ are well-defined, $\partial^* F_2$ has nonempty images, and $\partial^* F_2(u)$ is uniformly bounded for $u \in (L^p(\Omega))^{|\mathbf{A}|}$.

The finiteness of $\mathbf{A}$ implies that any $u \in (L^p(\Omega))^{|\mathbf{A}|}$ can also be viewed as a Carathéodory function $u : \Omega \times \mathbf{A} \to \mathbb{R}$. Hence for any given $u \in (L^p(\Omega))^{|\mathbf{A}|}$, we can deduce from Theorem B.1.3 the existence of a measurable function $\alpha^u : \Omega \to \mathbf{A}$ satisfying (4.3.10). Moreover, for any given measurable function $\alpha^u : \Omega \to \mathbf{A}$ and $v \in (L^p(\Omega))^{|\mathbf{A}|}$, the function $\partial^* F_2(u)v$ remains Lebesgue measurable (see Remark 4.3.1). Then, for any given $u \in (L^p(\Omega))^{|\mathbf{A}|}$ with $p > 2$, one can easily check that $F_2(u) \in L^2(\Omega)$, and $\partial^* F_2(u) \in \mathcal{L}((L^p(\Omega))^{|\mathbf{A}|}, L^2(\Omega))$, which subsequently implies that F_2 and $\partial^* F_2$ are well-defined, and the image of $\partial^* F_2$ is nonempty on $(L^p(\Omega))^{|\mathbf{A}|}$. Moreover, for any $u, v \in (L^p(\Omega))^{|\mathbf{A}|}$, Hölder's inequality leads to the following estimate:

$$\int_\Omega |v(x, \alpha^u(x))|^2 \, dx \le \int_\Omega \sum_{\alpha \in \mathbf{A}} |v(x,\alpha)|^2 \, dx \le \sum_{\alpha \in \mathbf{A}} |\Omega|^{(p-2)/p} \|v(\cdot,\alpha)\|^2_{L^p(\Omega)},$$

which shows that $\|\partial^* F_2(u)\|_{\mathcal{L}((L^p(\Omega))^{|\mathbf{A}|}, L^2(\Omega))} \le |\Omega|^{(p-2)/(2p)}$ for all $u \in (L^p(\Omega))^{|\mathbf{A}|}$.

Now we prove by contradiction that the operator F_2 is $\partial^* F_2$-semismooth. Suppose there exists a constant $\delta > 0$ and functions $u, \{v_k\}_{k=1}^\infty \in (L^p(\Omega))^{|\mathbf{A}|}$ such that $\|v_k\|_{p,\mathbf{A}} \to 0$ as $k \to \infty$, and

$$\|F_2(u + v_k) - F_2(u) - \partial^* F_2(u + v_k) v_k\|_{L^2(\Omega)} / \|v_k\|_{p,\mathbf{A}} \ge \delta > 0, \quad k \in \mathbb{N}, \tag{4.3.11}$$

where for each $k \in \mathbb{N}$, $\partial^* F_2(u + v_k)$ is defined with some measurable function $\alpha^{u+v_k} : \Omega \to \mathbf{A}$. Then, by passing to a subsequence, we may assume that for all $\alpha \in \mathbf{A}$, the sequence $\{v_k(\cdot,\alpha)\}_{k \in \mathbb{N}}$ converges to zero pointwise a.e. in Ω, as $k \to \infty$.

For notational simplicity, we define $\Sigma(x, u) := \arg\max_{\alpha \in \mathbf{A}} (u(x,\alpha))$ for all $u \in (L^p(\Omega))^{|\mathbf{A}|}$ and $x \in \Omega$. Then for a.e. $x \in \Omega$, we have $\lim_{k\to\infty} v_k(x,\alpha) = 0$ for all $\alpha \in \mathbf{A}$, $\alpha^{u+v_k}(x) \in \Sigma(x, u + v_k)$ for all $k \in \mathbb{N}$. By using the finiteness of $\mathbf{A}$ and the convergence of $\{v_k(\cdot,\alpha)\}_{k \in \mathbb{N}}$, it is straightforward to prove by contradiction that for all such $x \in \Omega$, $\alpha^{u+v_k}(x) \in \Sigma(x, u)$ for all large enough k.

We now derive an upper bound of the left-hand side of (4.3.11). For a.e. $x \in \Omega$, we have

$$
\begin{aligned}
&F_2(u+v_k)(x) - F_2(u)(x) - \big(\partial^* F_2(u+v_k)v_k\big)(x) \\
&\qquad \le (u+v_k)(x, \alpha^{u+v_k}(x)) - u(x, \alpha^{u+v_k}(x)) - v_k(x, \alpha^{u+v_k}(x)) = 0, \\
&F_2(u+v_k)(x) - F_2(u)(x) - \big(\partial^* F_2(u+v_k)v_k\big)(x) \\
&\qquad \ge (u+v_k)(x, \alpha^{u}(x)) - u(x, \alpha^{u}(x)) - v_k(x, \alpha^{u+v_k}(x)) \\
&\qquad = v_k(x, \alpha^{u}(x)) - v_k(x, \alpha^{u+v_k}(x)),
\end{aligned}
$$

from any $\alpha^u(x) \in \Sigma(x,u)$. Thus, for each $k \in \mathbb{N}$, we have for a.e. $x \in \Omega$ that,

$$
\begin{aligned}
&|F_2(u+v_k)(x) - F_2(u)(x) - \big(\partial^* F_2(u+v_k)v_k\big)(x)| \\
&\le \phi_k(x) \coloneqq \inf_{\alpha^u \in \Sigma(x,u)} |v_k(x, \alpha^u) - v_k(x, \alpha^{u+v_k}(x))|,
\end{aligned}
$$

where, by applying Theorem B.1.3 twice, we can see that both the set-valued mapping $x \rightrightarrows \Sigma(x,u)$ and the function ϕ_k are measurable.

We then introduce the set $\Omega_k = \{x \in \Omega \mid \alpha^{u+v_k}(x) \notin \Sigma(x,u)\}$ for each $k \in \mathbb{N}$. The measurability of the set-valued mapping $x \rightrightarrows \Sigma(x,u)$ implies the associated distance function $\rho(x,\alpha) \coloneqq \operatorname{dist}(\alpha, \Sigma(x,u))$ is a Carathéodory function (see [1, Theorem 18.5]), which subsequently leads to the measurability of Ω_k for all k. Hence we can deduce that

$$
\begin{aligned}
&\|F_2(u+v_k) - F_2(u) - \partial^* F_2(u+v_k)v_k\|^2_{L^2(\Omega)} \\
&\le \int_{\Omega_k} \inf_{\alpha^u \in \Sigma(x,u)} |v_k(x,\alpha^u) - v_k(x, \alpha^{u+v_k}(x))|^2\, dx \\
&\le 2 \int_{\Omega_k} \sum_{\alpha \in \mathbf{A}} |v_k(x,\alpha)|^2\, dx \le 2 \sum_{\alpha \in \mathbf{A}} |\Omega_k|^{(p-2)/p} \|v_k(\cdot,\alpha)\|^2_{L^p(\Omega)},
\end{aligned}
$$

which leads to the following estimate:

$$
\|F_2(u+v_k) - F_2(u) - \partial^* F_2(u+v_k)v_k\|_{L^2(\Omega)} / \|v_k\|_{p,\mathbf{A}} \le \sqrt{2} |\Omega_k|^{(p-2)/(2p)} \to 0, \quad \text{as } k \to \infty,
$$

where we have used the bounded convergence theorem and the fact that for a.e. $x \in \Omega$, $1_{\Omega_k}(x) = 0$ for all large enough k. This contradicts to the hypothesis (4.3.11), and hence finishes our proof. □

Now we are ready to conclude the semismoothness of the HJBI operator. Note that the argument in [96] does not apply directly to the HJBI operator, due to the nonconvexity of the Hamiltonian G defined as in (4.2.3).

Theorem 4.3.3. *Suppose Assumption 10 holds, and let $F : H^2(\Omega) \to L^2(\Omega)$ be the HJBI operator defined as in (4.2.2a). Then F is semismooth in $H^2(\Omega)$, with a generalized differential $\partial^* F : H^2(\Omega) \to \mathcal{L}(H^2(\Omega), L^2(\Omega))$ defined as follows: for any $u \in H^2(\Omega)$,*

$$\partial^* F(u) := -a^{ij}(\cdot)\partial_{ij} + b^i(\cdot, \alpha(\cdot), \beta^u(\cdot, \alpha(\cdot)))\partial_i + c(\cdot, \alpha(\cdot), \beta^u(\cdot, \alpha(\cdot))), \tag{4.3.12}$$

where $\beta^u : \Omega \times \boldsymbol{A} \to \boldsymbol{B}$ is any jointly measurable function satisfying (4.3.8), and $\alpha : \Omega \to \boldsymbol{A}$ is any measurable function such that for a.e. $x \in \Omega$,

$$\alpha(x) \in \arg\max_{\alpha \in \mathbf{A}} \left[\min_{\beta \in \boldsymbol{B}} \left(b^i(x, \alpha, \beta)\partial_i u(x) + c(x, \alpha, \beta)u(x) - f(x, \alpha, \beta) \right) \right], \tag{4.3.13}$$

Proof. Note that we can decompose the HJBI operator $F : H^2(\Omega) \to L^2(\Omega)$ into $F = F_0 + F_2 \circ F_1$, where $F_0 : H^2(\Omega) \to L^2(\Omega)$ is the linear operator $u \mapsto -a^{ij}\partial_{ij}u$, $F_1 : H^2(\Omega) \to (L^p(\Omega))^{|\mathbf{A}|}$ is the HJB operator defined in Proposition 4.3.1, $F_2 : (L^p(\Omega))^{|\mathbf{A}|} \to L^2(\Omega)$ is the pointwise maximum operator defined in Proposition 4.3.2, and p is a constant satisfying $p > 2$ if $n \leq 2$, and $p \in (2, 2n/(n-2))$ if $n > 2$.

Proposition 4.3.1 shows that F_1 is Lipschitz continuous and semismooth with respect to the generalized differential $\partial^* F_1$ defined by (4.3.7), while Proposition 4.3.2 shows that F_2 is semismooth with respect to the uniformly bounded generalized differential $\partial^* F_2$ defined by (4.3.9). Hence, we know the composed operator $F_2 \circ F_1$ is semismooth with respect to the composition of the generalized differentials (see [99, Proposition 3.8]), i.e., $\partial^*(F_2 \circ F_1)(u) = \partial^* F_2(F_1(u)) \circ \partial^* F_1(u)$ for all $u \in H^2(\Omega)$. Consequently, by using the fact that F_0 is Fréchet differentiable with the derivative $-a^{ij}\partial_{ij} \in \mathcal{L}(H^2(\Omega), L^2(\Omega))$, we can conclude from Propositions 4.3.1 and 4.3.2 that $F : H^2(\Omega) \to L^2(\Omega)$ is semismooth on $H^2(\Omega)$, and that (4.3.12) is a desired generalized differential of F at u. □

Note that the above characterization of the generalized differential of the HJBI operator involves a jointly measurable function $\beta^u : \Omega \times \mathbf{A} \to \mathbf{B}$, satisfying (4.3.8) for all $(x, \alpha) \in \Omega \times \mathbf{A}$. We now present a technical lemma, which allows us to view the control law β^k in (4.3.2) as such a feedback control on $x \in \Omega$ and $\alpha \in \mathbf{A}$.

Lemma 4.3.4. *Suppose Assumption 10 holds. Let $h, \{h_i\}_{i=1}^n : \Omega \to \mathbb{R}$, $\alpha^h : \Omega \to \boldsymbol{A}$ be given measurable functions, and $\beta^h : \Omega \to \boldsymbol{B}$ be a measurable function such that for all $x \in \Omega$,*

$$\beta^h(x) \in \arg\min_{\beta \in \boldsymbol{B}} \left(b^i(x, \alpha^h(x), \beta)h_i(x) + c(x, \alpha^h(x), \beta)h(x) - f(x, \alpha^h(x), \beta) \right). \tag{4.3.14}$$

Then there exists a jointly measurable function $\tilde{\beta}^h : \Omega \times \boldsymbol{A} \to \boldsymbol{B}$ *such that* $\beta^h(x) = \tilde{\beta}^h(x, \alpha^h(x))$ *for all* $x \in \Omega$, *and it holds for all* $x \in \Omega$ *and* $\alpha \in \boldsymbol{A}$ *that*

$$\tilde{\beta}^h(x, \alpha) \in \arg\min_{\beta \in \boldsymbol{B}} \left(b^i(x, \alpha, \beta)h_i(x) + c(x, \alpha, \beta)h(x) - f(x, \alpha, \beta)\right). \tag{4.3.15}$$

Proof. Let $\beta^h : \Omega \to \mathbf{B}$ be a given measurable function satisfying (4.3.14) for all $x \in \Omega$ (see Remark 4.3.1 and Theorem B.1.3 for the existence of such a measurable function). As shown in the proof of Proposition 4.3.1, there exists a jointly measurable function $\bar{\beta} : \Omega \times \mathbf{A} \to \mathbf{B}$ satisfying the property (4.3.15) for all $(x, \alpha) \in \Omega \times \mathbf{A}$. Now suppose that $\mathbf{A} = \{\alpha_i\}_{i=1}^{|\mathbf{A}|}$ with $|\mathbf{A}| < \infty$ (see Assumption 10), we shall define the function $\tilde{\beta}^h(x, \alpha) : \Omega \times \mathbf{A} \to \mathbf{B}$, such that for all $(x, \alpha) \in \Omega \times \mathbf{A}$,

$$\tilde{\beta}^h(x, \alpha) = \begin{cases} \beta^h(x), & (x, \alpha) \in \mathcal{C} := \bigcup_{i=1}^{|\mathbf{A}|} \left(\{x \in \Omega \mid \alpha^h(x) = \alpha_i\} \times \{\alpha_i\}\right), \\ \bar{\beta}(x, \alpha), & \text{otherwise.} \end{cases}$$

The measurability of α^h and the finiteness of $\mathbf{A}$ imply that the set $\mathcal{C}$ is measurable in the product σ-algebra on $\Omega \times \mathbf{A}$, which along with the joint measurability of $\bar{\beta}$ leads to the joint measurability of the function $\tilde{\beta}^h$.

For any given $x \in \Omega$, we have $(x, \alpha^h(x)) \in \{y \in \Omega \mid \alpha^h(y) = \alpha^h(x)\} \times \{\alpha^h(x)\}$, from which we can deduce from the definition of $\tilde{\beta}^h$ that $\tilde{\beta}^h(x, \alpha^h(x)) = \beta^h(x)$ for all $x \in \Omega$. Finally, for any given $\alpha_i \in \mathbf{A}$, we shall verify (4.3.15) for all $x \in \Omega$ and $\alpha = \alpha_i$. Let $x \in \Omega$ be fixed. If $\alpha^h(x) = \alpha_i$, then the fact that $(x, \alpha_i) \in \mathcal{C}$ and the definition of $\tilde{\beta}^h$ imply that $\tilde{\beta}^h(x, \alpha_i) = \beta^h(x)$, which along with (4.3.14) and $\alpha^h(x) = \alpha_i$ shows that (4.3.15) holds for the point (x, α_i). On the other hand, if $\alpha^h(x) \neq \alpha_i$, then $(x, \alpha_i) \notin \mathcal{C}$ and $\tilde{\beta}^h(x, \alpha_i) = \bar{\beta}(x, \alpha_i)$ satisfies the condition (4.3.15) due to the selection of $\bar{\beta}$. □

As a direct consequence of the above extension result, we now present an equivalent characterization of the generalized differential of the HJBI operator.

Corollary 4.3.5. *Suppose Assumption 10 holds, and let* $F : H^2(\Omega) \to L^2(\Omega)$ *be the HJBI operator defined as in (4.2.2a). Then* F *is semismooth in* $H^2(\Omega)$, *with a generalized differential* $\partial^* F : H^2(\Omega) \to \mathcal{L}(H^2(\Omega), L^2(\Omega))$ *defined as follows: for any* $u \in H^2(\Omega)$,

$$\partial^* F(u) := -a^{ij}(\cdot)\partial_{ij} + b^i(\cdot, \alpha^u(\cdot), \beta^u(\cdot))\partial_i + c(\cdot, \alpha^u(\cdot), \beta^u(\cdot)), \tag{4.3.16}$$

where $\alpha^u : \Omega \to \boldsymbol{A}$ *and* $\beta^u : \Omega \to \boldsymbol{B}$ *are any measurable functions satisfying for all*

$x \in \Omega$ *that*

$$\begin{aligned}\alpha^u(x) &\in \arg\max_{\alpha\in\mathbf{A}} \left[\min_{\beta\in\boldsymbol{B}} \left(b^i(x,\alpha,\beta)\partial_i u(x) + c(x,\alpha,\beta)u(x) - f(x,\alpha,\beta)\right)\right],\\ \beta^u(x) &\in \arg\min_{\beta\in\boldsymbol{B}} \left(b^i(x,\alpha^u(x),\beta)\partial_i u(x) + c(x,\alpha^u(x),\beta)u(x) - f(x,\alpha^u(x),\beta)\right).\end{aligned} \tag{4.3.17}$$

Proof. Let $u \in H^2(\Omega)$, and let α^u and β^u be given measurable functions satisfying (4.3.17) (see Remark 4.3.1 and Theorem B.1.3 for the existence of such measurable functions). Then by using Lemma 4.3.4, we know there exists a jointly measurable function $\tilde{\beta}^u : \Omega \times \mathbf{A} \to \mathbf{B}$ such that $\tilde{\beta}^u$ satisfies (4.3.8) for all $(x,\alpha) \in \Omega \times \mathbf{A}$, and $\tilde{\beta}^u(x,\alpha^u(x)) = \beta^u(x)$ for all $x \in \Omega$. Hence we see the linear operator defined in (4.3.16) is equal to the following operator

$$-a^{ij}(\cdot)\partial_{ij} + b^i(\cdot,\alpha^u(\cdot),\tilde{\beta}^u(\cdot,\alpha^u(\cdot)))\partial_i + c(\cdot,\alpha^u(\cdot),\tilde{\beta}^u(\cdot,\alpha^u(\cdot))) \in \mathcal{L}(H^2(\Omega),L^2(\Omega)),$$

which is a generalized differential of the HJBI operator F at u due to Theorem 4.3.3. □

The above characterization of the generalized differential of the HJBI operator enables us to demonstrate the superlinear convergence of Algorithm 1 by reformulating it as a semismooth Newton method for an operator equation. Here we allow **A** and **B** to be finite sets, and focus on the superlinear convergence of Algorithm 1. If **A** and **B** are convex and the coefficients of (4.2.2) are sufficiently regular with respect to the control variables, then one can establish further regularity of the HJBI operator and a quadratic convergence of Algorithm 1 (see e.g. [99, Theorem 3.13]).

Theorem 4.3.6. *Suppose Assumption 10 holds and let $u^* \in H^2(\Omega)$ be a strong solution to the Dirichlet problem (4.2.2). Then there exists a neighborhood $\mathcal{N}$ of u^*, such that for all $u^0 \in \mathcal{N}$, Algorithm 1 either terminates with $u^k = u^*$ for some $k \in \mathbb{N}$, or generates a sequence $\{u^k\}_{k\in\mathbb{N}}$ that converges q-superlinearly to u^* in $H^2(\Omega)$, i.e., $\lim_{k\to\infty} \|u^{k+1}-u^*\|_{H^2(\Omega)}/\|u^k-u^*\|_{H^2(\Omega)} = 0$.*

Proof. Note that the Dirichlet problem (4.2.2) can be written as an operator equation $\tilde{F}(u) = 0$ with the following operator

$$\tilde{F} : u \in H^2(\Omega) \to (F(u), \tau u - g) \in L^2(\Omega) \times H^{3/2}(\partial\Omega),$$

where F is the HJBI operator defined as in (4.2.2a), and $\tau : H^2(\Omega) \to H^{3/2}(\partial\Omega)$ is the trace operator. Moreover, one can directly check that given an iterate $u^k \in H^2(\Omega)$, $k \geq 0$, the next iterate u^{k+1} solves the following Dirichlet problem:

$$L_k(u^{k+1}-u^k) = -F(u^k), \quad \text{in } \Omega; \quad \tau(u^{k+1}-u^k) = -(\tau u^k - g), \quad \text{on } \partial\Omega.$$

with the differential operator $L_k \in \partial^* F(u^k)$ defined as in (4.3.16). Since $F : H^2(\Omega) \to L^2(\Omega)$ is $\partial^* F$-semismooth (see Corollary 4.3.5) and $\tau \in \mathcal{L}(H^2(\Omega), H^{3/2}(\partial\Omega))$, we can conclude that Algorithm 1 is in fact a semismooth Newton method for solving the operator equation $\tilde{F}(u) = 0$.

Note that the boundedness of coefficients and the classical theory of elliptic regularity (see Theorem B.1.1) imply that under Assumption 10, there exists a constant $C > 0$, such that for any $u \in H^2(\Omega)$ and any $L \in \partial^* F(u)$, the inverse operator $(L, \tau)^{-1} : L^2(\Omega) \times H^{3/2}(\partial\Omega) \to H^2(\Omega)$ is well-defined, and the operator norm $\|(L, \tau)^{-1}\|$ is bounded by C, uniformly in $u \in H^2(\Omega)$. Hence one can conclude from [99, Theorem 3.13] (see also Theorem B.1.5) that the iterates $\{u^k\}_{k\in\mathbb{N}}$ converges superlinearly to u^* in a neighborhood $\mathcal{N}$ of u^*. $\square$

The next theorem strengthens Theorem 4.3.6, and establishes a novel global convergence result of Algorithm 1 applied to the Dirichlet problem (4.2.2), which subsequently provides a constructive proof for the existence of solutions to (4.2.2). As pointed out in [19, Remark 5.8], policy iteration may not converge globally for nonconvex HJBI equations, even in a finite-dimensional setting. The following additional condition is essential for our proof of the global convergence of Algorithm 1, which ensures that Algorithm 1 behaves like a fixed-point algorithm at the initial stage.

Assumption 11. *Let the function c in Assumption 10 be given as: $c(x, \alpha, \beta) = \bar{c}(x, \alpha, \beta) + \underline{c}_0$, for all $(x, \alpha, \beta) \in \Omega \times \boldsymbol{A} \times \boldsymbol{B}$, where $\underline{c}_0$ is a sufficiently large constant, depending on Ω, $\{a^{ij}\}_{i,j=1}^n$, $\{b^i\}_{i=1}^n$ and $\|\bar{c}\|_{L^\infty(\Omega\times\boldsymbol{A}\times\boldsymbol{B})}$.*

In practice, Assumption 11 can be satisfied if (4.2.2) arises from an infinite-horizon stochastic game with a large discount factor (see e.g. [24]), or if (4.2.2) stems from an implicit (time-)discretization of parabolic HJBI equations with a small time stepsize.

Theorem 4.3.7. *Suppose Assumptions 10 and 11 hold, then the Dirichlet problem (4.2.2) admits a unique strong solution $u^* \in H^2(\Omega)$. Moreover, for any initial guess $u^0 \in H^2(\Omega)$, Algorithm 1 either terminates with $u^k = u^*$ for some $k \in \mathbb{N}$, or generates a sequence $\{u^k\}_{k\in\mathbb{N}}$ that converges q-superlinearly to u^* in $H^2(\Omega)$, i.e., $\lim_{k\to\infty} \|u^{k+1} - u^*\|_{H^2(\Omega)}/\|u^k - u^*\|_{H^2(\Omega)} = 0$.*

Proof. If Algorithm 1 terminates in iteration k, we have $F(u^k) = L_k u^k - f_k = 0$ and $\tau u^k = g$, from which, we obtain from the uniqueness of strong solutions to (4.2.2) (Proposition 4.2.1) that $u^k = u^*$ is the strong solution to the Dirichlet problem (4.2.2). Hence we shall assume without loss of generality that Algorithm 1 runs infinitely.

We now establish the global convergence of Algorithm 1 by first showing the iterates $\{u^k\}_{k\in\mathbb{N}}$ form a Cauchy sequence in $H^2(\Omega)$. For each $k \geq 0$, we deduce from (4.3.6) and Assumption 11 that $\tau u^{k+1} = g$ on $\partial\Omega$ and

$$\begin{aligned}
&-a^{ij}\partial_{ij}u^{k+1} + b^i_k\partial_i u^{k+1} + c_k u^{k+1} - f_k = -a^{ij}\partial_{ij}u^{k+1} + b^i_k\partial_i u^{k+1} + (\bar{c}_k + \underline{c}_0)u^{k+1} - f_k \\
&\quad = -a^{ij}\partial_{ij}u^{k+1} + \underline{c}_0 u^{k+1} + b^i_k\partial_i(u^{k+1} - u^k) + \bar{c}_k(u^{k+1} - u^k) + \bar{G}(\cdot, u^k, \nabla u^k) = 0,
\end{aligned} \tag{4.3.18}$$

for a.e. $x \in \Omega$, where the function $\bar{c}_k(x) \coloneqq \bar{c}(x, \alpha^k(x), \beta^k(x))$ for all $x \in \Omega$, and the modified Hamiltonian is defined as:

$$\bar{G}(x, u, \nabla u) = \max_{\alpha\in\mathbf{A}}\min_{\beta\in\mathbf{B}}\big(b^i(x,\alpha,\beta)\partial_i u(x) + \bar{c}(x,\alpha,\beta)u(x) - f(x,\alpha,\beta)\big). \tag{4.3.19}$$

Hence, by taking the difference of equations corresponding to the indices $k-1$ and k, one can obtain that

$$\begin{aligned}
&-a^{ij}\partial_{ij}(u^{k+1} - u^k) + \underline{c}_0(u^{k+1} - u^k) = -b^i_k\partial_i(u^{k+1} - u^k) - \bar{c}_k(u^{k+1} - u^k) \\
&\quad + b^i_{k-1}\partial_i(u^k - u^{k-1}) + \bar{c}_{k-1}(u^k - u^{k-1}) - [\bar{G}(\cdot, u^k, \nabla u^k) - \bar{G}(\cdot, u^{k-1}, \nabla u^{k-1})],
\end{aligned} \tag{4.3.20}$$

for $x \in \Omega$, and $\tau(u^{k+1} - u^k) = 0$ on $\partial\Omega$.

It has been proved in Theorem 9.14 of [43] that, there exist positive constants C and γ_0, depending only on $\{a^{ij}\}_{i,j=1}^n$ and Ω, such that it holds for all $u \in H^2(\Omega)$ with $\tau u = 0$, and for all $\gamma \geq \gamma_0$ that

$$\|u\|_{H^2(\Omega)} \leq C\| - a^{ij}\partial_{ij}u + \gamma u\|_{L^2(\Omega)},$$

which, together with the identity that $\gamma u = (-a^{ij}\partial_{ij}u + \gamma u) + a^{ij}\partial_{ij}u$ and the boundedness of $\{a^{ij}\}_{ij}$, implies that the same estimate also holds for $\|u\|_{H^2(\Omega)} + \gamma\|u\|_{L^2(\Omega)}$:

$$\|u\|_{H^2(\Omega)} + \gamma\|u\|_{L^2(\Omega)} \leq C\| - a^{ij}\partial_{ij}u + \gamma u\|_{L^2(\Omega)}.$$

Thus, by assuming $\underline{c}_0 \geq \gamma_0$ and using the boundedness of the coefficients, we can deduce from (4.3.20) that

$$\begin{aligned}
&\|u^{k+1} - u^k\|_{H^2(\Omega)} + \underline{c}_0\|u^{k+1} - u^k\|_{L^2(\Omega)} \\
&\leq C\Big(\| - b^i_k\partial_i(u^{k+1} - u^k) - \bar{c}_k(u^{k+1} - u^k) + b^i_{k-1}\partial_i(u^k - u^{k-1}) \\
&\qquad + \bar{c}_{k-1}(u^k - u^{k-1}) - [\bar{G}(\cdot, u^k, \nabla u^k) - \bar{G}(\cdot, u^{k-1}, \nabla u^{k-1})]\|_{L^2(\Omega)}\Big) \\
&\leq C\big(\|u^{k+1} - u^k\|_{H^1(\Omega)} + \|u^k - u^{k-1}\|_{H^1(\Omega)}\big),
\end{aligned} \tag{4.3.21}$$

for some constant C independent of $\underline{c}_0$ and the index k.

Now we apply the following interpolation inequality (see [43, Theorem 7.28]): there exists a constant C, such that for all $u \in H^2(\Omega)$ and $\varepsilon > 0$, we have $\|u\|_{H^1(\Omega)} \leq \varepsilon\|u\|_{H^2(\Omega)} + C\varepsilon^{-1}\|u\|_{L^2(\Omega)}$. Hence, for any given $\varepsilon_1 \in (0,1), \varepsilon_2 > 0$, we have

$$\begin{aligned}&(1-\varepsilon_1)\|u^{k+1}-u^k\|_{H^2(\Omega)} + \underline{c}_0\|u^{k+1}-u^k\|_{L^2(\Omega)}\\ &\leq \varepsilon_2\|u^k-u^{k-1}\|_{H^2(\Omega)} + C\varepsilon_1^{-1}\|u^{k+1}-u^k\|_{L^2(\Omega)} + C\varepsilon_2^{-1}\|u^k-u^{k-1}\|_{L^2(\Omega)}.\end{aligned}$$

Then, by taking $\varepsilon_1 \in (0,1)$, $\varepsilon_2 < 1-\varepsilon_1$, and assuming that $\underline{c}_0$ satisfies $(\underline{c}_0 - C/\varepsilon_1)/(1-\varepsilon_1) \geq C/\varepsilon_2^2$, we can obtain for $c' = C/\varepsilon_2^2$ that

$$\|u^{k+1}-u^k\|_{H^2(\Omega)} + c'\|u^{k+1}-u^k\|_{L^2(\Omega)} \leq \frac{\varepsilon_2}{1-\varepsilon_1}\big(\|u^k-u^{k-1}\|_{H^2(\Omega)} + c'\|u^k-u^{k-1}\|_{L^2(\Omega)}\big),$$

which implies that $\{u^k\}_{k\in\mathbb{N}}$ is a Cauchy sequence with the norm $\|\cdot\|_{c'} := \|\cdot\|_{H^2(\Omega)} + c'\|\cdot\|_{L^2(\Omega)}$.

Since $\|\cdot\|_{c'}$ is equivalent to $\|\cdot\|_{H^2(\Omega)}$ on $H^2(\Omega)$, we can deduce that $\{u^k\}_{k\in\mathbb{N}}$ converges to some $\bar{u}$ in $H^2(\Omega)$. By passing $k \to \infty$ in (4.3.18) and using Proposition 4.2.1, we can deduce that $\bar{u} = u^*$ is the unique strong solution of (4.2.2). Finally, for a sufficiently large $K_0 \in \mathbb{N}$, we can conclude the superlinear convergence of $\{u^k\}_{k\geq K_0}$ from Theorem 4.3.6. □

Remark 4.3.3. If one of the sets $\mathbf{A}$ and $\mathbf{B}$ is a singleton, and $a^{ij} \in C^{0,1}(\bar{\Omega})$ for all i,j, then Algorithm 1 applied to the Dirichlet problem (4.2.2) is in fact monotonically convergent with an arbitrary initial guess. Suppose, for instance, that $\mathbf{A}$ is a singleton, then for each $k \in \mathbb{N}\cup\{0\}$, we have that

$$0 = L_k u^{k+1} - f_k \geq F(u^{k+1}) = -L_{k+1}(u^{k+2}-u^{k+1}), \quad \text{for a.e. } x \in \Omega.$$

Hence we can deduce that $w^{k+1} := u^{k+1} - u^{k+2}$ is a weak subsolution to $L_{k+1}w = 0$, i.e.,

$$\int_\Omega \Big[a^{ij}\partial_j w^{k+1}\partial_i\phi + \big((\partial_i a^{ij} + b^i_{k+1})\partial_i w^{k+1} + c_{k+1}w^{k+1}\big)\phi\Big]\,dx \leq 0, \quad \forall\phi \geq 0,\ \phi \in C^1_0(\Omega).$$

Thus, the weak maximal principle (see [42, Theorem 1.3.7]) and the fact that $u^{k+1} - u^{k+2} = 0$ a.e. $x \in \partial\Omega$ (with respect to the surface measure), leads to the estimate $\operatorname{ess\,sup}_\Omega u^{k+1} - u^{k+2} \leq 0$, which consequently implies that $u^k \leq u^{k+1}$ for all $k \in \mathbb{N}$ and a.e. $x \in \Omega$.

4.4 Inexact policy iteration for HJBI Dirichlet problems

Note that at each policy iteration, Algorithm 1 requires us to obtain an exact solution to a linear Dirichlet boundary value problem, which is generally infeasible. Moreover, an accurate computation of numerical solutions to linear Dirichlet boundary value problems could be expensive, especially in a high-dimensional setting. In this section, we shall propose an inexact policy iteration algorithm for (4.2.2), where we compute an approximate solution to (4.3.3) by solving an optimization problem over a family of trial functions, while maintaining the superlinear convergence of policy iteration.

We shall make the following assumption on the trial functions of the optimization problem.

Assumption 12. *The collections of trial functions $\{\mathcal{F}_M\}_{M\in\mathbb{N}}$ satisfies the following properties: $\mathcal{F}_M \subset \mathcal{F}_{M+1}$ for all $M \in \mathbb{N}$, and $\mathcal{F} = \{\mathcal{F}_M\}_{M\in\mathbb{N}}$ is dense in $H^2(\Omega)$.*

It is clear that Assumption 12 is satisfied by any reasonable H^2-conforming finite element spaces (see e.g. [22]) and high-order polynomial spaces or kernel-function spaces used in global spectral methods (see e.g. [13, 12, 29, 65, 66]). We now demonstrate that Assumption 12 can also be easily satisfied by the sets of multi-layer feedforward neural networks, which provides effective trial functions for high-dimensional problems. Let us first recall the definition of a feedforward neural network.

Definition 4.4.1 (Artificial neural networks)**.** Let $L, N_0, N_1, \ldots, N_L \in \mathbb{N}$ be given constants, and $\varrho : \mathbb{R} \to \mathbb{R}$ be a given function. For each $l = 1, \ldots, L$, let $T_l : \mathbb{R}^{N_{l-1}} \to \mathbb{R}^{N_l}$ be an affine function given as $T_l(x) = W_l x + b_l$ for some $W_l \in \mathbb{R}^{N_l \times N_{l-1}}$ and $b_l \in \mathbb{R}^{N_l}$. A function $F : \mathbb{R}^{N_0} \to \mathbb{R}^{N_L}$ defined as

$$F(x) = T_L \circ (\varrho \circ T_{L-1}) \circ \cdots (\varrho \circ T_1), \quad x \in \mathbb{R}^{N_0},$$

is called a feedforward neural network. Here the activation function ϱ is applied componentwise. We shall refer the quantity L as the depth of F, $N_1, \ldots N_{L-1}$ as the dimensions of the hidden layers, and N_0, N_L as the dimensions of the input and output layers, respectively. We also refer to the number of entries of $\{W_l, b_l\}_{l=1}^N$ as the complexity of F.

Let $\{L^{(M)}\}_{M\in\mathbb{N}}, \{N_1^{(M)}\}_{M\in\mathbb{N}}, \ldots, \{N_{L^{(M)}-1}^{(M)}\}_{M\in\mathbb{N}}$ be some nondecreasing sequences of natural numbers, we define for each M the set $\mathcal{F}_M$ of all neural networks with depth $L^{(M)}$, input dimension being equal to n, output dimension being equal to 1, and dimensions of hidden layers being equal to $\{N_1^{(M)}, \dots, N_{L^{(M)}-1}^{(M)}\}_{M\in\mathbb{N}}$. It is clear that if $L^{(M)} \equiv L$ for all $M \in \mathbb{N}$, then we have $\mathcal{F}_M \subset \mathcal{F}_{M+1}$. The following proposition is proved in [50, Corollary 3.8], which shows neural networks with one hidden layer are dense in $H^2(\Omega)$.

Proposition 4.4.1. *Let $\Omega \subset \mathbb{R}^n$ be an open bounded starshaped domain, and $\varrho \in C^2(\mathbb{R})$ satisfying $0 < |D^l\varrho|_{L^1(\Omega)} < \infty$ for all $l = 1, 2$. Then the family of all neural networks with depth $L = 2$ is dense in $H^2(\Omega)$.*

Now we discuss how to approximate the strong solutions of Dirichlet problems by reformulating the equations into optimization problems over trial functions. The idea is similar to least squares finite-element methods (see e.g. [17]), and has been employed previously to develop numerical methods for PDEs based on neural networks (see e.g. [74, 16, 95]). However, compared to [16, 74], we do not impose additional constraints on the trial functions by requiring that the networks exactly agree with the boundary conditions, due to the lack of theoretical support that the constrained neural networks are still dense in the solution space. Moreover, to ensure the convergence of solutions in the $H^2(\Omega)$-norm, we include the $H^{3/2}(\partial\Omega)$-norm of the boundary data in the cost function, instead of the $L^2(\partial\Omega)$-norm used in [95] (see Remark 4.4.2 for more details).

For each $k \in \mathbb{N} \cup \{0\}$, let $u^{k+1} \in H^2(\Omega)$ be the unique solution to the Dirichlet problem (4.3.3):

$$L_k u - f_k = 0, \text{ in } \Omega; \quad \tau u = g, \text{ on } \partial\Omega,$$

where L_k and f_k denote the linear elliptic operator and the source term in (4.3.3), respectively. For each $M \in \mathbb{N}$, we shall consider the following optimization problems:

$$J_{k,M} := \inf_{u\in\mathcal{F}_M} J_k(u), \quad \text{with } J_k(u) = \|L_k u - f_k\|^2_{L^2(\Omega)} + \|\tau u - g\|^2_{H^{3/2}(\partial\Omega)}. \tag{4.4.1}$$

The following result shows that the cost function J_k provides a computable indicator of the error.

Proposition 4.4.2. *Suppose Assumptions 10 and 12 hold. For each $k \in \mathbb{N} \cup \{0\}$ and $M \in \mathbb{N}$, let $u^{k+1} \in H^2(\Omega)$ be the unique solution to (4.3.3), and J_k, $J_{k,M}$ be defined as in (4.4.1). Then there exist positive constants C_1 and C_2, such that we have for each $u \in H^2(\Omega)$ and $k \in \mathbb{N} \cup \{0\}$ that*

$$C_1 J_k(u) \le \|u - u^{k+1}\|^2_{H^2(\Omega)} \le C_2 J_k(u).$$

Consequently, it holds for each $k \in \mathbb{N} \cup \{0\}$ that $\lim_{M\to\infty} J_{k,M} = 0$.

Proof. Let $k \in \mathbb{N} \cup \{0\}$ and $u \in H^2(\Omega)$. The definition of $J_k(u)$ implies that $L_k u - f_k = f^e \in L^2(\Omega)$, $\tau u - g = g^e \in H^{3/2}(\partial\Omega)$ and $J(u) = \|f^e\|^2_{L^2(\Omega)} + \|g^e\|^2_{H^{3/2}(\partial\Omega)}$. Then, by using the assumption that u^{k+1} solves (4.3.3), we deduce that the residual term satisfies the following Dirichlet problem:

$$L_k(u - u^{k+1}) = f^e, \quad \text{in } \Omega; \quad \tau(u - u^{k+1}) = g^e, \quad \text{on } \partial\Omega.$$

Hence the boundedness of coefficients and the regularity theory of elliptic operators (see Theorem B.1.1) lead to the estimate that

$$C_1(\|f^e\|^2_{L^2(\Omega)} + \|g^e\|^2_{H^{3/2}(\partial\Omega)}) \leq \|u - u^{k+1}\|^2_{H^2(\Omega)} \leq C_2(\|f^e\|^2_{L^2(\Omega)} + \|g^e\|^2_{H^{3/2}(\partial\Omega)}),$$

where the constants $C_1, C_2 > 0$ depend only on the $L^\infty(\Omega)$-norms of a^{ij}, b^i_k, c_k, f_k, which are independent of k. The above estimate, together with the facts that $\{\mathcal{F}_M\}_{M\in\mathbb{N}}$ is dense in $H^2(\Omega)$ and $\mathcal{F}_M \subset \mathcal{F}_{M+1}$, leads to the desired result that $\lim_{M\to\infty} J_{k,M} = 0$. □

We now present the inexact policy iteration algorithm for the HJBI problem (4.2.2), where at each policy iteration, we solve the linear Dirichlet problem within a given accuracy.

Algorithm 2 Inexact policy iteration algorithm for Dirichlet problems.

---

1. Choose a family of trial functions $\mathcal{F} = \{\mathcal{F}_M\}_{M\in\mathbb{N}} \subset H^2(\Omega)$, an initial guess u^0 in $\mathcal{F}$, a sequence $\{\eta_k\}_{k\in\mathbb{N}\cup\{0\}}$ of positive scalars, and set $k = 0$.

2. Given the iterate u^k, update the control laws α^k and β^k by (4.3.1) and (4.3.2), respectively.

3. Find $u^{k+1} \in \mathcal{F}$ such that[1]

$$J_k(u^{k+1}) = \|L_k u^{k+1} - f_k\|^2_{L^2(\Omega)} + \|\tau u^{k+1} - g\|^2_{H^{3/2}(\partial\Omega)} \leq \eta_{k+1} \min(\|u^{k+1} - u^k\|^2_{H^2(\Omega)}, \eta_0), \tag{4.4.2}$$

   where L_k and f_k denote the linear operator and the source term in (4.3.3), respectively.

4. If $\|u^{k+1} - u^k\|_{H^2(\Omega)} = 0$, then terminate with outputs u^{k+1}, α^k and β^k, otherwise increment k by one and go to step 2.

---

[1]With a slight abuse of notation, we denote by u^{k+1} the inexact solution to the Dirichlet problem (4.3.3).

Remark 4.4.1. In practice, the evaluation of the squared residuals J_k in (4.4.2) depends on the choice of trial functions. For trial functions with linear architecture, e.g. if $\{\mathcal{F}_M\}_{M\in\mathbb{N}}$ are finite element spaces, high-order polynomial spaces, and kernel-function spaces (see [22, 29, 65, 66]), one may evaluate the norms by applying high-order quadrature rules to the basis functions involved.

For trial functions with nonlinear architecture, such as feedforward neural networks, we can replace the integrations in J_k by the empirical mean over suitable collocation points in Ω and on $\partial\Omega$, such as pseudorandom points or quasi-Monte Carlo points (see Section 4.6; see also [74, 16, 95]). In particular, due to the existence of local coordinate charts of the boundaries, we can evaluate the double integral in the definition of the $H^{3/2}(\partial\Omega)$-norm (see (4.2.1)) by first generating points in $\mathbb{R}^{2(n-1)}$ and then mapping the samples onto $\partial\Omega\times\partial\Omega$. The resulting empirical least-squares problem for the $k+1$-th policy iteration step (cf. (4.4.1)) can then be solved by stochastic gradient descent (SGD) algorithms; see Section 4.6. We remark that, instead of pre-generating all the collocation points in advance, one can perform gradient descent based on a sequence of mini-batches of points generated at each SGD iteration. This is particularly useful in higher dimensions, where many collocation points may be needed to cover the boundary, and using mini-batches avoids having to evaluate functions at all collocation points in each iteration.

It is well-known (see e.g. [99, 35]) that the residual term $\|u^{k+1}-u^k\|_{H^2(\Omega)}$ is crucial for the superlinear convergence of inexact Newton methods. This next theorem establishes the global superlinear convergence of Algorithm 2.

Theorem 4.4.3. *Suppose Assumptions 10, 11 and 12 hold, and* $\lim_{k\to\infty}\eta_k=0$ *in Algorithm 2. Let* $u^*\in H^2(\Omega)$ *be the solution to the Dirichlet problem* (4.2.2)*. Then for any initial guess* $u^0\in\mathcal{F}$*, Algorithm 2 either terminates with* $u^k=u^*$ *for some* $k\in\mathbb{N}$*, or generates a sequence* $\{u^k\}_{k\in\mathbb{N}}$ *that converges q-superlinearly to* u^* *in* $H^2(\Omega)$*, i.e.,* $\lim_{k\to\infty}\|u^{k+1}-u^*\|_{H^2(\Omega)}/\|u^k-u^*\|_{H^2(\Omega)}=0$*. Consequently, we have* $\lim_{k\to\infty}(u^k,\partial_i u^k,\partial_{ij}u^k)(x)=(u^*,\partial_i u^*,\partial_{ij}u^*)(x)$ *for a.e.* $x\in\Omega$*, and for all* $i,j=1,\dots,n$*.*

Proof. Let $u^0\in\mathcal{F}$ be an arbitrary initial guess. We first show that Algorithm 2 is always well-defined. For each $k\in\mathbb{N}\cup\{0\}$, if $u^k\in\mathcal{F}$ is the strong solution to (4.3.3), then we can choose $u^{k+1}=u^k$, which satisfies (4.4.2) and terminates the algorithm. If u^k does not solve (4.3.3), the fact that $\mathcal{F}$ is dense in $H^2(\Omega)$ enables us to find $u^{k+1}\in\mathcal{F}$ satisfying the criterion (4.4.2).

Moreover, one can clearly see from (4.4.2) that if Algorithm 2 terminates at iteration k, then u^k is the exact solution to the Dirichlet problem (4.2.2). Hence in the sequel we shall assume without loss of generality that Algorithm 2 runs infinitely, i.e., $\|u^{k+1} - u^k\|_{H^2(\Omega)} > 0$ and $u^k \neq u^*$ for all $k \in \mathbb{N} \cup \{0\}$.

We next show the iterates converge to u^* in $H^2(\Omega)$ by following similar arguments as those for Theorem 4.3.7. For each $k \geq 0$, we can deduce from (4.4.2) that there exists $f_k^e \in L^2(\Omega)$ and $g_k^e \in H^{3/2}(\partial\Omega)$ such that

$$L_k u^{k+1} - f_k = f_k^e, \quad \text{in } \Omega; \quad \tau u^{k+1} - g = g_k^e, \quad \text{on } \partial\Omega, \tag{4.4.3}$$

and $J_k(u^{k+1}) = \|f_k^e\|^2_{L^2(\Omega)} + \|g_k^e\|^2_{H^{3/2}(\Omega)} \leq \eta_{k+1}(\|u^{k+1} - u^k\|^2_{H^2(\Omega)})$ with $\lim_{k\to\infty} \eta_k = 0$. Then, by taking the difference between (4.4.3) and (4.2.2), we obtain that

$$\begin{aligned}-a^{ij}\partial_{ij}(u^{k+1} - u^*) + \underline{c}_0(u^{k+1} - u^*) &= -b_k^i \partial_i(u^{k+1} - u^k) - \bar{c}_k(u^{k+1} - u^k) \\ &\quad - [\bar{G}(\cdot, u^k, \nabla u^k) - \bar{G}(\cdot, u^*, \nabla u^*)] + f_k^e, \quad \text{in } \Omega,\end{aligned}$$

and $\tau(u^{k+1} - u^*) = g_k^e$ on $\partial\Omega$, where $\bar{G}$ is the modified Hamiltonian defined as in (4.3.19). Then, by proceeding along the lines of Theorem 4.3.7, we can obtain a positive constant C, independent of $\underline{c}_0$ and the index k, such that

$$\begin{aligned}&\|u^{k+1} - u^*\|_{H^2(\Omega)} + \underline{c}_0\|u^{k+1} - u^*\|_{L^2(\Omega)} \\ &\leq C\big(\|u^{k+1} - u^*\|_{H^1(\Omega)} + \|u^{k+1} - u^k\|_{H^1(\Omega)} + \|u^k - u^*\|_{H^1(\Omega)}\big) + o(\|u^{k+1} - u^k\|_{H^2(\Omega)}) \\ &\leq C\big(\|u^{k+1} - u^*\|_{H^1(\Omega)} + \|u^k - u^*\|_{H^1(\Omega)}\big) + o(\|u^{k+1} - u^*\|_{H^2(\Omega)} + \|u^k - u^*\|_{H^2(\Omega)})\end{aligned}$$

as $k \to \infty$, where the additional high-order terms are due to the residuals f_k^e and g_k^e. Then, by using the interpolation inequality and assuming $\underline{c}_0$ is sufficiently large, we can deduce that $\{u^k\}_{k\in\mathbb{N}}$ converge linearly to u^* in $H^2(\Omega)$.

We then reformulate Algorithm 2 into a quasi-Newton method for the operator equation $\tilde{F}(u) = 0$, with the operator $\tilde{F} : u \in H^2(\Omega) \to (F(u), \tau u - g) \in L^2(\Omega) \times H^{3/2}(\partial\Omega)$ defined in the proof of Theorem 4.3.6. Let $H^2(\Omega)^*$ denote the strong dual space of $H^2(\Omega)$, and $\langle \cdot, \cdot \rangle$ denote the dual product on $H^2(\Omega)^* \times H^2(\Omega)$. For each $k \in \mathbb{N} \cup \{0\}$, by using the fact that $\|u^{k+1} - u^k\|_{H^2(\Omega)} > 0$, we can choose $w_k \in H^2(\Omega)^*$ satisfying $\langle w_k, u^{k+1} - u^k \rangle = -1$, and introduce the following linear operators $\delta L_k \in \mathcal{L}(H^2(\Omega), L^2(\Omega))$ and $\delta\tau_k \in \mathcal{L}(H^2(\Omega), H^{3/2}(\partial\Omega))$:

$$\begin{aligned}&\delta L_k : v \in H^2(\Omega) \to \langle w_k, v \rangle f_k^e \in L^2(\Omega) \\ &\delta\tau_k : v \in H^2(\Omega) \to \langle w_k, v \rangle g_k^e \in H^{3/2}(\partial\Omega).\end{aligned}$$

Then, we can apply the identity $F(u^k) = L_k u^k - f_k$ and rewrite (4.4.3) as:

$$(L_k+\delta L_k)(u^{k+1}-u^k) = -F(u^k), \quad \text{in } \Omega; \quad (\tau+\delta\tau_k)(u^{k+1}-u^k) = -(\tau u^k - g), \quad \text{on } \partial\Omega,$$

with $(L_k, \tau) \in \partial^* \tilde{F}(u^k)$ as shown in Theorem 4.3.6. Hence one can clearly see that (4.4.3) is precisely a Newton step with a perturbed operator for the equation $\tilde{F}(u) = 0$.

Now we are ready to establish the superlinear convergence of $\{u^k\}_{k\in\mathbb{N}}$. For notational simplicity, in the subsequent analysis we shall denote by $Z := L^2(\Omega)\times H^{3/2}(\partial\Omega)$ the Banach space with the usual product norm $\|z\|_Z := \|z_1\|_{L^2(\Omega)} + \|z_2\|_{H^{3/2}(\partial\Omega)}$ for each $z = (z_1, z_2) \in Z$. By using the semismoothess of $\tilde{F} : H^2(\Omega) \to Z$ (see Theorem 4.3.6) and the strong convergence of $\{u^k\}_{k\in\mathbb{N}}$ in $H^2(\Omega)$, we can directly infer from Theorem B.1.5 that it remains to show that there exists a open neighborhood V of u^*, and a constant $L > 0$, such that

$$\|v - u^*\|_{H^2(\Omega)}/L \le \|\tilde{F}(v) - \tilde{F}(u^*)\|_Z \le L\|v - u^*\|_{H^2(\Omega)}, \ \forall v \in V, \tag{4.4.4}$$

and also

$$\lim_{k\to\infty} \|(\delta L_k s^k, \delta\tau_k s^k)\|_Z / \|s^k\|_{H^2(\Omega)} = 0, \quad \text{with } s^k = u^{k+1} - u^k \text{ for all } k \in \mathbb{N}. \tag{4.4.5}$$

The criterion (4.4.2) and the definitions of δL_k and $\delta\tau_k$ imply that (4.4.5) holds:

$$\left(\frac{\|(\delta L_k s^k, \delta\tau_k s^k)\|_Z}{\|s^k\|_{H^2(\Omega)}}\right)^2 = \left(\frac{\|f_k^e\|_{L^2(\Omega)} + \|g_k^e\|_{H^{3/2}(\partial\Omega)}}{\|s^k\|_{H^2(\Omega)}}\right)^2 \le \frac{2J_k(u^{k+1})}{\|s^k\|^2_{H^2(\Omega)}} \le 2\eta_0\eta_{k+1} \to 0,$$

as $k \to \infty$. Moreover, the boundedness of the coefficients a^{ij}, b^i, c, f shows that $\tilde{F}$ is Lipschitz continuous. Finally, the characterization of the generalized differential of $\tilde{F}$ in Theorem 4.3.6 and the regularity theory of elliptic operators (see Theorem B.1.1) show that for each $v \in H^2(\Omega)$, we can choose an invertible operator $M_v = (L_v, \tau) \in \partial^*\tilde{F}(v)$ such that $\|M_v^{-1}\|_{\mathcal{L}(Z,H^2(\Omega))} \le C < \infty$, uniformly in v. Thus we can conclude from the semismoothness of $\tilde{F}$ at u^* that

$$\begin{aligned}
\|\tilde{F}(v) - \tilde{F}(u^*)\|_Z &= \|M_v(v - u^*) + o(\|v - u^*\|_{H^2(\Omega)})\|_Z \\
&\ge \|M_v(v - u^*)\|_Z - o(\|v - u^*\|_{H^2(\Omega)}) \\
&\ge \|v - u^*\|_{H^2(\Omega)}/C - o(\|v - u^*\|_{H^2(\Omega)}) \\
&\ge \|v - u^*\|_{H^2(\Omega)}/(2C),
\end{aligned}$$

for all v in some neighborhood V of u^*, which completes our proof for q-superlinear convergence of $\{u^k\}_{k\in\mathbb{N}}$.

Finally, we establish the pointwise convergence of $\{u^k\}_{k=1}^\infty$ and their derivatives. For any given $\gamma \in (0,1)$, the superlinear convergence of $\{u^k\}_{k=1}^\infty$ implies that there exists a constant $C > 0$, depending on γ, such that $\|u^k - u^*\|_{H^2(\Omega)}^2 \le C\gamma^{2k}$ for all $k \in \mathbb{N}$. Taking the summation over the index k, we have

$$\begin{aligned}&\int_\Omega \sum_{k=1}^\infty \left(|u^k - u^*|^2 + \sum_{i,j=1}^n [|\partial_i u^k - \partial_i u^*|^2 + |\partial_{ij} u^k - \partial_{ij} u^*|^2] \right) dx \\ &= \sum_{k=1}^\infty \|u^k - u^*\|_{H^2(\Omega)}^2 \le \frac{C\gamma^2}{1-\gamma^2} < \infty,\end{aligned}$$

where we used the monotone convergence theorem in the first equality. Thus, we have

$$\sum_{k=1}^\infty \left(|u^k - u^*|^2 + \sum_{i,j=1}^n [|\partial_i u^k - \partial_i u^*|^2 + |\partial_{ij} u^k - \partial_{ij} u^*|^2] \right)(x) < \infty, \quad \text{for a.e. } x \in \Omega,$$

which leads us to the pointwise convergence of u^k and its partial derivatives with respect to k. □

Remark 4.4.2. We reiterate that merely including the $L^2(\partial\Omega)$-norm of the boundary data in the cost functional (4.4.2) in general cannot guarantee the convergence of the derivatives of the numerical solutions $\{u^k\}_{k=1}^\infty$, which can be seen from the following simple example. Let $\{g_k\}_{k=1}^\infty \subseteq H^{3/2}(\partial\Omega)$ be a sequence such that $g_k \to 0$ in $L^2(\partial\Omega)$ but not in $H^{1/2}(\partial\Omega)$, and for each $k \in \mathbb{N}$, let $h^k \in H^2(\Omega)$ be the strong solution to $-\Delta h^k = 0$ in Ω and $h^k = g_k$ on $\partial\Omega$.

The fact that $g_k \not\to 0$ in $H^{1/2}(\partial\Omega)$ implies that $h^k \not\to 0$ in $H^1(\Omega)$ as $k \to \infty$. We now show $\lim_{k\to\infty} h^k = 0$ in $L^2(\Omega)$. Let $w \in H^2(\Omega)$ be the solution to $-\Delta w = h^k$ in Ω and $w = 0$ on $\partial\Omega$, we can deduce from the integration by parts and the *a priori* estimate $\|w\|_{H^2(\Omega)} \le C\|h^k\|_{L^2(\Omega)}$ that

$$\begin{aligned}\|h^k\|_{L^2(\Omega)}^2 &= \int_\Omega (-\Delta w)\, h^k\, dx = \int_\Omega w\,(-\Delta h^k)\, dx + \int_{\partial\Omega} w \partial_n h^k \, d\sigma - \int_{\partial\Omega} h^k \partial_n w \, d\sigma \\ &\le C\|g_k\|_{L^2(\partial\Omega)} \|w\|_{H^2(\Omega)} \le C\|g_k\|_{L^2(\partial\Omega)} \|h^k\|_{L^2(\Omega)},\end{aligned}$$

which shows that $\|h^k\|_{L^2(\Omega)} \le C\|g_k\|_{L^2(\partial\Omega)} \to 0$ as $k \to \infty$. Now let $\mathcal{F}$ be a given family of trial functions, which is dense in $H^2(\Omega)$. One can find $\{u^k\}_{k=1}^\infty \subseteq \mathcal{F}$ satisfying $\lim_{k\to\infty} \|u^k - h^k\|_{H^2(\Omega)} = 0$, and consequently $u^k \not\to 0$ in $H^1(\Omega)$ as $k \to \infty$. However, we have

$$\|-\Delta u^k\|_{L^2(\Omega)}^2 + \|u^k\|_{L^2(\partial\Omega)}^2 = \|-\Delta(u^k - h^k)\|_{L^2(\Omega)}^2 + \|u^k - h^k + g_k\|_{L^2(\partial\Omega)}^2 \to 0, \quad \text{as } k \to \infty.$$

Similarly, one can construct functions $\{u^k\}_{k=1}^\infty \subseteq \mathcal{F}$ such that $\|-\Delta u^k\|_{L^2(\Omega)}^2 + \|u^k\|_{H^{1/2}(\partial\Omega)}^2 \to 0$ as $k \to \infty$, but $\{u^k\}_{k=1}^\infty$ does not converge to 0 in $H^2(\Omega)$.

We end this section with a convergent approximation of the optimal control strategies based on the iterates $\{u^k\}_{k=1}^\infty$ generated by Algorithm 2. For any given $u \in H^2(\Omega)$, we denote by $\mathbf{A}^u(x)$ and $\mathbf{B}^u(x,\alpha)$ the set of optimal control strategies for all $\alpha \in \mathbf{A}$ and for a.e. $x \in \Omega$, such that

$$\begin{aligned} \mathbf{B}^u(x,\alpha) &= \arg\min_{\beta\in\mathbf{B}} \big(b^i(x,\alpha,\beta)\partial_i u(x) + c(x,\alpha,\beta)u(x) - f(x,\alpha,\beta)\big), \\ \mathbf{A}^u(x) &= \arg\max_{\alpha\in\mathbf{A}} \min_{\beta\in\mathbf{B}} \big(b^i(x,\alpha,\beta)\partial_i u(x) + c(x,\alpha,\beta)u(x) - f(x,\alpha,\beta)\big). \end{aligned}$$

As an important consequence of the superlinear convergence of Algorithm 2, we now conclude that the feedback control strategies $\{\alpha^k\}_{k=1}^\infty$ and $\{\beta^k\}_{k=1}^\infty$ generated by Algorithm 2 are convergent to the optimal control strategies.

Corollary 4.4.4. *Suppose the assumptions of Theorem 4.4.3 hold, and let $u^* \in H^2(\Omega)$ be the solution to the Dirichlet problem (4.2.2). Assume further that there exist functions $\alpha^* : \Omega \to \boldsymbol{A}$ and $\beta^* : \Omega \to \boldsymbol{B}$ such that $\boldsymbol{A}^{u^*}(x) = \{\alpha^*(x)\}$ and $\boldsymbol{B}^{u^*}(x,\alpha^*(x)) = \{\beta^*(x)\}$ for a.e. $x \in \Omega$. Then the measurable functions $\alpha^k : \Omega \to \boldsymbol{A}$ and $\beta^k : \Omega \to \boldsymbol{B}$, $k \in \mathbb{N}$, generated by Algorithm 2 converge to the optimal feedback control (α^*, β^*) pointwise almost everywhere.*

Proof. Let ℓ and h be the Carathéodory functions defined by (4.3.4) and (4.3.5), respectively, and we consider the following set-valued mappings:

$$\begin{aligned} \Gamma_1 : (x,\mathbf{u}) \in \Omega \times \mathbb{R}^{n+1} &\rightrightarrows \Gamma_1(x,\mathbf{u}) \coloneqq \arg\max_{\alpha\in\mathbf{A}} h(x,\mathbf{u},\alpha), \\ \Gamma_2 : (x,\mathbf{u},\alpha) \in \Omega \times \mathbb{R}^{n+1} \times \mathbf{A} &\rightrightarrows \Gamma_2(x,\mathbf{u},\alpha) \coloneqq \arg\min_{\beta\in\mathbf{B}} \ell(x,\mathbf{u},\alpha,\beta). \end{aligned} \tag{4.4.6}$$

Theorem B.1.4 implies that the set-valued mappings $\Gamma_1(x,\cdot) : \mathbb{R}^{n+1} \rightrightarrows \mathbf{A}$ and $\Gamma_2(x,\cdot,\cdot) : \mathbb{R}^{n+1} \times \mathbf{A} \rightrightarrows \mathbf{B}$ are upper hemicontinuous. Then the result follows directly from the pointwise convergence of $(u^k, \nabla u^k)_{k=1}^\infty$ in Theorem 4.4.3, and the fact that $\mathbf{A}^{u^*}(x) = \{\alpha^*(x)\}$ and $\mathbf{B}^{u^*}(x,\alpha^*(x)) = \{\beta^*(x)\}$ are singleton for a.e. $x \in \Omega$. □

Remark 4.4.3. If we assume in addition that $\mathbf{A} \subset X_A$ and $\mathbf{B} \subset Y_B$ for some Banach spaces X_A and Y_B, then by using the compactness of $\mathbf{A}$ and $\mathbf{B}$ (see Assumption 10), we can conclude from the dominated convergence theorem that $\alpha^k \to \alpha^*$ in $L^p(\Omega; X_A)$ and $\beta^k \to \beta^*$ in $L^p(\Omega; Y_B)$, for any $p \in [1,\infty)$.

4.5 Inexact policy iteration for HJBI oblique derivative problems

In this section, we extend the algorithms introduced in previous sections to more general boundary value problems. In particular, we shall propose a neural network based policy iteration algorithm with global H^2-superlinear convergence for solving HJBI boundary value problems with oblique derivative boundary conditions. Similar arguments can be adapted to design superlinear convergent schemes for mixed boundary value problems with both Dirichlet and oblique derivative boundary conditions.

We consider the following HJBI oblique derivative problem:

$$F(u) := -a^{ij}(x)\partial_{ij}u + G(x, u, \nabla u) = 0, \quad \text{a.e. } x \in \Omega, \tag{4.5.1a}$$

$$Bu := \gamma^i \tau(\partial_i u) + \gamma^0 \tau u - g = 0, \quad \text{on } \partial\Omega. \tag{4.5.1b}$$

where (4.5.1a) is the HJBI equation given in (4.2.2a), and (4.5.1b) is an oblique boundary condition. Note that the boundary condition Bu on $\partial\Omega$ involves the traces of u and its first partial derivatives, which exist almost everywhere on $\partial\Omega$ (with respect to the surface measure).

The following conditions are imposed on the coefficients of (4.5.1):

Assumption 13. *Assume Ω, $\boldsymbol{A}$, $\boldsymbol{B}$, $(a^{ij})_{i,j=1}^n$, $(b^i)_{i=1}^n, c, f$ satisfy the same conditions as those in Assumption 10. Let $g \in H^{1/2}(\partial\Omega)$, $\{\gamma^i\}_{i=0}^n \subseteq C^{0,1}(\partial\Omega)$, $\gamma^0 \geq 0$ on $\partial\Omega$, and assume there exists a constant $\mu > 0$, such that $c \geq \mu$ on $\Omega \times \boldsymbol{A} \times \boldsymbol{B}$, and $\sum_{i=1}^n \gamma^i \nu_i \geq \mu$ on $\partial\Omega$, where $\{\nu_i\}_{i=1}^n$ are the components of the unit outer normal vector field on $\partial\Omega$.*

The next proposition establishes the well-posedness of the oblique derivative problem.

Proposition 4.5.1. *Suppose Assumption 13 holds. Then the oblique derivative problem (4.2.2) admits a unique strong solution $u^* \in H^2(\Omega)$.*

Proof. We shall establish the uniqueness of strong solutions to (4.5.1) in this proof, and then explicitly construct the solution in Theorem 4.5.2 with the help of policy iteration; see also Theorem 4.3.7. Suppose that $u, v \in H^2(\Omega)$ are two strong solutions to (4.5.1), then we can see $w = u - v$ is a strong solution to the following linear oblique derivative problem:

$$-a^{ij}\partial_{ij}w + \tilde{b}^i\partial_i w + \tilde{c}w = 0, \quad \text{a.e. in } \Omega; \quad \gamma^i\tau(\partial_i w) + \gamma^0\tau w = 0, \quad \text{on } \partial\Omega, \tag{4.5.2}$$

where $\tilde{b}^i$ is defined as in Proposition 4.2.1, and

$$\tilde{c}(x) = \begin{cases} \frac{G(x,u,\nabla u)-G(x,v,\nabla u)}{(u-v)(x)}, & \text{on } \{x \in \Omega \mid (u-v)(x) \neq 0\}, \\ \mu, & \text{otherwise.} \end{cases}$$

By following the same arguments as the proof of Proposition 4.2.1, we can show that $\tilde{b}^i, \tilde{c} \in L^\infty(\Omega)$, and $\tilde{c} \geq \mu > 0$ a.e. in Ω, which, along with Theorem B.1.2, implies that $w^* = 0$ is the unique strong solution to (4.5.2), and consequently $u = v$ in $H^2(\Omega)$. □

Now we present the neural network based policy iteration algorithm for solving the oblique derivative problem and establish its rate of convergence.

Algorithm 3 Inexact policy iteration algorithm for oblique derivative problems.

---

1. Choose a family of trial functions $\mathcal{F} = \{\mathcal{F}_M\}_{M\in\mathbb{N}} \subset H^2(\Omega)$, an initial guess u^0 in $\mathcal{F}$, a sequence $\{\eta_k\}_{k\in\mathbb{N}\cup\{0\}}$ of positive scalars, and set $k = 0$.

2. Given the iterate u^k, update the control laws α^k and β^k by (4.3.1) and (4.3.2), respectively.

3. Find $u^{k+1} \in \mathcal{F}$ such that

$$J_k(u^{k+1}) = \|L_k u^{k+1} - f_k\|^2_{L^2(\Omega)} + \|Bu^{k+1}\|^2_{H^{1/2}(\partial\Omega)} \leq \eta_{k+1} \min(\|u^{k+1} - u^k\|^2_{H^2(\Omega)}, \eta_0), \tag{4.5.3}$$

   where L_k, f_k, and B denote the linear operator in (4.3.3), the source term in (4.3.3) and the boundary operator in (4.5.1b), respectively.

4. If $\|u^{k+1} - u^k\|_{H^2(\Omega)} = 0$, then terminate with outputs u^{k+1}, α^k and β^k, otherwise increment k by one and go to step 2.

Note that the $H^{1/2}(\partial\Omega)$-norm of the boundary term is included in the cost function J_k, instead of the $H^{3/2}(\partial\Omega)$-norm as in Algorithm 2. It is straightforward to see that Algorithm 3 is well-defined under Assumptions 12 and 13. In fact, for each $k \in \mathbb{N} \cup \{0\}$, given the iterate $u^k \in \mathcal{F} \subset H^2(\Omega)$, Corollary 4.3.5 shows that one can select measurable control laws (α^k, β^k) such that the following linear oblique boundary value problem has measurable coefficients:

$$L_k u - f_k = 0, \text{ in } \Omega; \quad Bu = 0, \text{ on } \partial\Omega,$$

and hence admits a unique strong solution $\bar{u}^k$ in $H^2(\Omega)$ (see Theorem B.1.2). If $u^k = \bar{u}^k$, then u^k solve the HJBI oblique derivative problem (4.5.1), and we can select $u^{k+1} = u^k$ and terminate the algorithm. Otherwise, the facts that $J_k(u^{k+1}) \leq C\|\bar{u}^k - u^{k+1}\|^2_{H^2(\Omega)}$ and $\mathcal{F}$ is dense in $H^2(\Omega)$ allows us to choose $u^{k+1} \in \mathcal{F}$ sufficiently closed to $\bar{u}$ such that the criterion (4.5.3) is satisfied, and proceed to the next iteration.

The following result is analogue to Theorem 4.4.3, and shows the global superlinear convergence of Algorithm 3 for solving the oblique derivative problem (4.5.1). The proof follows precisely the lines given in Theorem 4.4.3, hence we shall only present the main steps in Appendix B.2 for the reader's convenience. The convergence of feedback control laws can be concluded similarly to Corollary 4.4.4 and Remark 4.4.3.

Theorem 4.5.2. *Suppose Assumptions 11, 12 and 13 hold, and* $\lim_{k\to\infty}\eta_k = 0$ *in Algorithm 3. Let* $u^* \in H^2(\Omega)$ *be the solution to the oblique derivative problem (4.5.1). Then for any initial guess* $u^0 \in \mathcal{F}$*, Algorithm 3 either terminates with* $u^k = u^*$ *for some* $k \in \mathbb{N}$*, or generates a sequence* $\{u^k\}_{k\in\mathbb{N}}$ *that converges q-superlinearly to* u^* *in* $H^2(\Omega)$*, i.e.,* $\lim_{k\to\infty}\|u^{k+1}-u^*\|_{H^2(\Omega)}/\|u^k-u^*\|_{H^2(\Omega)} = 0$*. Consequently, we have* $\lim_{k\to\infty}(u^k, \partial_i u^k, \partial_{ij}u^k)(x) = (u^*, \partial_i u^*, \partial_{ij}u^*)(x)$ *for a.e.* $x \in \Omega$*, and for all* $i, j = 1, \dots, n$.

4.6 Numerical experiments: Zermelo's Navigation Problem

In this section, we illustrate the theoretical findings and demonstrate the effectiveness of the schemes through numerical experiments. We present a two-dimensional convection-dominated HJBI Dirichlet boundary value problem in an annulus, which is related to stochastic minimum time problems.

In particular, we consider the stochastic Zermelo navigation problem (see e.g. [80]), which is a time-optimal control problem where the objective is to find the optimal trajectories of a ship/aircraft navigating a region of strong winds, modelled by a random vector field. Given a bounded open set $\Omega \subset \mathbb{R}^n$ and an adaptive control strategy $\{\alpha_t\}_{t\geq 0}$ taking values in $\mathbf{A}$, we assume the dynamics $X^{x,\alpha}$ of the ship is governed by the following controlled dynamics:

$$dX_t = b(X_t, \alpha_t)\,dt + \sigma\,dW_t, \quad t \in [0,\infty); \quad X_0 = x \in \Omega,$$

where the drift coefficient $b : \Omega \times \mathbf{A} \to \mathbb{R}^n$ is the sum of the velocity of the wind and the relative velocity of the ship, the nondegenerate diffusion coefficient $\sigma : \Omega \to \mathbb{R}^{n\times n}$ describes a random perturbation of the velocity field of the wind, and W is an n-dimensional Brownian motion defined on a probability space $(\tilde{\Omega}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$.

The aim of the controller is to minimize the expected exit time of the region Ω, taking model ambiguity into account in the spirit of [90]. More generally, we consider the following value function:

$$u(x) := \inf_{\alpha\in\mathcal{A}} \mathcal{E}\left[\int_0^{\tau_{x,\alpha}} f(X_t^{x,\alpha})\,dt + g(X_{\tau_{x,\alpha}}^{x,\alpha})\right] = \inf_{\alpha\in\mathcal{A}} \sup_{\mathbb{Q}\in\mathcal{M}} \mathbb{E}_{\mathbb{Q}}\left[\int_0^{\tau_{x,\alpha}} f(X_t^{x,\alpha})\,dt + g(X_{\tau_{x,\alpha}}^{x,\alpha})\right] \tag{4.6.1}$$

over all admissible choices of $\alpha \in \mathcal{A}$, where $\tau_{x,\alpha} := \inf\{t \geq 0 \mid X_t^{x,\alpha} \notin \Omega\}$ denotes the first exit time of the controlled dynamics $X^{x,\alpha}$, the functions f and g denote the running cost and the exit cost, respectively, which indicate the desired destinations, and $\mathcal{M}$ is a family of absolutely continuous probability measures with respect to $\mathbb{P}$ with density $M_t = \exp\big(\int_0^t \beta_t\,dW_t - \frac{1}{2}\int_0^t \beta_t^2\,dt\big)$, where $\{\beta_t\}_{t\geq 0}$ is a predictable process satisfying $\|\beta_t\|_\infty = \max_i |\beta_{t,i}| \leq \kappa$ for all t and a given parameter $\kappa \geq 0$. In other words, we would like to minimize a functional of the trajectory up to the exit time under the worst-case scenario, with uncertainty arising from the unknown law of the random perturbation.

By using the dual representation of $\mathcal{E}[\cdot]$ and the dynamic programming principle (see e.g. [90, 24]), we can characterize the value function u as the unique viscosity solution to an HJBI Dirichlet boundary value problem of the form (4.2.2). Moreover, under suitable assumptions, one can further show that u is the strong (Sobolev) solution to this Dirichlet problem (see e.g. [72]).

For our numerical experiments, we assume that the domain Ω is an annulus, i.e., $\Omega = \{(x,y) \in \mathbb{R}^2 \mid r^2 < x^2 + y^2 < R^2\}$, the wind blows along the positive x-axis with a magnitude v_c:

$$v_c(x,y) = 1 - a\sin\left(\pi\frac{x^2+y^2-r^2}{R^2-r^2}\right), \quad \text{for some constant } a \in [0,1),$$

which decreases in terms of the distance from the bank, and the random perturbation of the wind is given by the constant diffusion coefficient $\sigma = \operatorname{diag}(\sigma_x, \sigma_y)$. We also assume that the ship moves with a constant velocity v_s, and the captain can control the boat's direction instantaneously, which leads to the following dynamics of the boat in the region:

$$\begin{pmatrix} dX_t^{x,\alpha} \\ dY_t^{x,\alpha} \end{pmatrix} = \begin{pmatrix} v_c(X_t^{x,\alpha}, Y_t^{x,\alpha}) + v_s\cos(\alpha_t) \\ v_s\sin(\alpha_t) \end{pmatrix} dt + \begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_y \end{pmatrix} dW_t, \quad t \geq 0; \quad \begin{pmatrix} X_0^{x,\alpha} \\ Y_0^{x,\alpha} \end{pmatrix} = x,$$

where $\alpha_t \in \mathbf{A} = [0, 2\pi]$ represents the angle (measured counter-clockwise) between the positive x-axis and the direction of the boat. Finally, we assume the exit cost $g \equiv 0$ on $\partial B_r(0)$ and $g \equiv 1$ on $\partial B_R(0)$, which represents that the controller prefers to exit the domain through the inner boundary instead of the outer one (see Figure

4.1). Then the corresponding Dirichlet problem for the value function u in (4.6.1) is given by: $u \equiv 0$ on $\partial B_r(0)$, $u \equiv 1$ on $\partial B_R(0)$, and

$$\begin{aligned} &F(u) \\ &= -\frac{1}{2}(\sigma_x^2 u_{xx} + \sigma_y^2 u_{yy}) - v_c u_x - v_s \inf_{\alpha \in \mathbf{A}} \left[(\cos(\alpha), \sin(\alpha))^T \nabla u\right] - \sup_{\|\beta\|_\infty \leq \kappa} \left[\beta^T(\sigma \nabla u)\right] - f \\ &= -\frac{1}{2}(\sigma_x^2 u_{xx} + \sigma_y^2 u_{yy}) - v_c u_x + v_s \|\nabla u\|_{\ell^2} - \kappa \|\sigma \nabla u\|_{\ell^1} - f = 0, \quad \text{in } \Omega, \end{aligned} \tag{4.6.2}$$

where $\|\cdot\|_{\ell^1}$ and $\|\cdot\|_{\ell^2}$ denote the ℓ^1-norm and ℓ^2-norm on $\mathbb{R}^2$, respectively. The optimal feedback control laws can be further computed as

$$\alpha^* = \pi + \theta, \quad \beta^* = \kappa\big(\text{sgn}(\sigma_x u_x), \text{sgn}(\sigma_y u_y)\big)^T, \quad \text{a.e. in } \Omega, \tag{4.6.3}$$

where $u \in H^2(\Omega)$ is the strong solution to (4.6.2), and $\theta \in (-\pi, \pi]$ is the angle between ∇u and the positive x direction. Note that the equation (4.6.2) is neither convex nor concave in ∇u.

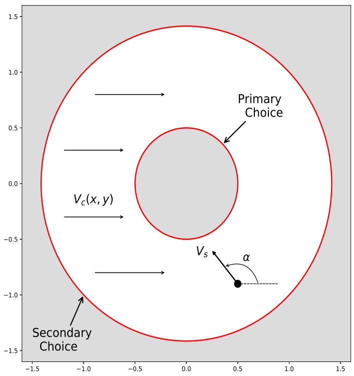


Figure 4.1: Zermelo navigation problem in an annulus.

4.6.1 Implementation details

In this section, we discuss the implementation details of Algorithm 2 for solving (4.6.2) with multi-layer neural networks (see Definition 4.4.1) as the trial functions. We shall now introduce the architecture of the neural networks, the involved hyper-parameters, and various computational aspects of the training process.

For simplicity, we shall adopt a fixed set of trial functions $\mathcal{F}_M$ for all policy iterations, which contains fully-connected networks with the activation function $\varrho(y) = \tanh(y)$, the depth L, and the dimension of each hidden layer H. The hyper-parameters L and H will be chosen depending on the complexity of the problem, which ensures that $\mathcal{F}_M$ admits sufficient flexibility to approximate the solutions within the desired accuracy. More complicated architectures of neural networks with shortcut connections can be adopted to further improve the performance of the algorithm (see e.g. [37, 95]).

We then proceed to discuss the computation of the cost functional J_k in (4.4.2) for each policy iteration. It is well-known that Sobolev norms of functions on sufficiently smooth boundaries can be explicitly computed via local coordinate charts of the boundaries (see e.g. [44]). In particular, due to the annulus shaped domain and the constant boundary conditions used in our experiment, we can express the cost functional J_k as follows: for all $k \in \mathbb{N}$ and $u \in \mathcal{F}_M$,

$$
\begin{aligned}
J_k(u) = \|L_k u - f_k\|^2_{L^2(\Omega)} + \sum_{l=r,R} \Bigg[\|u-g\|^2_{L^2(\partial B_l(0))} + \gamma \bigg(\int_{-\pi}^{\pi} |D_\theta (u\circ \Phi_l)|^2 \, d\theta \\
+ \int_{(-\pi,\pi)^2} \frac{|D_\theta(u\circ\Phi_l)(\theta_1) - D_\theta(u\circ\Phi_l)(\theta_2)|^2}{|\theta_1-\theta_2|^2} \, d\theta_1 d\theta_2 \bigg) \Bigg],
\end{aligned} \tag{4.6.4}
$$

where we define the map $\Phi_l : \theta \in (-\pi,\pi) \to (l\cos(\theta), l\sin(\theta)) \in \partial B_l(0)$ for $l = r, R$. Note that we introduce an extra weighting parameter $\gamma > 0$ in (4.6.4), which helps achieve the optimal balance between the residual of the PDE and the residuals of the boundary data. We set the parameter $\gamma = 0.1$ for all the computations.

The cost functional (4.6.4) is further approximated by an empirical cost via the collocation method (see [74, 16]), where we discretize Ω and $\Theta = (-\pi,\pi)^2$ by sets of collocation points $\Omega_d = \{x_i \in \Omega \mid 1 \le i \le N_d\}$ and $\Theta_d = \{\theta = (\theta_{1,i}, \theta_{2,i}) \in \Theta \mid 1 \le i \le N_b\}$, respectively, and write the discrete form of (4.6.4) as follows: for all $k \in \mathbb{N}$ and $u \in \mathcal{F}_M$,

$$
\begin{aligned}
&J_{k,d}(u) \\
&= \frac{|\Omega|}{N_d} \sum_{x_i \in \Omega_d} |L_k u(x_i) - f_k(x_i)|^2 + \sum_{l=r,R} \Bigg[\frac{|\partial B_l(0)|}{N_b} \sum_{\theta\in\Theta_d} |(u-g)\circ\Phi_l(\theta_{1,i})|^2 \\
&\quad + \gamma \bigg(\frac{2\pi}{N_b} \sum_{\theta\in\Theta_d} |D_\theta(u\circ\Phi_l)|^2(\theta_{1,i}) + \frac{(2\pi)^2}{N_b} \sum_{\theta\in\Theta_d} \frac{|D_\theta(u\circ\Phi_l)(\theta_{1,i}) - D_\theta(u\circ\Phi_l)(\theta_{2,i})|^2}{|\theta_{1,i}-\theta_{2,i}|^2} \bigg) \Bigg],
\end{aligned} \tag{4.6.5}
$$

where $|\Omega| = \pi(R^2 - r^2)$, and $|\partial B_l(0)| = 2\pi l$ for $l = r, R$ are, respectively, the Lebesgue measures of the domain and boundaries. Note that the choice of the smooth activation function $\varrho(y) = \tanh(y)$ implies that every trial function $u \in \mathcal{F}_M$ is smooth, hence all its derivatives are well-defined at any given point. For simplicity, we take the same number of collocation points in the domain and on the boundaries, i.e., $N_d = N_b = N$.

It is clear that the choice of collocation points is crucial for the accuracy and efficiency of the algorithm. Since the total number of points in a regular grid grows exponentially with respect to the dimension, such a construction is infeasible for high-dimensional problems. Moreover, it is well-known that uniformly distributed pseudorandom points in high dimensions tend to cluster on hyperplanes and lead to a suboptimal distribution by relevant measures of uniformity (see e.g. [26, 16]). Therefore, we shall generate collocation points by a quasi-Monte Carlo (QMC) method based on low-discrepancy sequences. In particular, we first define points in $[0,1]^2$ from the generalized Halton sequence (see [38]), and then map those points into the annulus via the polar map $(x, y) \mapsto (l\cos(\psi), l\sin(\psi))$, where $l = \sqrt{(R^2 - r^2)x + r^2}$ and $\psi = 2\pi y$ for all $(x, y) \in [0,1]^2$. The above transformation preserves fractional area, which ensures that a set of well-distributed points on the square will map to a set of points spread evenly over the annulus. We also use Halton points to approximate the (one-dimensional) boundary segments.

Now we are ready to describe the training process, i.e., how to optimize (4.6.5) over all trial functions in $\mathcal{F}_M$. The optimization is performed by using the well-known Adam stochastic gradient descent (SGD) algorithm [69] with a decaying learning rate schedule. At each SGD iteration, we randomly draw a mini-batch of points with size $B = 25$ from the collection of collocation points, and perform gradient descent based on these samples. We initialize the learning rate at 10^{-3} and decrease it by a factor of 0.5 for every 2000 SGD iterations for the examples with analytic solutions in Section 4.6.2, while for the examples without analytic solutions in Section 4.6.3 we decrease the learning rate by a factor of 0.5 once the total number of iterations reaches one of the milestones $2000, 4000, 6000, 10000, 20000, 30000$. Note that compared to the examples in Section 4.6.2, the examples in Section 4.6.3 have less regular solutions, hence one needs to employ neural networks with more complicated architectures to approximate the solutions, and SGD algorithm with more carefully chosen learning rate to fit the networks.

We implement Algorithm 2 using PyTorch and perform all computations on a NVIDIA Tesla K40 GPU with 12 GB memory. The entire algorithm can be briefly summarized as follows. Let $\{\eta_k\}_{k=0}^{\infty}$ be a given sequence, denoting the accuracy requirement for each policy iteration. For each $k \in \mathbb{N} \cup \{0\}$, given the previous iterate u^k, we compute the feedback controls as in (4.6.3) and obtain the controlled coefficients as defined in (4.3.3). Then we apply the SGD method with analytically derived gradient to optimize $J_{k,d}$ over $\mathcal{F}_M$ until we obtain a solution u^{k+1} satisfying $J_{k,d}(u^{k+1}) \leq \eta_k \min(\|u^{k+1} - u^k\|_{2,d}^2, \eta_0)$, where $\|\cdot\|_{2,d}$ denotes the discrete H^2-norm evaluated based on the training samples in Ω_d. We then proceed to the next policy iteration, and terminate Algorithm 2 once the desired accuracy is achieved.

4.6.2 Examples with analytical solutions

In this section, we shall examine the convergence of Algorithm 2 for solving Dirichlet problems of the form (4.6.2) with known solutions. In particular, we shall choose a running cost f such that the analytical solution to (4.6.2) is given by $u^*(x, y) = \sin(\pi r^2/2) - \sin(\pi(x^2 + y^2)/2)$ for all $(x, y) \in \Omega$. To demonstrate the generalizability and the superlinear convergence of the numerical solutions obtained by Algorithm 2, we generate a different set of collocation points in Ω of the size $N_{\text{val}} = 2000$, and use them to estimate the relative error and the q-factor of the numerical solution u^k obtained from the k-th policy iteration for all $k \in \mathbb{N}$:

$$\text{Err}_k = \frac{\|u^k - u^*\|_{2,\Omega,\text{val}}}{\|u^*\|_{2,\Omega,\text{val}}} \quad \text{and} \quad q_k = \frac{\|u^k - u^*\|_{2,\Omega,\text{val}}}{\|u^{k-1} - u^*\|_{2,\Omega,\text{val}}}.$$

We use neural networks with depth $L = 4$ and varying H as trial functions, initialize Algorithm 2 with $u^0 = 0$, and perform experiments with the following model parameters: $a = 0.04$, $\sigma_x = 0.5$, $\sigma_y = 0.2$, $r = 0.5$, $R = \sqrt{2}$, $\kappa = 0.1$ and $v_s = 0.6$.

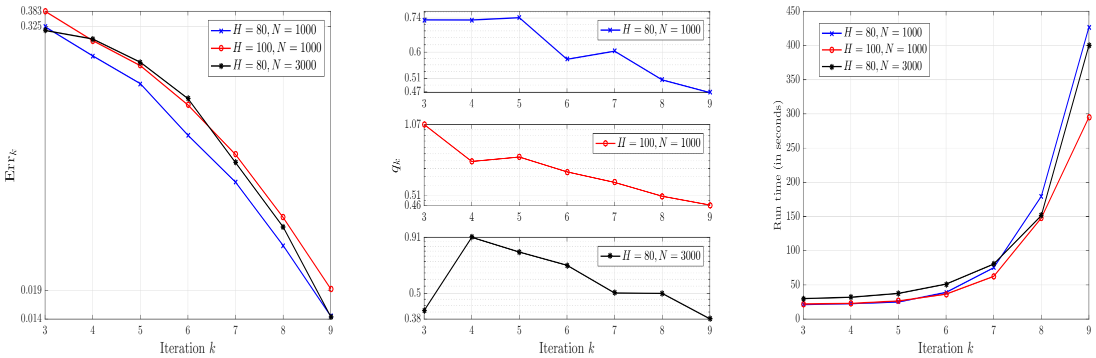


Figure 4.2: Impact of the training sample size N and the hidden width H on the performance of Algorithm 2; from left to right: relative errors (plotted in a log scale), q-factors and the overall runtime for all policy iterations.

Figure 4.2 depicts the performance of Algorithm 2 with different sizes of training samples and the dimensions of hidden layers, which are denoted by N and H respectively. The hyper-parameters $\{\eta_k\}_{k=0}^{\infty}$ are chosen as $\eta_0 = 10$ and $\eta_k = 2^{-k}$ for all $k \in \mathbb{N}$. One can clearly see from Figure 4.2 (left) and (middle) that, despite the fact that Algorithm 2 is initialized with a relatively poor initial guess, the numerical solutions converge superlinearly to the exact solution in the H^2-norm for all these combinations of H and N, which confirms the theoretical result in Theorem 4.4.3. It is interesting to observe from Figure 4.2 that, even though increasing either the complexity of the networks (the red lines) or the size of training samples (the black lines) seems to accelerate the training process slightly (right), neither of them ensures a higher generalization accuracy on the testing samples (left). In all our computations, the accuracies of numerical solutions in the L^2-norm and the H^1-norm are in general higher than the accuracy in the H^2-norm. For example, both the L^2-relative error and the H^1-relative error of the numerical solution obtained at the 9th policy iteration with $H = 80, N = 1000$ are 0.0045.

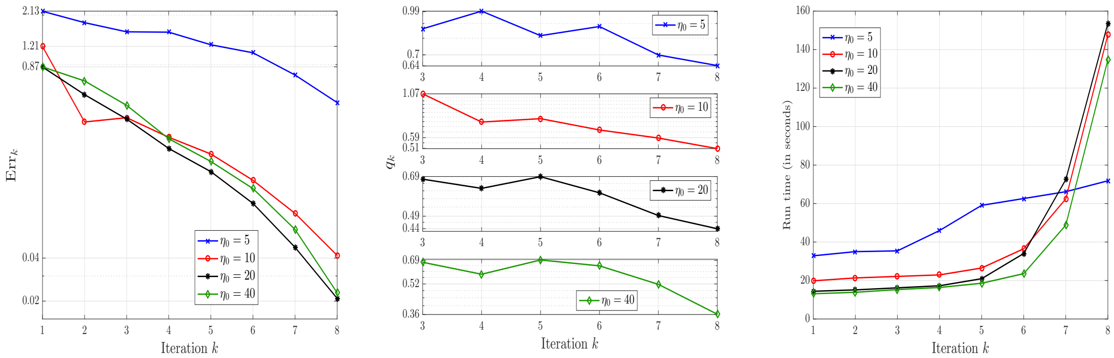


Figure 4.3: Impact of η_0 on the performance of Algorithm 2; from left to right: relative errors (plotted in a log scale), q-factors and the overall runtime for all policy iterations.

We then proceed to analyze the effects of the hyper-parameters $\{\eta_k\}_{k=0}^{\infty}$. Roughly speaking, the magnitude of η_0 indicates the accuracy of the iterates $\{u^k\}_{k=1}^{\infty}$ to the linear Dirichlet problems in the initial stage of Algorithm 2, while the decay of $\{\eta_k\}_{k=1}^{\infty}$ determines the speed at which the q-factors $\{q_k\}_{k=1}^{\infty}$ converge to 0, at an extra cost of solving the optimization problem in a given iteration more accurately for smaller q_k. Figure 4.3 presents the numerical results for different choices of η_0 with a fixed training sample size $N = 1000$, hidden width $H = 100$ and $\eta_k = 2^{-k}$ for all $k \geq 1$. Note that solving each linear equation extremely accurate in the initial stage, i.e., by choosing η_0 to be a small value (the blue line), may not be beneficial for the overall performance of the algorithm in terms of both the accuracy and computational efficiency. This is due to the fact that the initialization of the algorithm is in general far from the exact solution to the semilinear boundary value problem, and so are the solutions of the linear equations arising from the first few policy iterations. In fact, it appears in our experiments that the choices of $\eta_0 = 20, 40$ lead to the optimal performance of Algorithm 2, which solves the initial equations sufficiently accurately, and leverages the superlinear convergence of policy iteration to achieve a higher accuracy with a similar computational cost.

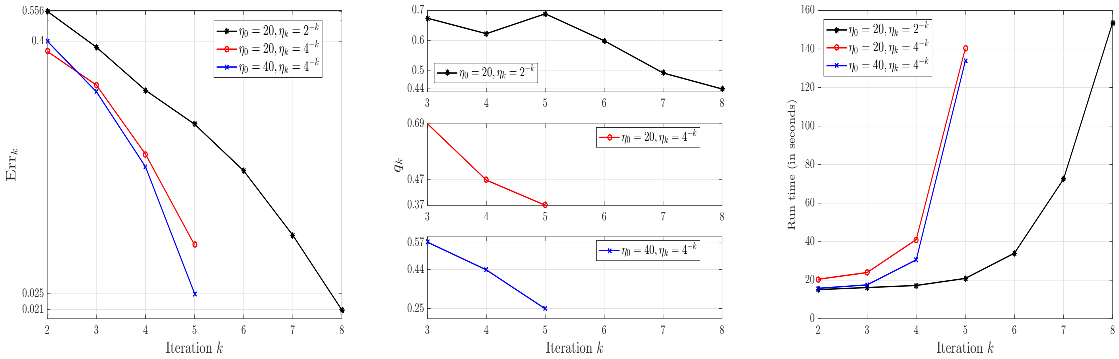


Figure 4.4: Impact of η_k on the performance of Algorithm 2; from left to right: relative errors (plotted in a log scale), q-factors and the overall runtime for all policy iterations.

We further perform computations with different choices of $\{\eta_k\}_{k=1}^{\infty}$ by fixing the training sample size $N = 1000$ and the hidden width $H = 100$. Numerical results are shown in Figure 4.4, from which we can clearly observe that the iterates obtained with $\eta_k = 4^{-k}$, $k \in \mathbb{N}$, converge more rapidly to the exact solution. Note that for $\eta_k = 4^{-k}$, the optimal performance of the algorithm is achieved at $\eta_0 = 40$ instead of $\eta_0 = 20$. This is due to the fact that we solve the first linear Dirichlet problem up to the accuracy $\eta_0\eta_1$ (if we ignore the requirement that $J_0(u^1) \leq \eta_1 \|u^1 - u^0\|^2_{H^2(\Omega)}$ in (4.4.2)), hence one needs to enlarge η_0 for a smaller η_k, such that $\eta_0\eta_1$ is of the same magnitude as before. We observe that the rapid convergence of policy iteration indeed improves the efficiency of the algorithm, in the sense that, to achieve the same accuracy, Algorithm 2 with $\eta_k = 4^{-k}$ requires slightly less computational time than Algorithm 2 with $\eta_k = 2^{-k}$, even though Algorithm 2 with $\eta_k = 4^{-k}$ takes more time to solve the linear equations for each policy iteration; see the last few iterations of the blue line and the black line. This efficiency improvement is more pronounced for the practical problems with complicated solutions in Section 4.6.3; see Figure 4.7 and Table 4.1.

Finally, we shall compare the efficiency of Algorithm 2 (with $\eta_0 = 40$, $\eta_k = 4^{-k}$) to that of the Direct Methods (see e.g. [74, 16, 37, 95]) by fixing the trial functions (4-layer networks with hidden width $H = 100$), the training samples (with size $N = 1000$) and the learning rate of the SGD algorithm. In the Direct Methods, we shall directly apply the SGD method to minimize the following (discretized) squared residual of the semilinear boundary value problem (4.6.2):[2]

$$\|F(u)\|_{0,\Omega,\text{tra}}^2 + \|u-g\|_{X,\partial\Omega,\text{tra}}^2, \tag{4.6.6}$$

where $\|\cdot\|_{0,\Omega,\text{tra}}$ is the discrete L^2 interior norm evaluated from the training samples in Ω, and $\|\cdot\|_{X,\partial\Omega,\text{tra}}$ is a certain discrete boundary norm evaluated from samples on the boundary. In particular, we shall perform computations by setting $\|\cdot\|_{X,\partial\Omega,\text{tra}}^2 = \|\cdot\|_{3/2,\partial\Omega,\text{tra}}^2$ (defined as in (4.6.5)) and $\|\cdot\|_{X,\partial\Omega,\text{tra}}^2 = \vartheta\|\cdot\|_{0,\partial\Omega,\text{tra}}^2$ with different choices of $\vartheta > 0$ ($\vartheta = 1$ in [95] and $\vartheta \in \{500, 1000\}$ in [37]), which will be referred to as "DM with $\|\cdot\|_{3/2,\partial\Omega}^2$" and "DM with $\vartheta\|\cdot\|_{0,\partial\Omega}^2$", respectively, in the following discussion. For both the Direct Methods and Algorithm 2, we shall estimate the H^2-relative error of the numerical solution $\hat{u}_i$ obtained from the i-th SGD iteration by using the same testing samples in Ω of the size $N_{\text{val}} = 2000$ as follows:

$$\text{SGD Err}_i = \|\hat{u}_i - u^*\|_{2,\Omega,\text{val}} / \|u^*\|_{2,\Omega,\text{val}},$$

where u^* denotes the analytical solution to (4.6.2).

Figure 4.5 (left) depicts the H^2-convergence of "DM with $\|\cdot\|_{3/2,\partial\Omega}^2$" and "DM with $\vartheta\|\cdot\|_{0,\partial\Omega}^2$" (with various choices of $\vartheta > 0$) as the number of SGD iterations tends to infinity, which clearly shows that, compared with using the L^2-boundary norm as in [37, 95], incorporating the $H^{3/2}$-boundary norm in the loss function helps achieve a higher H^2-accuracy of the numerical solutions. It is interesting to point out that, even though penalizing the L^2-norm of the boundary term with a suitable parameter ϑ helps improve the accuracy of "DM with $\vartheta\|\cdot\|_{0,\partial\Omega}^2$" as suggested in [37], in our experiments, $\vartheta = 10$ leads to the best H^2-convergence of "DM with $\vartheta\|\cdot\|_{0,\partial\Omega}^2$" (after 10^4 SGD iterations) among other choices of $\vartheta \in \{0.1, 1, 5, 10, 20, 50, 100, 500, 1000\}$.

Figure 4.5 (right) presents the decay of H^2-relative errors with respect to the number of SGD iterations used in "DM with $\|\cdot\|_{0,\partial\Omega}^2$", "DM with $\|\cdot\|_{3/2,\partial\Omega}^2$" and Algorithm 2, which clearly demonstrates that the superlinear convergence of policy iteration significantly accelerates the convergence of the algorithm. In particular, the accuracy enhancement of Algorithm 2 over "DM with $\|\cdot\|_{0,\partial\Omega}^2$" (or equivalently the Deep Galerkin Method proposed in [95]) is of a factor of 20 with 10^4 SGD iterations. We remark that the training time of Algorithm 2 is only slightly longer than that of "DM with $\|\cdot\|_{0,\partial\Omega}^2$" (the runtimes of Algorithm 2 and "DM with $\|\cdot\|_{0,\partial\Omega}^2$" with 10^4

[2] Strictly speaking, the squared residual (4.6.6) is not differentiable (with respect to the network parameters) at the samples where one of the first partial derivatives of the current iterate u is zero, due to the nonsmooth functions $\|\cdot\|_{\ell^1}, \|\cdot\|_{\ell^2} : \mathbb{R}^2 \to [0, \infty)$ in the HJBI operator F (see (4.6.2)). In practice, PyTorch will assign 0 as partial derivatives of $\|\cdot\|_{\ell^1}$ and $\|\cdot\|_{\ell^2}$ functions at their nondifferentiable points, and use it in the backward propagation.

SGD iterations are 333 and 308 seconds, respectively), since Algorithm 2 requires to determine whether a given iterate solves the policy evaluation equations sufficiently accurate (see (4.4.2)), in order to proceed to the next policy iteration step.

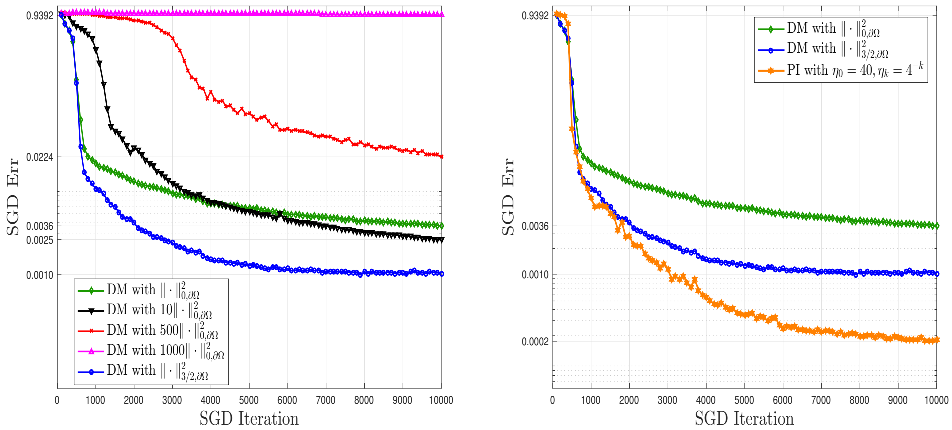


Figure 4.5: Relative errors of the Direct Methods and Algorithm 2 with different numbers of SGD iterations (plotted in a log scale); from left to right: improvements caused by the $H^{3/2}$-boundary norm and by policy iteration.

4.6.3 Examples without analytical solutions

In this section, we shall demonstrate the performance of Algorithm 2 by solving (4.6.2) with $f \equiv 1$. This corresponds to a minimum time problem with preferred targets, whose solution in general is not known analytically. Numerical simulations will be conducted with the following model parameters: $a = 0.2$, $\sigma_x = 0.5$, $\sigma_y = 0.2$, $r = 0.5$, $R = \sqrt{2}$ and $\kappa = 0.1$ but two different values of v_s, $v_s = 0.5$ and $v_s = 1.2$, which are associated with the two scenarios where the ship moves slower than and faster than the wind, respectively. The algorithm is initialized with $u^0 = 0$.

We remark that this is a numerically challenging problem due to the fact that the convection term in (4.6.2) dominates the diffusion term, which leads to a sharp change of the solution and its derivatives near the boundaries. However, as we shall see, these boundary layers can be captured effectively by the numerical solutions of Algorithm 2.

Figure 4.6 presents the numerical results for the two different scenarios obtained by Algorithm 2 with $N = 2000$, $\eta_0 = 40$ and $\eta_k = 1/k$ for all $k \in \mathbb{N}$. The set of trial functions consists of all fully-connected neural networks with depth $L = 7$ and hidden width $H = 50$ (the total number of parameters in this network is 12951).

We can clearly see from Figure 4.6 (left) and (middle) that for both scenarios, the numerical solution $\bar{u}$ and its derivatives are symmetric with respect to the axis $y = 0$, and change rapidly near the boundaries.

The feedback control strategies, computed by (4.6.3), are depicted in Figure 4.6 (right). If the ship starts from the left-hand side and travels toward the inner boundary, then the expected travel time to $\partial B_r(0)$ is around $\frac{R-r}{v_s+v_c}$, which is smaller than the exit cost along $\partial B_R(0)$. Hence the ship would move in the direction of the positive x-axis for both cases, $v_s < v_c$ and $v_s > v_c$. However, the optimal control is different for the two scenarios if the ship is on the right-hand side. For the case where the ship's speed is less than the wind ($v_s = 0.5$), if the ship is closed to $\partial B_r(0)$, then it would move in the direction of the negative x-axis, hoping the random perturbation of the wind would bring it to the preferred target, while if it is far from $\partial B_r(0)$, then it has less chance to reach $\partial B_r(0)$, so it would move along the positive x-axis. On the other hand, for the case where the ship's speed is larger than the wind ($v_s = 1.2$), the ship would in general try to reach the inner boundary. However, if the ship is sufficiently close to $\partial B_R(0)$ in the right-hand half-plane, then the expected travel time to $\partial B_r(0)$ is around $\frac{R-r}{v_s-v_c}$, which is larger than the exit cost along $\partial B_R(0)$. Hence the ship would choose to exit directly from the outer boundary.

We then analyze the convergence of Algorithm 2 by performing computations with 7-layer networks with different hidden width H (networks with wider hidden layers are employed such that every linear Dirichlet problem can be solved more accurately) and parameters $\{\eta_k\}_{k=0}^{\infty}$. For any given iterate u^k, we shall consider the following (squared) residual of the semilinear boundary value problem (4.6.2) :

$$\text{HJBI Residual} := \|F(u^k)\|_{0,\Omega,\text{val}}^2 + \|u^k - g\|_{3/2,\partial\Omega,\text{val}}^2, \tag{4.6.7}$$

which will be evaluated similar to (4.6.5) based on testing samples in Ω and on $(\partial\Omega)^2$ of the same size $N_{\text{val}} = 2000$. Figure 4.7 presents the decay of the residuals in terms of the number of policy iterations, which suggests the H^2-superlinear convergence of the iterates $\{u^k\}_{k=0}^{\infty}$ (the H^2-norms of the last iterates for $v_s = 0.5$ and $v_s = 1.2$ are 20.3 and 31.8, respectively). Note that the parameter $\eta_k = 1/k$, $k \in \mathbb{N}$ leads to a slower and more oscillating convergence of the iterates $\{u^k\}_{k=0}^{\infty}$, due to the fact that we apply a mini-batch SGD method to optimize the discrete cost functional $J_{k,d}$ for each policy iteration. A faster and smoother convergence can be achieved by choosing a more rapidly decaying $\{\eta_k\}_{k=1}^{\infty}$.

We further investigate the influence of the parameters $\{\eta_k\}_{k=1}^{\infty}$ on the accuracy and efficiency of Algorithm 2 in detail. Algorithm 2 is carried out with the same trial

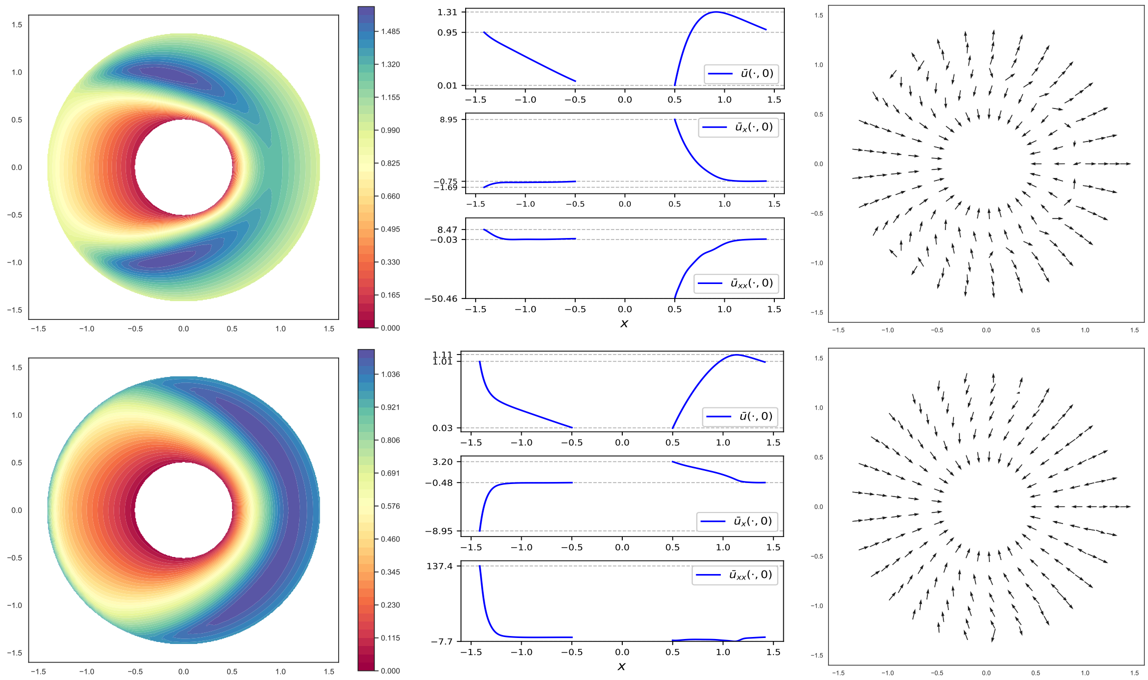


Figure 4.6: Numerical results for the two different scenarios; from top to bottom: the ship moves slower than the wind ($v_s = 0.5$) and faster than the wind ($v_s = 1.2$); from left to right: the value function $\bar{u}$, the numerical solutions along $y = 0$, and feedback control strategies.

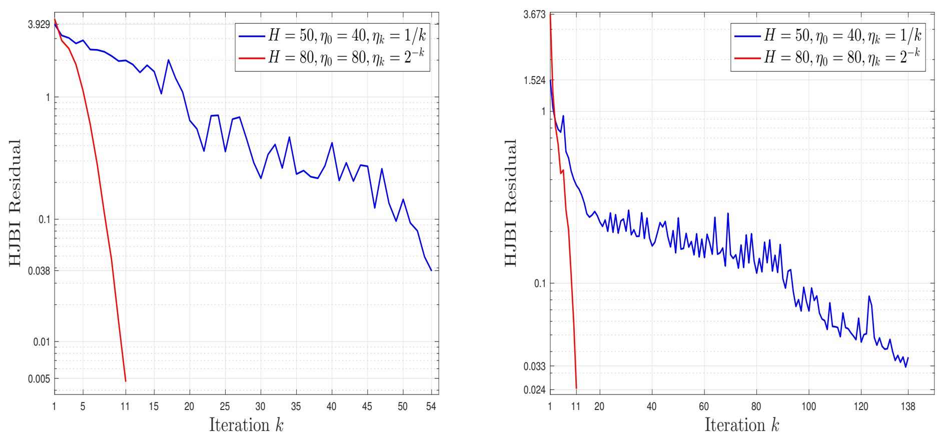


Figure 4.7: Residuals of the HJBI Dirichlet problems for the two different scenarios with respect to the number of policy iterations (plotted in a log scale); from left to right: the ship moves slower than the wind ($v_s = 0.5$) and faster than the wind ($v_s = 1.2$).

functions (7-layer networks with hidden width $H = 80$ and complexity 32721) but different $\{\eta_k\}_{k=1}^{\infty}$ (we choose different η_0 to keep the quantity $\eta_0\eta_1$ constant), and the numerical results are summarized in Table 4.1. One can clearly see that a more rapidly decaying $\{\eta_k\}_{k=1}^{\infty}$ results in a better overall performance (in terms of accuracy and computational time), even though it requires more time to solve the linear equation for each policy iteration. The rapid decay of $\{\eta_k\}_{k=1}^{\infty}$ not only accelerates the superlinear convergence of the iterates $\{u^k\}_{k=1}^{\infty}$, but also helps to eliminate the oscillation caused by the randomness in the SGD algorithm (see Figure 4.7), which enables us to achieve a higher accuracy with less total computational effort. However, we should keep in mind that smaller $\{\eta_k\}_{k=1}^{\infty}$ means that we need to solve all linear equations with higher accuracy, which subsequently requires more complicated networks and more careful choices of the optimizers for $J_{k,d}$. Therefore, in general, we need to tune the balance between the superlinear convergence rate of policy iteration and the computational costs of the linear solvers, in order to achieve optimal performance of the algorithm.

Table 4.1: Numerical results with different parameters $\{\eta_k\}_{k=0}^{\infty}$ for the scenario where the ship is slower than the wind ($v_s = 0.5$).

(a) $H = 80, \eta_0 = 80, \eta_k = 2^{-k}$

PI Itr	HJBI Residual	SGD Itr	Run time
9	0.0466	11030	922s
10	0.0144	20330	1710s
11	0.0046	45510	3820s

(b) $H = 80, \eta_0 = 40, \eta_k = 1/k$

PI Itr	HJBI Residual	SGD Itr	Run time
58	0.0252	32500	2729s
59	0.0201	39080	3275s
60	0.0156	45770	3836s

Finally, we shall compare the performance of Algorithm 2 (with $\eta_0 = 80, \eta_k = 2^{-k}$) and the Direct Method (with $\|\cdot\|_{X,\partial\Omega,\text{tra}} = \|\cdot\|_{3/2,\partial\Omega,\text{tra}}$ in (4.6.6)) by fixing the trial functions (7-layer networks with hidden width $H = 80$), the training samples and the learning rates of the SGD algorithms. For both methods, we shall consider the following squared residual for each iterate $\hat{u}_i$ obtained from the i-th SGD iteration (see (4.6.7)):

$$\text{HJBI Residual} := \|F(\hat{u}_i)\|_{0,\Omega,\text{val}}^2 + \|\hat{u}_i - g\|_{3/2,\partial\Omega,\text{val}}^2.$$

Figure 4.8 (left) presents the decay of the residuals as the number of SGD iterations tends to infinity, which demonstrates the efficiency improvement of Algorithm

2 over the Direct Method. The superlinear convergence of policy iteration helps to provide better initial guesses of the SGD algorithm, which leads to a more rapidly decaying loss curve with smaller noise (on the validation samples); the HJBI residuals obtained in the last 1000 SGD iterations of Algorithm 2 (resp. the Direct Method) oscillates around the value 0.0028 (resp. 0.0144) with a standard derivation 0.00096 (resp. 0.0093).

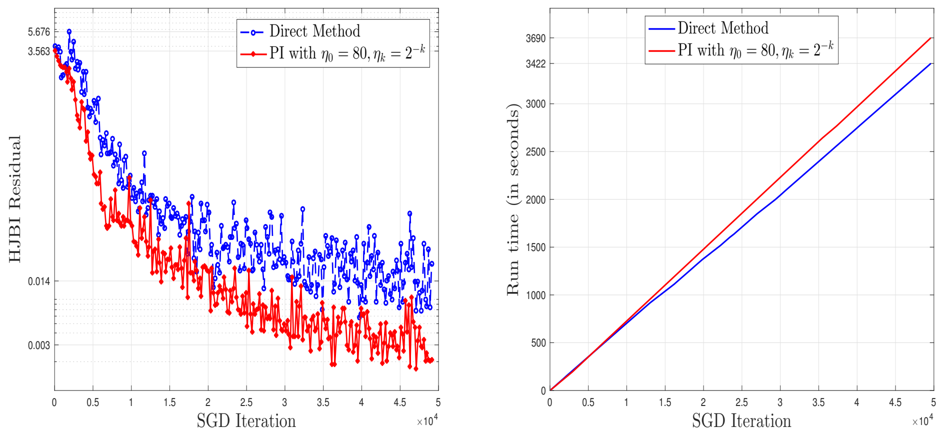


Figure 4.8: Performance comparison of the Direct Method and Algorithm 2 for the scenario where the ship is slower than the wind ($v_s = 0.5$); from left to right: residuals (plotted in a log scale) and overall runtime for all SGD iterations.

Chapter 5

Conclusion and future work

This thesis develops numerical approximations of stochastic hybrid control problems involving impulse controls. In particular, we construct the value functions, the action regions and the optimal impulse controls of hybrid control problems based on a sequence of penalized equations, for which the penalization error is estimated. The penalized equation is then solved by monotone schemes in the low-dimensional setting, and by an efficient neural network based policy iteration algorithm in the high-dimensional setting, whose convergence is proved. Numerical experiments are presented to illustrate the theoretical findings and demonstrate the efficiency improvement of the proposed schemes over existing algorithms.

There are several interesting directions that can be explored in a future work:

(1) In Chapter 4 (see Theorem 4.4.3), we establish the convergence of the neural network based policy iteration algorithm by assuming that at each policy iteration step, one can apply the stochastic gradient descent (SGD) algorithm to minimize the squared residual of a given linear PDE within an arbitrary accuracy. The convergence of the SGD algorithm for solving such a nonlinear optimization problem has recently been analyzed in [62] for linear boundary value problems in the unit cube, and for two-layer neural network ansatzes satisfying boundary conditions exactly. It is interesting to see whether similar techniques can be applied to analyze the performance of the SGD algorithm for solving general second-order PDEs with multi-layer neural network ansatzes.

(2) Optimal control problems with very large populations of interacting individuals and intricate feedback effects (such as crowd control and system risk management) are numerically challenging, even for the neural network based algorithms proposed in Chapter 4. One possible approach to overcome the difficulty is to

consider the corresponding mean field problem (see e.g. [25]), which essentially focuses on the behaviour of a representative agent and describes the inter-player interaction by a nonlinear dependence on the law of the state process. A promising research direction is to extend the techniques in Chapter 4 and to design efficient numerical solvers for mean field control problems.

Appendix A

Supplementary materials for Chapter 3

A.1 Proofs of Lemma 3.2.2 (3), Propositions 3.3.1 and 3.4.1, and Lemmas 3.4.5 and 3.4.9

Proof of Lemma 3.2.2 (3). Let $x^\rho, x \in \mathbb{R}^d$ for all $\rho \in \mathbb{N}$ and $\lim_{\rho\to\infty} x^\rho = x$, we first establish that $\limsup_{\rho\to\infty}(\mathcal{M}_i u^\rho)(x^\rho) \le (\mathcal{M}_i u^*)u(x)$. For any $\varepsilon > 0$, there exists $z^\varepsilon \in Z(x)$, such that $u_i^*(\Gamma_i(x, z^\varepsilon)) + K_i(x, z^\varepsilon) - \varepsilon \le (\mathcal{M}_i u^*)(x)$. Since $Z(x^\rho)$ converges to $Z(x)$ in the Hausdorff metric, we can find $z^{\rho,\varepsilon} \in Z(x^\rho)$, such that $\lim_{\rho\to\infty} z^{\rho,\varepsilon} = z^\varepsilon$. Then we conclude the desired result from the continuity of Γ_i, K_i and the following inequality: for all $\varepsilon > 0$,

$$\begin{aligned}\limsup_{\rho\to\infty}(\mathcal{M}_i u^\rho)(x^\rho) &\le \limsup_{\rho\to\infty}[u_i^\rho(\Gamma_i(x^\rho, z^{\rho,\varepsilon})) + K_i(x^\rho, z^{\rho,\varepsilon})] \le u_i^*(\Gamma_i(x, z^\varepsilon)) + K_i(x, z^\varepsilon)\\ &\le (\mathcal{M}_i u^*)(x) + \varepsilon.\end{aligned}$$

We then show $(\mathcal{M}_i u_*)(x) \le \liminf_{\rho\to\infty}(\mathcal{M}_i u^\rho)(x^\rho)$. For any $\varepsilon > 0$ and $\rho \in \mathbb{N}$, there exists $z^{\rho,\varepsilon} \in Z(x^\rho)$ such that $u_i^\rho(\Gamma_i(x^\rho, z^{\rho,\varepsilon})) + K_i(x^\rho, z^{\rho,\varepsilon}) - \varepsilon \le (\mathcal{M}_i u^\rho)(x^\rho)$. The fact that $Z(x^\rho)$ is convergent to the compact set $Z(x)$ implies that by passing to a subsequence, one can assume $(z^{\rho,\varepsilon})_{\rho\in\mathbb{N}}$ is convergent to some $z^\varepsilon \in Z(x)$. Then we have

$$\begin{aligned}\liminf_{\rho\to\infty}(\mathcal{M}_i u^\rho)(x^\rho) &\ge \liminf_{\rho\to\infty}[u_i^\rho(\Gamma_i(x^\rho, z^{\rho,\varepsilon})) + K_i(x^\rho, z^{\rho,\varepsilon}) - \varepsilon]\\ &\ge (u_*)_i(\Gamma_i(x, z^\varepsilon)) + K_i(x, z^\varepsilon) - \varepsilon\\ &\ge (\mathcal{M}_i u_*)(x) - \varepsilon,\end{aligned}$$

which completes the proof by letting $\varepsilon \to 0$. □

Proof of Proposition 3.3.1. Let u and v be a bounded subsolution and supersolution of (3.3.1) with a fixed penalty parameter $\rho \geq 0$, respectively. We observe that for sufficiently large constant $C > 0$, $w = -C$ is a subsolution to $F_i^\rho(x, w, Dw_i, D^2w_i) \leq -\kappa_0 < 0$, from which by using the fact that F^ρ is convex in u, Du and D^2u, we deduce that $u_m := (1 - \frac{1}{m})u + \frac{1}{m}w$ is a subsolution to $F_i^\rho(x, u_m, D(u_m)_i, D^2(u_m)_i) \leq -\kappa_0/m$ for all $m \in \mathbb{N}$. Note that it suffices to show $u_m - v \leq 0$ for all $m \in \mathbb{N}$, since one can deduce the desired comparison principle $u - v \leq 0$ by letting $m \to \infty$.

Now suppose that there exists $m_0 \in \mathbb{N}$ such that $M = \sup_{x\in\mathbb{R}^d, i\in\mathcal{I}}((u_{m_0})_i - v_i)(x) > 0$, and consider for each $\varepsilon > 0$ the following quantity

$$M_\varepsilon = \sup_{x,y\in\mathbb{R}^d, i\in\mathcal{I}}((u_{m_0})_i(x) - v_i(y) - \frac{1}{2\varepsilon}|x-y|^2). \tag{A.1.1}$$

Then, by assuming without loss of generality that there exists an $i \in \mathcal{I}$, independent of ε, such that the maximum is obtained at the index i and the point $(x^\varepsilon, y^\varepsilon)$ (otherwise one can modify the test function with an additional penalty term), one can deduce from the standard arguments (see [30]) that $\lim_{\varepsilon\to 0} M_\varepsilon = M$ and $\lim_{\varepsilon\to 0} x^\varepsilon = \lim_{\varepsilon\to 0} y^\varepsilon = x_0$ for some x_0. Thus by applying the maximum principle ([30, Theorem 3.2]), we have for any given $\theta > 1$ the matrices $X, Y \in \mathbb{S}^d$ such that $(p_x, X) \in \bar{J}^{2,+}u_m(x^\varepsilon)$ and $(-p_y, -Y) \in \bar{J}^{2,-}v(y^\varepsilon)$, where

$$(p_x, p_y) = \frac{1}{\varepsilon}(x^\varepsilon - y^\varepsilon, y^\varepsilon - x^\varepsilon), \quad \text{and}, \quad \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \leq \theta\frac{1}{\varepsilon}\begin{pmatrix} I & -I \\ -I & I \end{pmatrix},$$

from which, by using the sub- and supersolution properties, we have

$$\begin{aligned} &\sup_{\alpha\in\mathcal{A}_i} \mathcal{L}_i^\alpha(x^\varepsilon, u_{m_0}(x^\varepsilon), p_x, X) - \sup_{\alpha\in\mathcal{A}_i} \mathcal{L}_i^\alpha(y^\varepsilon, v(y^\varepsilon), -p_y, -Y) \\ &\qquad + \rho((u_{m_0})_i - \mathcal{M}_i u_{m_0})^+(x^\varepsilon) - \rho(v_i - \mathcal{M}_i v)^+(y^\varepsilon) + \kappa_0/m_0 \leq 0. \end{aligned} \tag{A.1.2}$$

Now we separate our discussions into two cases. Suppose for all small enough ε, we have

$$\rho((u_{m_0})_i - \mathcal{M}_i u_{m_0})^+(x^\varepsilon) - \rho(v_i - \mathcal{M}_i v)^+(y^\varepsilon) \leq -\kappa_0/m_0,$$

which implies $(v_i - \mathcal{M}_i v)(y^\varepsilon) \geq 0$ and

$$((u_{m_0})_i - \mathcal{M}_i u_{m_0})(x^\varepsilon) - (v_i - \mathcal{M}_i v)(y^\varepsilon) \leq -\kappa_0/(\rho m_0).$$

Then by rearranging the terms in the above inequality and using the definition of M_ε, we have

$$\begin{aligned} M &= \lim_{\varepsilon\to 0} M_\varepsilon = \lim_{\varepsilon\to 0} \left[(u_{m_0})_i(x^\varepsilon) - v_i(y^\varepsilon) - |x^\varepsilon - y^\varepsilon|^2/(2\varepsilon)\right] \\ &\leq \limsup_{\varepsilon\to 0}(\mathcal{M}_i u_{m_0})(x^\varepsilon) - \liminf_{\varepsilon\to 0}(\mathcal{M}_i v)(y^\varepsilon) - \liminf_{\varepsilon\to 0} |x^\varepsilon - y^\varepsilon|^2/(2\varepsilon) - \kappa_0/(\rho m_0) \\ &\leq (\mathcal{M}_i u_{m_0})(x_0) - (\mathcal{M}_i v)(x_0) - \kappa_0/(\rho m_0) \leq M - \kappa_0/(\rho m_0), \end{aligned}$$

where we have used Lemma 3.2.2 (3) and the fact that u_{m_0} and v are upper- and lower-semicontinuous, respectively. This clearly contradicts to the fact that $\kappa_0/(\rho m_0) > 0$.

On the other hand, suppose for all small enough ε, we have

$$\sup_{\alpha\in\mathcal{A}_i} \mathcal{L}_i^\alpha(y^\varepsilon, v(y^\varepsilon), -p_y, -Y) - \sup_{\alpha\in\mathcal{A}_i} \mathcal{L}_i^\alpha(x^\varepsilon, u_{m_0}(x^\varepsilon), p_x, X) \geq 0.$$

This is the classical case (see [57]). In particular, by using the estimate

$$\begin{aligned}
&\sup_{\alpha\in\mathcal{A}_i} \mathcal{L}_i^\alpha(x^\varepsilon, u_{m_0}(x^\varepsilon), p_x, X) - \sup_{\alpha\in\mathcal{A}_i} \mathcal{L}_i^\alpha(x^\varepsilon, v(y^\varepsilon), p_x, X) \\
&\geq \lambda_0((u_{m_0})_i(x^\varepsilon) - v_i(y^\varepsilon)) = \lambda_0\left(M_\varepsilon + \frac{|x^\varepsilon - y^\varepsilon|^2}{2\varepsilon}\right)
\end{aligned}$$

and letting $\varepsilon \to 0$, we can deduce that $M \leq 0$, which is a contradiction. □

Proof of Proposition 3.4.1. We first introduce the solution operator $Q : [C^{0,1}(\mathbb{R}^d)]^M \mapsto [C^{0,1}(\mathbb{R}^d)]^M$ to (3.4.2). That is, for any given u, Qu solves the system of variational inequalities of the form (3.4.2), where the obstacle $\mathcal{M}_i u^{n-1}$ is replaced by $\mathcal{M}_i u$. Then the comparison principle of (3.4.2) and Lemma 3.2.2 (2) imply that Q is monotone: $Qu \geq Qv$ if $u \geq v$. Moreover, one can show Q is concave. In fact, for any given $u, v \in [C^{0,1}(\mathbb{R}^d)]^M$ and $\lambda \in [0, 1]$, we can deduce from Lemma 3.2.2 (1) that for all $i \in \mathcal{I}$,

$$(1-\lambda)(Qu)_i + \lambda(Qv)_i - \mathcal{M}_i[(1-\lambda)u + \lambda v] \leq (1-\lambda)((Qu)_i - \mathcal{M}_i u) + \lambda((Qv)_i - \mathcal{M}_i v). \tag{A.1.3}$$

Moreover, since the HJB equation (3.4.1) is convex in u, Du and D^2u, by applying [10, Lemma A.3] (note the weakly coupled term $\sum_{j\in\mathcal{I}^{-i}} d_{ij}^\alpha u_j$ is linear in u_j, $j \in \mathcal{I}$), we see $(1-\lambda)Qu + \lambda Qv$ is a subsolution to (3.4.2) with an obstacle $\mathcal{M}_i[(1-\lambda)u + \lambda v]$, and consequently conclude the concavity of the operator Q from the comparison principle of (3.4.2).

Now let C be a sufficiently large constant such that $w = (w_i)_{i\in\mathcal{I}}$ with $w_i = -C$ for all $i \in \mathcal{I}$ is a strict subsolution to (3.2.1), that is, $F_i(x, w, Dw_i, D^2 w_i) \leq -\kappa_0$ for all $i \in \mathcal{I}$. We proceed to establish a contractive property of the iterates $(u^n)_{n\in\mathbb{N}}$, where u^n is a viscosity solution to (3.4.2) for each n. By using the monotonicity and concavity of the operator Q, we can show that if $u^{n-1} - u^n \leq \lambda(u^{n-1} - w)$ for some $\lambda \in [0, 1]$ and $n \in \mathbb{N}$, then it holds for any constants $C \geq |(u^n)^+|_0 + |w|_0$ and $0 < \mu \leq \min(1, \kappa_0/C)$ that $u^n - u^{n+1} \leq \lambda(1-\mu)(u^n - w)$ (cf. [88, Lemma 3.3]). Since $w \leq u^n \leq u^0$ for all n and w is bounded, there exists a constant $\mu \in (0, 1]$ such that $0 \leq u^{n-1} - u^n \leq (1-\mu)^{n-1}(u^0 - w)$ for all $n \geq 0$. Consequently we can show $(u^n)_{n\geq 0}$ converges uniformly to some continuous function u, which is the unique viscosity solution to (3.2.1). Then the contractive property enables us to conclude the desired error estimate. □

Proof of Lemma 3.4.5. For $\delta,\gamma>0$, we define for all $x,y\in\mathbb{R}^d$ that

$$\Phi_i(x,y)=(Q^\rho u)_i(x)-(Q^\rho v)_i(y)-\phi(x,y),\quad \phi(x,y)=\delta|x-y|^2+\gamma|x|^2,$$

and let $\Phi_i(\bar{x},y)=m_{\delta,\gamma}:=\sup_{i,x,y}\Phi_i(x,y)$ for some $(\bar{x},\bar{y})\in\mathbb{R}^{2d}$ and $i\in\mathcal{I}$, where we omit the dependence on δ,γ for notational simplicity. Since $\mathcal{I}$ is a finite set, we shall assume without loss of generality that the index i is independent of δ,γ. Then for any $\theta>1$, we deduce from the maximum principle [30, Theorem 3.2] that for any $\theta>1$, we have

$$\begin{aligned}\sup_{\alpha\in\mathcal{A}_i}&\mathcal{L}_i^\alpha(\bar{x},(Q^\rho u)(\bar{x}),p_x,X)-\sup_{\alpha\in\mathcal{A}_i}\mathcal{L}_i^\alpha(\bar{y},(Q^\rho v)(\bar{y}),-p_y,-Y)\\&+\rho((Q^\rho u)_i-\mathcal{M}_iu)^+(\bar{x})-\rho((Q^\rho v)_i-\mathcal{M}_iv)^+(\bar{y})\leq 0,\end{aligned}$$

where $(p_x,p_y)=(D_x\phi(\bar{x},\bar{y}),D_y\phi(\bar{x},\bar{y}))$, and $\begin{pmatrix}X&0\\0&Y\end{pmatrix}\leq\theta D^2\phi(\bar{x},\bar{y})$.

We now discuss two cases. Suppose $((Q^\rho u)_i-\mathcal{M}_iu)^+(\bar{x})-((Q^\rho v)_i-\mathcal{M}_iv)^+(\bar{y})<0$, then we have $((Q^\rho u)_i-\mathcal{M}_iu)(\bar{x})\leq((Q^\rho v)_i-\mathcal{M}_iv)(\bar{y})$, and consequently

$$\begin{aligned}(Q^\rho u)_i(\bar{x})-(Q^\rho v)_i(\bar{y})&\leq(\mathcal{M}_iu)(\bar{x})-(\mathcal{M}_iv)(\bar{x})+(\mathcal{M}_iv)(\bar{x})-(\mathcal{M}_iv)(\bar{y})\\&\leq|(u_i-v_i)^+|_0+[\mathcal{M}_iv]_1|\bar{x}-\bar{y}|,\end{aligned}$$

where we used the definition (3.2.3) of $\mathcal{M}_i$. This implies that

$$m_{\delta,\gamma}\leq|(u_i-v_i)^+|_0+[\mathcal{M}_iv]_1|\bar{x}-\bar{y}|-\delta|\bar{x}-\bar{y}|^2\leq|(u_i-v_i)^+|_0+[\mathcal{M}_iv]_1^2/(4\delta).$$

Then, by passing $\gamma\to 0$, we deduce for any $x,y\in\mathbb{R}^d$ and $\delta>0$ that,

$$(Q^\rho u)_i(x)-(Q^\rho v)_i(y)\leq|(u_i-v_i)^+|_0+[\mathcal{M}_iv]_1^2/(4\delta)+\delta|x-y|^2,$$

which, along with the assumption $[\mathcal{M}_iv]_1\leq[v]_1+C$, leads to the desired conclusion by minimizing over $\delta>0$ and then setting $x=y$.

On the other hand, if $((Q^\rho u)_i-u_j-k_{ij})^+(\bar{x})-((Q^\rho v)_i-v_j-k_{ij})^+(\bar{y})\geq 0$, then the classical results for weakly coupled system gives us that $(Q^\rho u)_i\leq(Q^\rho v)_i$ (see e.g. [57]). □

Proof of Lemma 3.4.9. Note that for any given $\alpha > 0$, $\mu \in (0,1)$ and $\gamma \in \mathbb{N}$, we have $(\phi^\alpha)' = \alpha\gamma x^{\gamma-1} + \mu^x \log \mu$, which is increasing on $(0, \infty)$. Suppose that α is sufficiently small such that $\alpha\gamma < -\log \mu$, then we can show $(\phi^\alpha)'(n^\alpha) \geq 0$, with the natural number n^α defined as:

$$n^\alpha := \left\lceil \frac{\log(-\alpha\gamma/\log(\mu))}{\log \mu} \right\rceil \leq \frac{\log(-\alpha\gamma/\log \mu)}{\log \mu} + 1.$$

Consequently, ϕ^α is increasing on (n^α, ∞), which leads to the estimate that for all small enough α,

$$m^\alpha \leq \phi^\alpha(n^\alpha) \leq \alpha\left(\frac{\log(-\alpha\gamma/\log \mu)}{\log \mu} + 1\right)^\gamma + \mu\frac{-\alpha\gamma}{\log \mu} \leq C\alpha(-\log \alpha)^\gamma,$$

where the constant C depends only on γ and μ. $\square$

Appendix B

Supplementary materials for Chapter 4

B.1 Some fundamental results

Here, we collect some well-known results which are used frequently in Chapter 4.

We start with the well-posedness of strong solutions to Dirichlet boundary value problems. In the sequel, we shall denote by τ the trace operator.

Theorem B.1.1. *([42, Theorem 1.2.19]) Let Ω be a bounded $C^{1,1}$ domain. Suppose that for all $1 \le i,j \le n$, a^{ij} is in $C(\bar{\Omega})$, and b^i, c are in $L^\infty(\Omega)$, satisfying $c \ge 0$ and*

$$\sum_{i,j=1}^{n} a^{ij}(x)\xi_i\xi_j \ge \lambda|\xi|^2, \quad \text{for all } \xi \in \mathbb{R}^n \text{ and for almost every } x \in \Omega, \tag{B.1.1}$$

for some constant $\lambda > 0$. Then for every $f \in L^2(\Omega)$ and $g \in H^{3/2}(\partial\Omega)$, there exists a unique strong solution $u \in H^2(\Omega)$ to the Dirichlet problem

$$-a^{ij}\partial_{ij}u + b^i\partial_i u + cu = f, \text{ in } \Omega; \quad \tau u = g, \text{ on } \partial\Omega,$$

and the following estimate holds with a constant C independent of f and g:

$$\|u\|_{H^2(\Omega)} \le C\big(\|f\|_{L^2(\Omega)} + \|g\|_{H^{3/2}(\partial\Omega)}\big).$$

The next theorem shows the well-posedness of oblique boundary value problems.

Theorem B.1.2. *([42, Theorem 1.2.20]) Let Ω be a bounded $C^{1,1}$ domain. Suppose that for all $1 \le i,j \le n$, a^{ij} is in $C(\bar{\Omega})$, and b^i, c are in $L^\infty(\Omega)$, satisfying $c \ge 0$ and the uniform elliptic condition (B.1.1) for some constant $\lambda > 0$.*

Assume in addition, that $\{\gamma^j\}_{j=0}^n \subseteq C^{0,1}(\partial\Omega)$, $\gamma^0 \ge 0$ on $\partial\Omega$, $\operatorname{ess\,sup}_\Omega c + \max_{\partial\Omega}\gamma^0 > 0$, and $\sum_{j=1}^n \gamma^j\nu_j \ge \mu$ on $\partial\Omega$ for some constant $\mu > 0$, where $\{\nu_j\}_{j=1}^n$ are the components of the unit outer normal vector field on $\partial\Omega$. Then for every $f \in L^2(\Omega)$ and

$g \in H^{1/2}(\partial\Omega)$, there exists a unique strong solution $u \in H^2(\Omega)$ to the following oblique derivative problem:

$$-a^{ij}\partial_{ij}u + b^i\partial_i u + cu = f, \text{ in } \Omega; \quad \gamma^j\tau(\partial_j u) + \gamma^0\tau u = g, \text{ on } \partial\Omega,$$

and the following estimate holds with a constant C independent of f and g:

$$\|u\|_{H^2(\Omega)} \leq C\Big(\|f\|_{L^2(\Omega)} + \|g\|_{H^{1/2}(\partial\Omega)}\Big).$$

We then recall several important measurability results. The following measurable selection theorem follows from Theorems 18.10 and 18.19 in [1], and ensures the existence of a measurable selector maximizing (or minimizing) a Carathéodory function.

Theorem B.1.3. *Let (S, Σ) a measurable space and X be a separable metrizable space. Let $\Gamma : S \rightrightarrows X$ be a measurable set-valued mapping with nonempty compact values, and suppose $g : S \times X \to \mathbb{R}$ is a Carathéodory function. Define the value function $m : S \to \mathbb{R}$ by $m(s) = \max_{x \in \Gamma(s)} g(s, x)$, and the set-valued map $\mu : S \rightrightarrows X$ by $\mu(s) = \{x \in \Gamma(s) \mid g(s, x) = m(s)\}$. Then we have*

1. *The value function m is measurable.*
2. *The set-valued mapping μ is measurable, has nonempty and compact values. Moreover, there exists a measurable function $\psi : S \to X$ satisfying $\psi(s) \in \mu(s)$ for each $s \in S$.*

The following theorem shows the arg max set-valued mapping is upper hemicontinuous.

Theorem B.1.4. *([1, Theorem 17.31]) Let X, Y be topological spaces, $\Gamma \subset Y$ be a nonempty compact subset, and $g : X \times \Gamma \to \mathbb{R}$ be a continuous function. Define the value function $m : X \to \mathbb{R}$ by $m(x) = \max_{y \in \Gamma} g(x, y)$, and the set-valued map $\mu : X \rightrightarrows Y$ by $\mu(x) = \{y \in \Gamma \mid g(x, y) = m(x)\}$. Then μ has nonempty and compact values. Moreover, if Y is Hausdorff, then μ is upper hemicontinuous, i.e., for every $x \in X$ and every neighborhood U of $\mu(x)$, there is a neighborhood V of x such that $z \in V$ implies $\mu(z) \subset U$.*

Finally, we present a special case of [35, Theorem 2], which characterizes q-superlinear convergence of quasi-Newton methods for a class of semismooth operator-valued equations.

Theorem B.1.5. *Let Y, Z be two Banach spaces, and $F : Y \to Z$ be a given function with a zero $y^* \in Y$. Suppose there exists an open neighborhood V of y^* such that F is semismooth with a generalized differential $\partial^* F$ in V, and there exists a constant $L > 0$ such that*

$$\|y - y^*\|_Y / L \leq \|F(y) - F(y^*)\|_Z \leq L\|y - y^*\|_Y, \quad \forall y \in V.$$

For some starting point y^0 in V, let the sequence $\{y^k\}_{k\in\mathbb{N}} \subset V$ satisfy $y^k \neq y^*$ for all k, and be generated by the following quasi-Newton method:

$$B_k s^k = -F(y^k), \quad y^{k+1} = s^k + y^k, \quad k = 0, 1, \ldots$$

where $\{B_k\}_{k\in\mathbb{N}}$ is a sequence of bounded linear operators in $\mathcal{L}(Y, Z)$. Let $\{A_k\}_{k\in\mathbb{N}}$ be a sequence of generalized differentials of F such that $A_k \in \partial^* F(y^k)$ for all k, and let $E_k = B_k - A_k$. Then $y^k \to y^*$ q-superlinearly if and only if $\lim_{k\to\infty} y^k = y^*$ and $\lim_{k\to\infty} \|E_k s^k\|_Z / \|s^k\|_Y = 0$.

B.2 Proof of Theorem 4.5.2

Let $u^0 \in \mathcal{F}$ be an arbitrary initial guess, we shall assume without loss of generality that Algorithm 3 runs infinitely, i.e., $\|u^{k+1} - u^k\|_{H^2(\Omega)} > 0$ and $u^k \neq u^*$ for all $k \in \mathbb{N} \cup \{0\}$.

We first show $\{u^k\}_{k\in\mathbb{N}}$ converges to the unique solution u^* in $H^2(\Omega)$. For each $k \geq 0$, we can deduce from (4.5.3) that there exists $f_k^e \in L^2(\Omega)$ and $g_k^e \in H^{1/2}(\partial\Omega)$ such that

$$L_k u^{k+1} - f_k = f_k^e, \quad \text{in } \Omega; \quad B u^{k+1} = g_k^e, \quad \text{on } \partial\Omega, \tag{B.2.1}$$

and $\|f_k^e\|^2_{L^2(\Omega)} + \|g_k^e\|^2_{H^{1/2}(\Omega)} \leq \eta_{k+1}(\|u^{k+1} - u^k\|^2_{H^2(\Omega)})$ with $\lim_{k\to\infty} \eta_k = 0$. Then, we can proceed as in the proof of Theorem 4.4.3, and conclude that if $c \geq \underline{c}_0$ with a sufficiently large $\underline{c}_0$, then $\{u^k\}_{k\in\mathbb{N}}$ converges to the solution u^* of (4.5.1).

The q-superlinear convergence of Algorithm 3 can then be deduced by interpreting the algorithm as a quasi-Newton method for the operator equation $\bar{F}(u) = 0$, with the operator $\bar{F} : u \in H^2(\Omega) \to (F(u), Bu) \in Z$, where we introduce the Banach space $Z := L^2(\Omega) \times H^{1/2}(\partial\Omega)$ with the usual product norm $\|z\|_Z := \|z_1\|_{L^2(\Omega)} + \|z_2\|_{H^{1/2}(\partial\Omega)}$ for each $z = (z_1, z_2) \in Z$. Since $B \in \mathcal{L}(H^2(\Omega), H^{1/2}(\partial\Omega))$, we can directly infer from Corollary 4.3.5 that $\bar{F} : H^2(\Omega) \to Z$ is semismooth in $H^2(\Omega)$, with a generalized differential $M_k = (L_k, \gamma^i \tau(\partial_i) + \gamma^0 \tau) \in \partial^* \tilde{F}(u^k) \subset \mathcal{L}(H^2(\Omega), Z)$ for all $k \in \mathbb{N} \cup \{0\}$. Then, for each $k \geq 0$, by following the same arguments as in Theorem

4.4.3, we can construct a perturbed operator $\delta M_k \in \mathcal{L}(H^2(\Omega), Z)$, such that (B.2.1) can be equivalently written as $(M_k + \delta M_k)s_k = -\bar{F}(u^k)$ with $s_k = u^{k+1} - u^k$, and $\|\delta M_k s_k\|/\|s^k\|_{H^2(\Omega)} \leq \sqrt{2\eta_0\eta_{k+1}} \to 0$, as $k \to \infty$. Finally, the regularity theory of elliptic oblique derivative problems (see Theorem B.1.2) shows that M_k is nonsingular for each k, and $\|M_k^{-1}\|_{\mathcal{L}(Z,H^2(\Omega))} \leq C$ for some constant C independent of k. Hence we can verify that there exists a neighborhood V of u^* and a constant $L > 0$, such that

$$\|u - u^*\|_{H^2(\Omega)}/L \leq \|\bar{F}(u) - \bar{F}(u^*)\|_Z \leq L\|u - u^*\|_{H^2(\Omega)}, \quad \forall u \in V,$$

which allows us to conclude from Theorem B.1.5 the q-superlinear convergence of $\{u^k\}_{k\in\mathbb{N}}$.

Bibliography

[1] C. D. Aliprantis and K. C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, 3rd ed., Springer-Verlag, Berlin, 2006.

[2] A. Alla, M. Falcone, and D. Kalise, *An efficient policy iteration algorithm for dynamic programming equations*, SIAM J. Sci. Comput., 37 (2015), pp. A181–A200,

[3] J.-P. Aubin and H. Frankowska, *Set-Valued Analysis*, Birkhäuser, Basel, 1990.

[4] P. Azimzadeh, *A zero-sum stochastic differential game with impulses, precommitment, and unrestricted cost functions*, Appl. Math. Optim., to appear.

[5] P. Azimzadeh, E. Bayraktar, and G. Labahn, *Convergence of implicit schemes for Hamilton-Jacobi-Bellman quasi-variational inequalities*, SIAM J. Control Optim., 56 (2018), pp. 3994–4016.

[6] P. Azimzadeh and P. A. Forsyth, *Weakly chained matrices, policy iteration, and impulse control*, SIAM J. Numer. Anal., 54 (2016), pp. 1341–1364.

[7] J. Babbin, P. A. Forsyth and G. Labahn, *A comparison of iterated optimal stopping and local policy iteration for American options under regime switching*, J. Sci. Comput., 58 (2014), pp. 409–430.

[8] L. Bai and J. Paulsen, *Optimal dividend policies with transaction costs for a class of diffusion processes*, SIAM J. Control Optim., 48 (2010), pp. 4987–5008.

[9] M. Bardi and I. Capuzzo-Dolcetta, *Optimal Control and Viscosity Solutions of HamiltonJacobi-Bellman Equations*, Systems Control Found. Appl., Birkhäuser Boston, Boston, 1997.

[10] G. Barles and E. R. Jakobsen, *On the convergence rate of approximation schemes for Hamilton-Jacobi-Bellman equations*, M2AN Math. Model. Numer. Anal., 36 (2002), pp. 33–54.

[11] G. Barles and E. R. Jakobsen, *Error bounds for monotone approximation schemes for Hamilton-Jacobi-Bellman equations*, SIAM J. Numer. Anal., 43 (2005), pp. 540–558.

[12] R. W. Beard and T. W. Mclain, *Successive Galerkin approximation algorithms for nonlinear optimal and robust control*, Internat. J. Control, (71) 1998, pp. 717–743.

[13] R. W. Beard, G. N. Saridis, and J. T. Wen, *Galerkin approximation of the Generalized Hamilton-Jacobi-Bellman equation*, Automatica, (33) 1997, pp. 2159–2177.

[14] A. Bensoussan and J. L. Lions, *Contrôle impulsionnel et inéquations quasi variationnelles*, vol. 11 of Méthodes Mathématiques de l'Informatique [Mathematical Methods of Information Science], Gauthier-Villars, Paris, 1982.

[15] A. Bensoussan and J.L. Menaldi, *Hybrid control and dynamic programming*, Dynam. Contin. Discrete Impuls. Systems, 3 (1997), pp. 395–442.

[16] J. Berg and K. Nyström, *A unified deep artificial neural network approach to partial differential equations in complex geometries*, Neurocomputing, (317) 2018, pp. 28–41.

[17] P. B. Bochev and M. D. Gunzburger, *Least-Squares Finite Element Methods*, Springer, New York, 2009.

[18] O. Bokanowski, B. Bruder, S. Maroso, and H. Zidani, *Numerical approximation for a superreplication problem under gamma constraints*, SIAM J. Numer. Anal., 47 (2009), pp. 2289–2320,

[19] O. Bokanowski, S. Maroso and H. Zidani, *Some convergence results for Howard's algorithm*, SIAM J. Numer. Anal., 47 (2009), pp. 3001–3026.

[20] F. Bonnans, S. Maroso and H. Zidani, *Error estimates for a stochastic impulse control problem*, Appl. Math. Optim., 55 (2007), pp. 327–357.

[21] M. Boulbrachene and M. Haiour, *The finite element approximation of Hamilton-Jacobi-Bellman equations*, Comput. Math. Applic., 41 (2001), 993–1007.

[22] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag, New York, 1994.

[23] A. Briani, F. Camilli, H. Zidani, *Approximation schemes for monotone systems of nonlinear second order differential equations: convergence result and error estimate*, Diff. Equ. Appl., 4 (2012), pp. 297–317.

[24] R. Buckdahn and T. Y. Nie, *Generalized Hamilton–Jacobi–Bellman equations with Dirichlet boundary condition and stochastic exit time optimal control problem*, SIAM J. Control Optim., 54 (2016), pp. 602–631.

[25] R. Carmona and F. Delarue, *Probabilistic theory of mean field games with applications I: Mean-field FBSDEs, control, and games*, Springer International Publishing, Switzerland, 2018.

[26] C. Cervellera and M. Muselli, *Deterministic design for neural network learning: An approach based on discrepancy*, IEEE Trans. Neural Networks, 15 (2004), pp. 533–544.

[27] J. P. Chancelier, M. Messaoud, and A. Sulem, *A policy iteration algorithm for fixed point problems with nonexpansive operators*, Math. Methods Oper. Res., 65 (2007), pp. 239–259.

[28] X. Chen, Z. Nashed and L. Qi, *Smoothing methods and semismooth methods for nondifferentiable operator equations*, SIAM J. Numer. Anal., 38 (2000), pp. 1200–1216.

[29] K.C. Cheung, L. Ling, R. Schaback, *H^2-convergence of least-squares kernel collocation methods*, SIAM J. Numer. Anal. 56 (2018) 614–633.

[30] M. G. Crandall, H. Ishii, and P.-L. Lions, *User's guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc. (N.S.), 27 (1992), pp. 1–67.

[31] M. H. A. Davis, X. Guo, and G. Wu, *Impulse controls of multidimensional jump diffusions*, SIAM J. Control Optim., 48 (2010), pp. 5276–5293.

[32] K. Debrabant and E. R. Jakobsen, *Semi-Lagrangian schemes for linear and fully nonlinear diffusion equations*, Math. Comp., 82 (2012), pp. 1433–1462.

[33] B. Djehiche, S. Hamadene, M. A. Morlais and X. Zhao, *On the equality of solutions of max-min and min-max systems of variational inequalities with interconnected bilateral obstacles*. J. Math. Anal. Appl., 452 (2017), pp. 148–175.

[34] E. Dockner, S. Jorgensen, N. Van Long and G. Sorger, *Differential Games in Economics and Management Science*, Cambridge University Press, Cambridge, UK, 2000.

[35] A. L. Dontchev, *Generalizations of the Dennis-Moré theorem*, SIAM J. Optim., 22 (2012), pp. 821–830.

[36] W. E, J. Han, and A. Jentzen, *Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations*, Commun. Math. Stat., 5 (2017), pp. 349–380.

[37] W. E and B. Yu, *The deep Ritz method: a deep learning-based numerical algorithm for solving variational problems*, Commun. Math. Stat., 6 (2018), pp. 1–12.

[38] H. Faure and C. Lemieux, *Generalized Halton sequence in 2008: a comparative study*, ACM Trans. Model. Comput. Simul., 19 (2009).

[39] R. Ferretti, A. Sassi and H. Zidani, *Error estimates for numerical approximation of Hamilton-Jacobi equations related to hybrid control systems*, Appl. Math. Optim., to appear.

[40] P. A. Forsyth and G. Labahn, *Numerical methods for controlled Hamilton-Jacobi-Bellman PDEs in finance*, J. Computational Finance, 11 (2007/2008, Winter), pp. 1–43.

[41] P. A. Forsyth and K. R. Vetzal, *Quadratic convergence for valuing American options using a penalty method*, SIAM J. Sci. Comput., 23 (2002), pp. 2095–2122.

[42] M. G. Garroni and J. L. Menaldi, *Second order elliptic integro-differential problems*, Chapman & Hall/CRC, Boca Raton, FL, 2002.

[43] D. Gilbarg and N. Trudinger, *Elliptic Partial Differential Equations of Second Order*, 2nd edition, Springer-Verlag, Berlin, New York, 1983.

[44] E. Grisvard, *Elliptic problems in nonsmooth domains*, Pitman, Boston, MA, 1985.

[45] X. Guo and G. L. Wu, *Smooth fit principle for impulse control of multidimensional diffusion processes*, SIAM J. Control Optim., 48 (2009), pp. 594–617.

[46] J. Han and W. E, *Deep learning approximation for stochastic control problems*, preprint, arXiv:1611.07422, 2016.

[47] J. Han and J. Long, *Convergence of the deep BSDE method for coupled FBSDEs*, preprint, arXiv:1811.01165v1, 2018.

[48] M. Hintermüller, *Mesh-independence and fast local convergence of a primal-dual active set method for mixed control-state constrained elliptic problems*, ANZIAM J., 49 (2007), pp. 1–38.

[49] M. Hintermüller, K. Ito, and K. Kunisch, *The primal-dual active set strategy as a semismooth Newton method*, SIAM J. Optim., 13 (2002), pp. 865–888.

[50] K. Hornik, M. Stinchcombe, and H. White, *Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks*, Neural Networks, 3 (1990), pp. 551–560.

[51] C. C. Huang and S. Wang, *A power penalty approach to a nonlinear complementarity problem*, Oper. Res. Lett., 38 (2010), pp. 72–76.

[52] C. Huré, H. Pham, A. Bachouch, and N. Langrené, *Deep neural networks algorithms for stochastic control problems on finite horizon, part I: convergence analysis*, preprint, arXiv:1812.04300, 2018.

[53] H. Ishii, *On the equivalence of two notions of weak solutions, viscosity solutions and distribution solutions*, Funkcial. Ekvac., 38 (1995), pp. 101–120.

[54] K. Ishii, *Viscosity solutions of nonlinear second order elliptic PDEs associated with impulse control problems*, Funkcial. Ekvac., 36 (1993), pp. 123–141.

[55] H. Ishii and P. L. Lions, *Viscosity solutions of fully nonlinear second-order elliptic partial differential equations*, J. Differential Equations, 83 (1990), pp. 26–78.

[56] H. Ishii and S. Koike, *Viscosity solutions of a system of nonlinear second-order elliptic PDEs arising in switching games*, Funkcial. Ekvac., 34 (1991), pp. 143–155.

[57] H. Ishii and S. Koike, *Viscosity solutions for monotone systems of second-order elliptic PDEs*, Comm. Partial Differential Equations, 16 (1991), pp. 1095–1128.

[58] K. Ito and K. Kunisch, *Semi-smooth Newton methods for variational inequalities of the first kind*, M2AN Math. Model. Numer. Anal., 37 (2003), pp. 41–62.

[59] K. Ito and K. Kunisch, *Parabolic variational inequalities: The Lagrange multiplier approach*, J. Math. Pures Appl., 85 (2006), pp. 415–449.

[60] K. Ito, C. Reisinger, and Y. Zhang, *A neural network based policy iteration algorithm with global H^2-superlinear convergence for stochastic games on domains*, Found. Comput. Math., (2020), Available at arXiv:1906.02304v3.

[61] K. Ito and J. Zou, *Identification of some source densities of the distribution type*, J. Comput. Appl. Math., 132 (2001), pp. 295–308.

[62] T. Luo and H. Yang, *Two-layer neural networks for partial differential equations: Optimization and generalization theory*, preprint, arXiv:2006.15733, (2020).

[63] E. R. Jakobsen, *On the rate of convergence of approximation schemes for Bellman equations associated with optimal stopping time problems*, Math. Models Methods Appl. Sci., 13 (2003), pp. 613–644.

[64] E. R. Jakobsen, *On error bounds for monotone approximation schemes for multi-dimensional Isaacs equations*, Asymptot. Anal., 49 (2006), pp. 249–273.

[65] D. Kalise and K. Kunisch, *Polynomial approximation of high-dimensional Hamilton-Jacobi-Bellman equations and applications to feedback control of semi-linear parabolic PDEs*, SIAM J. Sci. Comput., 40 (2018), A629–A652.

[66] D. Kalise, S. Kundu, and K. Kunisch, *Robust feedback control of nonlinear PDEs by numerical approximation of high-dimensional Hamilton-Jacobi-Isaacs equations*, preprint, arXiv:1905.06276, 2019.

[67] B. Kerimkulov, D. Šiška, and Ł. Szpruch, *Exponential convergence and stability of Howard's policy improvement algorithm for controlled diffusions*, preprint, arXiv:1812.07846, 2018.

[68] I. Kharroubi, J. Ma, H. Pham, and J. Zhang, *Backward SDEs with contrained jumps and quasi-variational inequalities*, Ann. Probab., 38 (2010), pp. 794–840.

[69] D. P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, CoRR preprint, https://arxiv.org/abs/1412.6980, 2014.

[70] R. Korn, *Some applications of impulse control in mathematical finance*, Math. Methods Oper. Res., 50 (1999), pp. 493–518.

[71] N. V. Krylov, *On the dynamic programming principle for uniformly nondegenerate stochastic differential games in domains and the Isaacs equations*, Probab. Theory Related Fields, 158 (2014), pp. 751–783.

[72] N.V. Krylov, *Sobolev and Viscosity Solutions for Fully Nonlinear Elliptic and Parabolic Equations*, Mathematical Surveys and Monographs, 233, Amer. Math. Soc., Providence, RI, 2018.

[73] H. Kushner and P. Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer-Verlag, New York, 1991.

[74] E. Lagaris, A. Likas, and D. I. Fotiadis, *Artificial neural networks for solving ordinary and partial differential equations*, IEEE Trans. Neural Netw., 9 (1998) pp. 987–1000.

[75] G. Liang, *Stochastic control representations for penalized backward stochastic differential equations*, SIAM J. Control Optim., 53 (2015), pp. 1440–1463.

[76] G. Liang and W. Wei, *Optimal switching at Poisson random intervention times*, Discrete Contin. Dyn. Syst. Ser. B, 21 (2016), pp. 1483–1505.

[77] P.-L. Lions and J.-L Menaldi, *Optimal control of stochastic integrals and Hamilton-Jacobi-Bellman equations (part I)*, SIAM J. Control Optim., 20 (1982), pp. 58–81.

[78] P.-L. Lions and A.-S. Sznitman, *Stochastic differential equations with reflecting boundary conditions*, Comm. Pure Appl. Math., 37 (1984) pp. 511– 553.

[79] N. Lundström, K. Nyström, and M. Olofsson, *Systems of variational inequalities in the context of optimal switching problems and operators of Kolmogorov type*, Ann. Mat. Pura Appl., 4 (2014), pp. 1213–1247.

[80] P. Mohajerin Esfahani, D. Chatterjee, and J. Lygeros, *The stochastic reach-avoid problem and set characterization for diffusions*, Automatica, 70 (2016), pp. 43–56.

[81] B. Øksendal and A. Sulem, *Applied Stochastic Control of Jump Diffusions*, Universitext, Springer, Berlin, 2005.

[82] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Classics Appl. Math. 30, SIAM, Philadelphia, 2000.

[83] H. Pham, *Continuous-time Stochastic Control and Optimization with Financial Applications*, Stoch. Model. Appl. Probab. 61, Springer Verlag, Berlin, 2009.

[84] M.L. Puterman and S.L. Brumelle, *On the convergence of policy iteration in stationary dynamic programming*, Math. Oper. Res., 4 (1979), pp. 60–69.

[85] C. Reisinger, W. Stockinger, and Y. Zhang, *A posteriori error estimates for fully coupled McKean-Vlasov forward-backward SDEs*, arXiv preprint, arXiv:2007.07731, 2020.

[86] C. Reisinger and J. H. Witte, *On the use of policy iteration as an easy way of pricing American options*, SIAM J. Financ. Math., 3 (2012), pp. 459–478.

[87] C. Reisinger and Y. Zhang, *A penalty scheme and policy iteration for nonlocal HJB variational inequalities with monotone drivers*, preprint, arXiv:1805.06255 [math.NA], 2018.

[88] C. Reisinger and Y. Zhang, *A penalty scheme for monotone systems with interconnected obstacles: convergence and error estimates*, SIAM J. Numer. Anal., 57 (2019), pp. 1625–1648.

[89] C. Reisinger and Y. Zhang, *Error estimates of penalty schemes for quasi-variational inequalities arising from impulse control problems*, SIAM J. Control Optim., 58 (2020), pp. 243–276.

[90] M. Royer, *Backward stochastic differential equations with jumps and related non-linear expectations*, Stochastic Process. Appl., 116 (2006), pp. 1358–1376.

[91] M.S. Santos and J. Rust, *Convergence properties of policy iteration*, SIAM J. Control Optim., 42 (2004), pp. 2094–2115.

[92] R. C. Seydel, *Impulse control for jump-diffusions: viscosity solutions of quasi-variational inequalities and applications in bank risk management*, PhD Thesis, Leipzig University, 2009.

[93] R. C. Seydel, *Existence and uniqueness of viscosity solutions for QVI associated with impulse control of jump-diffusions*, Stochastic Process. Appl., 119 (2009), pp. 3719–3748.

[94] O. Shamir and T. Zhang, *Stochastic gradient descent for non-smooth optimization: Convergence results and optimal averaging schemes*, in Proceedings of the International Conference on Machine Learning, 2013.

[95] J. Sirignano and K. Spiliopoulos, *DGM: A deep learning algorithm for solving partial differential equations*, J. Comput. Phys., 375 (2018), pp. 1339–1364.

[96] I. Smears and E. Süli, *Discontinuous Galerkin finite element approximation of Hamilton-Jacobi-Bellman equations with Cordes coefficients*, SIAM J. Numer. Anal., 52 (2014), pp. 993–1016,

[97] L. R. Sotomayor and A. Cadenillas, *Stochastic impulse control with regime switching for the optimal dividend policy when there are business cycles, taxes and fixed costs*, Stochastics, 85 (2013), pp. 707–722.

[98] S. Tan, Z. Jin, and G. Yin, *Optimal dividend payment strategies with debt constraint in a hybrid regime-switching jump-diffusion model*, Nonlinear Anal. Hybrid Syst., 27 (2018), pp. 141–156.

[99] M. Ulbrich, *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*, MOS-SIAM Ser. Optim. 11, SIAM, Philadelphia, 2011.

[100] J. Wei, H. Yang, and R. Wang, *Classical and impulse control for the optimization of dividend and proportional reinsurance policies with regime switching*, J. Optim. Theory Appl., 147 (2010), pp. 358–377.

[101] J. H. Witte and C. Reisinger, *A penalty method for the numerical solution of Hamilton-Jacobi- Bellman (HJB) equations in finance*, SIAM J. Numer. Anal., 49 (2011), pp. 213–231.

[102] J. H. Witte and C. Reisinger, *Penalty methods for the solution of discrete HJB equations: Continuous control and obstacle problems*, SIAM J. Numer. Anal., 50 (2012), pp. 595–625.

[103] G. Yin, C. Zhu, *Hybrid Switching Diffusions: Properties and Applications*, Springer, New York, 2010.