

VARIOUS TOPICS IN STOCHASTIC CONTROL AND
MEASURE THEORY: SINGULAR STOCHASTIC CONTROL,
OPTIMAL STOPPING AND DECOMPOSITIONS OF
MEASURES

A THESIS PRESENTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY OF IMPERIAL COLLEGE LONDON
BY
ALESSANDRO MILAZZO

DEPARTMENT OF MATHEMATICS
IMPERIAL COLLEGE LONDON
180 QUEEN'S GATE
LONDON SW7 2BZ

APRIL 2021

Degree: PhD

Thesis title: Various topics in stochastic control and measure theory: singular stochastic control, optimal stopping and decompositions of measures

Candidate: Alessandro Milazzo

Internal supervisor: Dr. Pietro Siorpaes

External supervisor: Dr. Tiziano De Angelis

Department of Mathematics, Imperial College London, 180 Queen's Gate, London SW7 2BZ

Copyright

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Statement of Originality

I, Alessandro Milazzo, certify that this thesis and the research to which it refers are the product of my own work, and that any ideas or quotations from the work of other people, published or otherwise, are fully acknowledged in accordance with the standard referencing practices of the discipline.

Signed: Alessandro Milazzo

To my family.

Abstract

Chapter II describes my doctoral work on the problem of optimally stopping the exponential of a Brownian bridge which was posed by Ernst and Shepp in their paper [42] and was motivated by bond selling with non-negative prices. We show how to obtain an optimal stopping rule for this problem and we prove regularity of the value function and of the optimal boundary.

Chapter III describes my doctoral work on problems of singular control with discretionary stopping. We develop a rigorous probabilistic study for this class of problems which leads to the dynamic programming principle.

Chapter IV describes my doctoral work on some decompositions of measures and processes. In particular, we study a decomposition of measures introduced by Dellacherie from which we show how to obtain well-known decompositions of measures and processes and which we extend to vector measures.

This page was intentionally left blank.

Contents

Copyright	2
Statement of Originality	3
Abstract	5
Frequently Used Notation	10
I Introduction	11
I.1 Optimal stopping for the exponential of a Brownian bridge	11
I.2 Singular stochastic control with discretionary stopping	12
I.3 On a remarkable decomposition of measures by Dellacherie	14
II Optimal stopping for the exponential of a Brownian bridge	16
II.1 Introduction	16
II.2 Problem formulation	18
II.3 Initial regularity of the value function	21
II.4 Existence of an optimal boundary	23
II.5 Regularity of the optimal boundary	25
II.6 Further regularity of the value function	30
II.7 Integral equations for the optimal boundary	35
II.8 Numerical results	39
II.9 Conclusions	43
IIISingular stochastic control with discretionary stopping	45
III.1 Introduction	45
III.2 Notation and terminology	47

III.3 Problem formulation	49
III.4 Independence of the reference probability system	50
III.5 Regular conditional probabilities	55
III.6 Value process	63
III.7 Dynamic programming principle	70
III.8 An optimal stopping rule	75
III.9 Regularity of the value function	84
III.10 Miscellaneous results	93
III.11 Conclusions	95
IV On a remarkable decomposition of measures by Dellacherie	97
IV.1 Introduction	97
IV.2 Dellacherie's decomposition for positive measures	98
IV.3 Dellacherie's decomposition for vector measures	101
IV.4 Regularity properties for Dellacherie's decomposition	105
IV.5 Conclusions	109
Bibliography	112

List of Figures

II.1	Geometry of the optimal stopping for the exponential of a Brownian bridge.	41
II.2	Trajectory of the numerical error for the Picard scheme used in the optimal stopping for the exponential of a Brownian bridge.	41
II.3	The value function of the optimal stopping for the exponential of a Brownian bridge.	42
II.4	Boundary functions of the optimal stopping for the exponential of a Brownian bridge with different terminal times.	42
II.5	Comparison between the Picard scheme and the traditional scheme to solve nonlinear integral equations.	43
II.6	Fitting an ansatz for the optimal boundary of the optimal stopping for the expo- nential of a Brownian bridge.	44
III.1	Existence of the canonical map for a process adapted to the Brownian filtration. .	52

Frequently Used Notation

I. General Notation

Let $a, b \in \mathbb{R}$, $d \in \mathbb{N}$, $x = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d$ and $y = (y_1, y_2, \dots, y_d) \in \mathbb{R}^d$.

- $:=$ means “is defined to be”.
- $a \vee b := \max\{a, b\}$.
- $a \wedge b := \min\{a, b\}$.
- $[a]^+ := \max\{a, 0\}$.
- $[a]^- := \max\{-a, 0\}$.
- $|x| := (\sum_{i=1}^d x_i^2)^{1/2}$ denotes the Euclidean norm.
- $\langle x, y \rangle := \sum_{i=1}^d x_i y_i$ denotes the inner product.

II. Sets and spaces

Let $d, n \in \mathbb{N}$, Ω be a general set and E be a Polish space.

- $\mathbb{R}_+ := [0, \infty)$.
- $A^c := \Omega \setminus A$ denotes the complement of $A \subseteq \Omega$.
- $\bar{B}$ denotes the topological closure of $B \subseteq E$.
- ∂B denotes the boundary of $B \subseteq E$.
- $\mathcal{B}(E)$ denotes the Borel σ -algebra on E .
- $M(\mathbb{R}^d, \mathbb{R}^n)$ denotes the collection of the $d \times n$ -dimensional real matrices.
- $C(E; \mathbb{R}^d)$ denotes the space of continuous functions defined on E with values in $\mathbb{R}^d$.
- $C^{1,2}(A)$ denotes the space of real-valued functions whose first time derivative and second spatial derivative exist and are continuous on a set $A \subseteq E$.

III. Probability and random variables

Let $X : \Omega \rightarrow E$ be a random variable defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and with values in a Polish space E .

- $\mathcal{L}_{\mathbb{P}}(X)$ is the law of X , i.e., the probability measure $\mathbb{P} \circ X^{-1}$ on $(E, \mathcal{B}(E))$. When no confusion may arise, we will simply denote $\mathcal{L}_{\mathbb{P}}(X)$ by $\mathcal{L}(X)$.

Chapter I

Introduction

This thesis is the result of four years of study and research as a PhD candidate in the Department of Mathematics at Imperial College London, under the supervision of Dr. Pietro Siorpaes and Dr. Tiziano De Angelis. It is divided into three chapters which describe three independent projects. Chapter II and III concern problems of stochastic optimal control and optimal stopping, whereas Chapter IV involves decompositions of measures and processes. In what follows I provide some background notions, motivation and results of the three projects.

I.1 Optimal stopping for the exponential of a Brownian bridge

In a problem of optimal stopping an agent aims to stop a process at a specific time in order to optimise his/her expected payoff. Mathematically, this is formulated by choosing an optimal stopping time within a specific class to optimise the expectation of a stopped process. These problems arise from various applications, for instance, in statistics, mathematical finance and engineering. One of the first examples appeared in Wald's sequential analysis [84] for a method of statistical inference where the number of observations needed for an experiment is not determined in advance. A great contribution to the optimal stopping theory was provided by Snell [83] who formulated a general optimal stopping problem for discrete-time stochastic processes and characterised the solution by means of the smallest supermartingale (the so called "Snell's envelope") dominating the reward sequence. Another key feature of optimal stopping problems is their connection to free-boundary problems (as we will see in Chapter II), which was first highlighted in a paper by Mikhalevich [71]. The general theory of optimal stopping was then developed and extended to continuous time, with several contributions especially in the 1960's and 1970's. In our work we will mostly refer to the more recent books by Peskir & Shiryaev [79] and by Karatzas & Shreve [65, Appendix D], who provide a comprehensive general theory.

Problems of optimal stopping have found several natural applications in mathematical finance. The most famous example is perhaps the pricing of an American option which is equivalent to solving an optimal stopping problem (see, e.g., [55]). This optimal stopping approach has another interesting feature as the optimal stopping time of the problem is the optimal exercise time of the option and it is proved to be the first time that the stock hits a the so-called optimal boundary.

Mathematically, this boundary can be characterised as the unique continuous solution to a family of integral equations.

Another interesting optimal stopping problem from the mathematical finance perspective is the optimal stopping of a Brownian bridge. Historically, in mathematical finance, Brownian bridges have been used to model bonds as their pinning feature reflects the face value of a bond. Thus, previous works by Shepp [82] on the optimal stopping of a Brownian bridge were applied by Boyce [22] in 1970 to study optimal selling times for bonds. It was then shown in more recent years that Brownian bridges are not only suitable to model bonds but also stocks when the so called “stock pinning” effect arises. As highlighted in a paper by Avellaneda & Lipkin [5], this is a phenomenon that has been empirically observed in equity options markets: when the number of outstanding contracts corresponding to a particular strike is unusually large, several stock prices close to maturity tend to end up (and thus pin) at the strike price. However, adopting a Brownian bridge to model a bond or stock price has the unrealistic disadvantage that the price could become negative. To overcome this unpleasant feature, Ernst and Shepp [42] posed the problem of solving the optimal stopping for the exponential of a Brownian bridge, where they also highlighted its difficult analytical tractability. Motivated by this challenge, we decided to study this problem which is presented in details in Chapter II.

We begin by proving continuity of the value function and showing that our problem admits an optimal stopping boundary which we aim to study. For time-homogeneous optimal stopping problems it can be relatively easy to characterise the optimal boundary but this is usually very challenging for time-inhomogeneous problems. For this reason, there are only a few examples of fully-solved time-inhomogeneous optimal stopping problems. One of these examples is the optimal stopping of a Brownian bridge which has been studied by several authors using different methods. The first results are due to Dvoretzky [38] and Shepp [82] and were later extended by Ekström and Wanhörp [41]. They were able to show the explicit form of the optimal boundary by applying a change of measure in order to formulate an auxiliary equivalent stopping problem for a time-homogeneous diffusion. However, these techniques are not applicable to the problem we study in Chapter II due to the exponential nature of the gain function. Hence, to prove the monotonicity of the optimal boundary - which is central in the analysis of the problem - we develop novel techniques based on the pathwise properties of the Brownian bridge. We believe that this method could be applied to more general time-inhomogeneous optimal stopping problem and this extension is in our future plans of research. By proving the monotonicity of the optimal boundary, we are able to obtain C^1 -regularity for the value function and to characterise the optimal boundary as a unique continuous solution to a specific family of Volterra integral equations. We then solve numerically these integral equations and obtain an approximation of the optimal boundary.

I.2 Singular stochastic control with discretionary stopping

In stochastic optimal control problems, an agent aims to optimise his/her expected payoff over a class of stochastic processes called the controls. These controls can affect the underlying dynamics in different ways. In classical stochastic control problems, the controls acts on the drift and diffusion coefficients of the underlying stochastic differential equation. These problems have several applications especially in mathematical finance: the most famous example is perhaps the Merton problem [70] where an investor chooses how much of his/her wealth to consume and much to allocate between a stock and a risk-free asset as to maximise his/her expected utility. In singular

stochastic control problems, the controls are characterised by being of “singular” nature (i.e., in general not absolutely continuous with respect to the Lebesgue measure as a function of time) and by acting additively on the underlying dynamics. To exercise a control the optimiser pays a cost which is proportional to the total variation of the control and represents the cost of fuel used to activate the control or a transaction cost in mathematical finance.

In a stochastic control problem, if the agent can also choose when to end the optimisation we then refer to stochastic control with discretionary stopping. Problems of singular control with discretionary stopping naturally arise from at least two applications: target tracking problems in engineering and consumption/investment problems with transaction cost in mathematical finance. In target tracking problems, the aim of the optimiser is to arrive as close as possible to a fixed target by paying a cost proportional to the fuel used, and decide to stop when the target is sufficiently close. Specific examples of these problems have been studied in, e.g., [54] and [18]. Consumption/investment problems involving transaction cost are a more realistic version of the Merton problem, where the investor pays a cost proportional to the amount transacted. In a famous paper by Davis and Norman [26] it was shown that optimal strategies in these problems are of singular nature. In this context, the stopping decision variable acts as an early exit from the investment or the exercise time of an American option held in the portfolio.

A classic result in stochastic control theory is the dynamic programming principle which was first formulated by Bellman [11]. Loosely speaking, this result guarantees that a control which is optimal at an initial time $t \in [0, T]$ will still be optimal after following the optimisation up to a future time $u \in [t, T]$ and then re-starting with the same control from there. This is very intuitive in several situations: for instance, say we want to go from a point A to a point B and we delineate the optimal path which passes through a point C; once we have reached point C it will still be optimal to follow the same path, otherwise this would contradict the optimality of the initial path. As one can expect, the dynamic programming principle turns out to be a powerful tool to study optimal strategies. Despite its simplicity, a mathematical proof is very technical, in particular when the problem has no intrinsic Markovian structure. For this reason, rigorous proofs have often been omitted and thus the dynamic programming principle has often been used naively.

Whilst investigating a problem of singular control with discretionary stopping, we realised no rigorous proof of the dynamic programming principle was available for this specific class of problems. Hence, we have developed a detailed probabilistic study of singular control with discretionary stopping problems which leads to a rigorous proof of the dynamic programming principle for this class, and which we present in Chapter III. In particular, we extend the work by Hausmann and Suo [52], who obtain the dynamic programming principle for singular control problems, by additionally including in the framework: (i) discretionary stopping, (ii) a terminal gain function, (iii) an exit time from a fixed domain. In fact, Bouchard and Touzi [21] established (a weak version of) the dynamic programming principle for some general stochastic control problems which can fit different framework. However, given the generality of their work, they need to set some abstract assumptions (e.g., on the class of admissible controls) which an interested reader should verify every time before applying their result. Moreover, fitting a specific control problem in their general framework may not always be straightforward. Thus, our aim here is to provide a self-contained probabilistic study of singular control with discretionary stopping problems which leads to a detailed proof of their dynamic programming principle: in this way anyone who is working on a specific problem of singular control with discretionary stopping could directly apply our results. Since our framework is in general non-Markovian, another purpose of our work is to rigorously show how to use the

technical concept of regular conditional probabilities which provides a pseudo-Markovian structure essential to obtain the dynamic programming principle. To this extent, we follow closely the analysis by Fabbri, Gozzi & Swiech [43] who establish the dynamic programming principle for classical control problems in infinite dimension. Thus, our work seems extendable to include also classical controls and is designed to fill the gap between the literature of dynamic programming principle established (and sometimes used naively) for specific control problems (e.g., for Markovian control problems) and some rather abstract works such as [21] and [19]. In addition, we also provide an optimal stopping rule for our problem. However, at the moment we are only able to show the optimality of the stopping time under some strong technical assumptions which are not easy to prove and thus we aim to extend our result under some milder assumptions.

I.3 On a remarkable decomposition of measures by Dellacherie

In measure theory and stochastic analysis, decompositions have been built in order to characterise measures and processes and simplify their study. The different parts constituting a decomposition satisfy specific properties and are, in some mathematical sense, orthogonal to each other. Since most of these decompositions were built in distinct frameworks, they reveal different concepts of orthogonality.

Whilst studying the connections between some of these decompositions and trying to interpret them in a more unifying framework, we discovered a surprisingly interesting decomposition of measures introduced by Dellacherie [34, Ch. II, Sec. 2, Th. T14]: given a σ -finite measure $\mu : \Omega \rightarrow [0, \infty]$ on a measurable space $(\Omega, \mathcal{F})$ and a sub-family $\mathcal{G} \subseteq \mathcal{F}$ closed under countable unions, there exists a set $G_\mu \in \mathcal{G}$ and a unique decomposition

$$\mu = \mu_{\mathcal{G}} + \mu_{\mathcal{G}}^\perp, \tag{I.3.1}$$

where $\mu(G_\mu^c) = 0$ and $\mu_{\mathcal{G}}^\perp(G) = 0$ for every $G \in \mathcal{G}$.

In accordance with Dellacherie, who states decomposition (I.3.1) deserves to be classical (“méritait d’être classique”), we decided to study this decomposition more into details and we report our results in Chapter IV. Despite the simplicity of its statement, the proof in Dellacherie’s book was initially not clear to us and relies on several non-obvious properties of the order between measures with no specific mention. We thus begin our study by proposing a simple original proof based only on the concept of essential supremum. We then show why this decomposition is remarkably powerful: we obtain several well-known decompositions of measures (e.g., the Lebesgue decomposition) and some new potentially interesting decompositions as a simple application of (I.3.1).

Our contribution is also to vastly extend the range of applicability of Dellacherie’s decomposition by establishing a version of the theorem for vector measures. As a consequence, we show how this extension can be applied to signed measures to derive the Jordan decomposition and to the stochastic integral to obtain a decomposition of semimartingales due to Bichteler [15, Theorem 4.4.5]. Finally, we also study some regularity properties of the decomposition with respect to the order and with respect to the topology.

Our work enhances the strength of this decomposition which deserves to be better known by the

mathematical community. We also believe further (well-known or new) decompositions of measures and processes could be obtained as corollaries of this decomposition and we are currently working in this direction.

Chapter II

Optimal stopping for the exponential of a Brownian bridge

In this chapter I present my doctoral research on the optimal stopping for the exponential of a Brownian bridge, where we study the problem of stopping a Brownian bridge in order to maximise the expected value of an exponential gain function.

II.1 Introduction

Problems of optimal stopping involving the Brownian bridge have a long history, dating back to the early days of modern optimal stopping theory. The first results were obtained by Dvoretzky [38] and Shepp [82]. Both authors considered stopping of a Brownian bridge to maximise its expected value. Dvoretzky proved existence of an optimal stopping time and Shepp provided an explicit solution in terms of the first time the Brownian bridge (pinned at zero at time $T = 1$) exceeds a boundary of the form $t \mapsto a\sqrt{1-t}$, for $t \in [0, 1]$ and a suitable $a > 0$.

Few years later, Föllmer [45] extended the study to the case of a Brownian bridge whose pinning point is random with normal distribution. He showed that the optimal stopping time is the first time the process crosses a time-dependent boundary and the stopping set may lie either above or below the boundary, depending on the variance of the pinning point's distribution.

More recently, Ekström and Wamntorp [41] studied optimal stopping of a Brownian bridge via the solution of associated free boundary problems. They recovered results by Shepp and extended the analysis by finding explicit solutions to some examples with more general gain functions than the linear case.

Optimal stopping of a Brownian bridge with random pinning point or random pinning time were also studied in [40] and [48], respectively. In [40], the authors consider more general versions of the problem addressed in [45] and, among their results, they give general sufficient conditions for optimal stopping rules in the form of a hitting time to a one-sided stopping region. In [48], the author provides sufficient conditions for a one-sided stopping set and is able to solve the problem in closed form for some choices of the pinning time's distribution.

Problems of optimal stopping for Brownian bridge have attracted significant attention from the mathematical finance community because of their application to trading. Already in 1970, Boyce [22] proposed applications of Shepp's results to bond trading. In that context, the pinning effect of the Brownian bridge captures the well-known *pull-to-par* mechanism of bonds. Several other applications to finance have appeared in recent years, motivated by phenomena of stock pinning (see, e.g., [5] and [58]). Explicit results for some problems of optimal double stopping of a Brownian bridge, also inspired by finance, were obtained in [9].

In our paper we study a problem that was posed by Ernst and Shepp in [42, Section 3]. In particular, we are interested in finding the optimal stopping rule that maximises the expected value of the exponential of a Brownian bridge which is constrained to be equal to zero at time $T = 1$. Besides the pure mathematical interest, this problem is better suited to model bond/stock trading situations than its predecessors with linear gain function. Indeed, the exponential structure avoids the unpleasant feature of negative asset prices, whilst retaining the pinning effect discussed above. Questions concerning stopping the exponential of a Brownian bridge were also considered in [69] in a model inspired by financial applications. In fact, in [69] authors consider a more general model than ours and allow a random pinning point. However, the complexity of the model is such that their analysis is carried out mostly from a numerical point of view.

In this work, we prove that the optimal stopping time for our problem is the first time the Brownian bridge exceeds a time-dependent optimal boundary $t \mapsto b(t)$, which is non-negative, continuous and non-increasing on $[0, 1]$. The boundary is uniquely characterised as the continuous solution to a family of integral equations of Volterra type, which can be computed numerically (see Section II.7). The full analysis that we perform relies on four equivalent formulations of the problem (see (II.2.7), (II.2.14), (II.2.16) and (II.4.4)), which are of interest in their own right, and offer different points of view on the problem.

Our study reveals interesting features of the value function v . Indeed, we can prove that v is continuously differentiable on $[0, 1) \times \mathbb{R}$, both with respect to time and space, with second order spatial derivative which is continuous up to the optimal boundary (regularity which goes beyond the standard smooth-fit condition in optimal stopping). Notice, however, that the value function is *not* continuous at $\{1\} \times (-\infty, 0)$, due to the pinning behaviour of the Brownian bridge as $t \rightarrow 1$.

We extend the existing literature in several directions. The non-linear structure of the exponential gain function does not allow to use scaling properties that are central in all the papers where explicit solutions were obtained (see, e.g., [82], [41], [9], [48]). For this reason we must deal directly with a stopping problem for a time-inhomogeneous diffusion. Optimal boundaries for such problems are hard to come by in the literature and, in order to prove monotonicity of the boundary (which is the key to the subsequent analysis), we have developed a method based on pathwise properties of the Brownian bridge and martingale theory (see Theorem II.5.1). The task is challenging because there is no obvious comparison principle for sample paths of Brownian bridges $X^{t,x}$ and $X^{t',x}$ starting from a point $x \in \mathbb{R}$ at different times $t \neq t'$. Hence, our approach could be used in other optimal stopping problems involving time-inhomogeneous diffusions.

It is worth noticing that, in Section 5 of [40], authors also obtain a characterisation of the optimal boundary via integral equations. However, in that case a time-change of the Brownian bridge and linearity of the gain function are used to infer monotonicity of the boundary.

The chapter is organised as follows. In Section II.2 we provide some background notions on the Brownian bridge and we formulate the stopping problem. In Section II.3 we prove some

initial regularity of the value function, in particular its continuity. In Section II.4 we show the existence of an optimal boundary. In Section II.5 we prove that the boundary is monotonic non-increasing, continuous and bounded on $[0, 1]$ and find its limit at time $t = 1$. In Section II.6 we find C^1 -regularity for the value function. In Section II.7 we derive the family of integral equations that uniquely characterise the optimal boundary. In Section II.8 we solve the integral equation numerically using a Picard's iteration scheme and we provide plots of the optimal boundary and of the value function. We also illustrate numerically the convergence of the algorithm for the boundary and the dependence of the boundary on the pinning time of the Brownian bridge.

II.2 Problem formulation

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ be a complete filtered probability space, where $(\mathcal{F}_t)_{t \geq 0}$ is the filtration generated by a Brownian motion $W := (W_t)_{t \geq 0}$ and augmented with $\mathbb{P}$ -null sets. For $(t, x) \in [0, 1] \times \mathbb{R}$, we denote by $X := (X_s)_{s \in [t, 1]}$ a Brownian bridge starting from $X_t = x$ and pinning at zero at time $T = 1$, i.e., such that $X_1 = 0$. This is described by the stochastic differential equation (SDE)

$$\begin{aligned} dX_s &= -\frac{X_s}{1-s} ds + dW_s, \quad s \in [t, 1), \\ X_t &= x. \end{aligned} \tag{II.2.1}$$

In order to keep track of the initial conditions, we sometimes denote X by $X^{t,x}$. The unique strong solution to the SDE (II.2.1) is given by

$$X_s^{t,x} = (1-s) \left(\frac{x}{1-t} + \int_t^s \frac{dW_u}{1-u} \right), \quad s \in [t, 1]. \tag{II.2.2}$$

The expression in (II.2.2) allows to identify (in law) the process X with the process $Z := (Z_s)_{s \in [t, 1]}$ given by

$$Z_s := \frac{1-s}{1-t} x + \sqrt{\frac{1-s}{1-t}} W_{s-t}, \quad s \in [t, 1]. \tag{II.2.3}$$

That is, we have

$$\mathcal{L}(X) = \mathcal{L}(Z). \tag{II.2.4}$$

Analogously, we will sometimes denote Z by $Z^{t,x}$. Using this identity in law between X and Z , along with well-known distributional properties of the Brownian motion, we obtain

$$\mathbb{E} \left[\sup_{t \leq s \leq 1} e^{X_s} \right] \leq e^{|x|} \mathbb{E} [e^{S_1}] < +\infty, \tag{II.2.5}$$

where $S_1 := \sup_{0 \leq s \leq 1} W_s$. The random variable S_1 will be used several times in what follows and we denote

$$c_1 := \mathbb{E} [e^{S_1}] \quad \text{and} \quad c_2 := \mathbb{E} [S_1 e^{S_1}], \tag{II.2.6}$$

which are both finite.

Our objective is to study the optimal stopping problem

$$v(t, x) = \sup_{0 \leq \tau \leq 1-t} \mathbb{E} \left[e^{X_{t+\tau}^{t,x}} \right], \quad \forall (t, x) \in [0, 1] \times \mathbb{R}, \quad (\text{II.2.7})$$

where τ is a random time such that $t + \tau$ is a $(\mathcal{F}_s)_{s \geq t}$ -stopping time (in what follows we simply say that τ is a $(\mathcal{F}_s)_{s \geq t}$ -stopping time, as no confusion shall arise).

Remark II.2.1. *There is no loss of generality in our choice of a pinning time $T = 1$ and a pinning point $\alpha = 0$. We could equivalently choose a generic pinning time $T > t \geq 0$ and a generic pinning point $\alpha \in \mathbb{R}$ and consider the dynamics*

$$dX_s = -\frac{X_s - \alpha}{T - s} ds + dW_s, \quad s \in [t, T). \quad (\text{II.2.8})$$

Then, the analysis of the problem would remain valid up to obvious tweaks.

Since $Y := e^X$ is continuous, non-negative and the integrability condition (II.2.5) holds, we can rely upon standard optimal stopping theory to give some initial results (see, e.g., [65, Appendix D]). In particular, we split the state space $[0, 1] \times \mathbb{R}$ into a continuation region $\mathcal{C}$ and a stopping region $\mathcal{D}$, respectively given by

$$\mathcal{C} := \{(t, x) \in [0, 1] \times \mathbb{R} : v(t, x) > e^x\}, \quad (\text{II.2.9})$$

$$\mathcal{D} := \{(t, x) \in [0, 1] \times \mathbb{R} : v(t, x) = e^x\}. \quad (\text{II.2.10})$$

Then, for every $(t, x) \in [0, 1] \times \mathbb{R}$, the smallest optimal stopping time for problem (II.2.7) is given by (see, e.g., [65, Thm. D.12, Appendix D])

$$\tau^* := \inf\{s \in [0, 1 - t] : (t + s, X_{t+s}^{t,x}) \in \mathcal{D}\}, \quad \mathbb{P}\text{-a.s.}, \quad (\text{II.2.11})$$

i.e.,

$$v(t, x) = \mathbb{E} \left[e^{X_{t+\tau^*}^{t,x}} \right]. \quad (\text{II.2.12})$$

We will sometimes denote τ^* by $\tau_{t,x}^*$ to keep track of the initial condition of the time-space process (t, X) .

Moreover, standard theory of the Snell envelope also guarantees (see, e.g., [65, Thm. D.9, Appendix D]) that the process $V := (V_t)_{t \in [0, 1]}$ defined by $V_t := v(t, X_t)$ is a right-continuous, non-negative, $\mathbb{P}$ -supermartingale and that the stopped process $V^* := (V_{t \wedge \tau^*})_{t \in [0, 1]}$ is a right-continuous, non-negative, $\mathbb{P}$ -martingale.

Remark II.2.2. *Notice that here we have preferred to skip the technical details concerning the Snell envelope in order to focus on the subsequent analysis of the problem. In fact, by [65, Thm. D.12, Appendix D], the optimal stopping time τ^* is of the form*

$$\tau^* = \inf\{s \in [t, 1] : V_s^0 = e^{X_s}\}, \quad (\text{II.2.13})$$

where V^0 is the Snell envelope for our problem. However, in the next section we will show (only by using the existence of an optimal stopping time but not its nature) that $(t, x) \mapsto v(t, x)$ is continuous. Under this condition, our definition (II.2.11) of τ^* is guaranteed to be optimal by, e.g., [79, Corollary 2.9].

To conclude this section, we show two further formulations of problem (II.2.7) that will become useful in our analysis. The former is provided by the equality in law (II.2.4), which yields

$$v(t, x) = \sup_{0 \leq \tau \leq 1-t} \mathbb{E} \left[e^{Z_{t+\tau}^{t,x}} \right], \quad \forall (t, x) \in [0, 1] \times \mathbb{R}. \quad (\text{II.2.14})$$

The second formulation, instead, uses ideas originally contained in the proof of [57, Lemma 3.9]. In particular, for any fixed $t \in [0, 1]$ and any $(\mathcal{F}_s)_{s \geq t}$ -stopping time $\tau \in [0, 1-t]$, we can define an $(\hat{\mathcal{F}}_s)_{0 \leq s \leq 1}$ stopping time $\theta \in [0, 1]$ such that $\tau =: \theta(1-t)$ and $\hat{\mathcal{F}}_s := \mathcal{F}_{s(1-t)}$. In addition to this, notice that

$$\mathcal{L}(W_{s(1-t)}, s \geq 0) = \mathcal{L}(\sqrt{1-t} W_s, s \geq 0). \quad (\text{II.2.15})$$

Therefore, problem (II.2.14) and, hence, problem (II.2.7) can be rewritten as

$$v(t, x) = \sup_{0 \leq \theta \leq 1} \mathbb{E} \left[\exp \left((1-\theta)x + \sqrt{(1-\theta)(1-t)} W_\theta \right) \right]. \quad (\text{II.2.16})$$

This last formulation of the problem has the advantage that the domain of admissible stopping times θ is independent of the initial time t .

Remark II.2.3. *We have skipped an intermediate technical step to pass from (II.2.14) to (II.2.16). In fact, for every $(\mathcal{F}_s)$ -stopping time $\tau \in [0, 1-t]$ $\mathbb{P}$ -a.s., we can define a $(\hat{\mathcal{F}}_s)$ -stopping time $\rho \in [0, 1]$ $\mathbb{P}$ -a.s. such that $\tau = \rho(1-t)$. Then, by (II.2.14), we would obtain that*

$$v(t, x) = \sup_{0 \leq \rho \leq 1} \mathbb{E} \left[\exp \left((1-\rho)x + \sqrt{1-\rho} W_{\rho(1-t)} \right) \right], \quad (\text{II.2.17})$$

where ρ are $(\hat{\mathcal{F}}_s)$ -stopping times. Since $(\hat{\mathcal{F}}_s)$ is generated by the process $\hat{W}_s := W_{s(1-t)}$ then a $(\hat{\mathcal{F}}_s)$ -stopping time ρ can be equivalently represented as a measurable map of the process $\hat{W}$ (for a technical discussion see later Lemma III.4.3). The same measurable map evaluated in the process $\tilde{W}_s := \sqrt{1-t} W_s$ gives rise to a $(\mathcal{F}_s)$ -stopping time $\theta \in [0, 1]$ $\mathbb{P}$ -a.s. which maintains the equality in law because of (II.2.15). Hence,

$$\begin{aligned} v(t, x) &= \sup_{0 \leq \rho \leq 1} \mathbb{E} \left[\exp \left((1-\rho)x + \sqrt{1-\rho} W_{\rho(1-t)} \right) \right] \\ &= \sup_{0 \leq \theta \leq 1} \mathbb{E} \left[\exp \left((1-\theta)x + \sqrt{(1-\theta)(1-t)} W_\theta \right) \right], \end{aligned} \quad (\text{II.2.18})$$

where θ are $(\mathcal{F}_s)$ -stopping times, which leads to (II.2.16).

To conclude this section we make one last remark. Since, as discussed, our optimisation problem is independent of change of processes that are equal in law, we will not distinguish between the integral form of (II.2.1), i.e.,

$$X_{t+s}^{t,x} = x - \int_0^s \frac{X_{t+u}^{t,x}}{1-(t+u)} du + W_{t+s} - W_s, \quad s \in [0, 1-t], \quad (\text{II.2.19})$$

and

$$X_{t+s}^{t,x} = x - \int_0^s \frac{X_{t+u}^{t,x}}{1-(t+u)} du + W_s, \quad s \in [0, 1-t]. \quad (\text{II.2.20})$$

II.3 Initial regularity of the value function

In this section we show some properties of the value function, including its continuity.

It follows immediately from (II.2.5) that the value function is non-negative and uniformly bounded on compact sets. In particular, we have

$$0 < v(t, x) \leq c_1 e^{|x|}, \quad \forall (t, x) \in [0, 1] \times \mathbb{R}, \quad (\text{II.3.1})$$

where $c_1 > 0$ is given by (II.2.6).

In the following two propositions we prove some regularity for the maps $x \mapsto v(t, x)$ and $t \mapsto v(t, x)$, which will, in particular, yield the continuity of the value function $(t, x) \mapsto v(t, x)$.

Proposition II.3.1. *The map $x \mapsto v(t, x)$ is convex and non-decreasing. Moreover, for any compact set $K \subset \mathbb{R}$ there exists $L_K > 0$ such that*

$$\sup_{t \in [0, 1]} |v(t, y) - v(t, x)| \leq L_K |y - x|, \quad \forall x, y \in K. \quad (\text{II.3.2})$$

Proof. Convexity of $x \mapsto v(t, x)$ follows from linearity of $x \mapsto Z_s^{t, x}$ (see (II.2.3)), convexity of the map $x \mapsto e^x$ and the well-known inequality $\sup(a + b) \leq \sup a + \sup b$.

Monotonicity can be easily deduced by, e.g., the explicit dependence of (II.2.16) on $x \in \mathbb{R}$. As for the Lipschitz continuity (II.3.2), the claim is trivial for $t = 1$ since $v(1, x) = e^x$. For the remaining cases, fix $t \in [0, 1)$ and let us fix $x, y \in \mathbb{R}$ with $y > x$ (the arguments are symmetric for $x > y$). For simplicity, we denote the optimal stopping time $\tau_{t, y}^*$ by τ_y . Then, we obtain

$$\begin{aligned} 0 &\leq v(t, y) - v(t, x) \\ &\leq \mathbb{E} \left[e^{Z_{t+\tau_y}^{t, y}} - e^{Z_{t+\tau_y}^{t, x}} \right] \\ &= \mathbb{E} \left[\left(\exp \left(\frac{1 - (t + \tau_y)}{1 - t} y \right) - \exp \left(\frac{1 - (t + \tau_y)}{1 - t} x \right) \right) \exp \left(\sqrt{\frac{1 - (t + \tau_y)}{1 - t}} W_{\tau_y} \right) \right] \\ &\leq \mathbb{E} \left[\left(\frac{1 - (t + \tau_y)}{1 - t} \right) \exp \left(\sqrt{\frac{1 - (t + \tau_y)}{1 - t}} W_{\tau_y} \right) \right] e^{x \vee y} (y - x) \\ &\leq \mathbb{E} \left[e^{S_1} \right] e^{x \vee y} (y - x), \end{aligned} \quad (\text{II.3.3})$$

where the first inequality follows from monotonicity of $x \mapsto v(t, x)$; the second inequality follows from the fact that τ_y is sub-optimal for $v(t, x)$; the third inequality follows from $|e^a - e^b| \leq e^{a \vee b} |a - b|$. Hence, the claim follows with $L_K := c_1 \max_{x \in K} e^x$. $\square$

Next, we show that the value function is locally Lipschitz in time on $[0, 1) \times \mathbb{R}$. However, it fails to be continuous at $\{1\} \times (-\infty, 0)$.

Proposition II.3.2. *For every $T < 1$ and every $0 \leq t_1 < t_2 \leq T$, we have*

$$|v(t_2, x) - v(t_1, x)| \leq \frac{c_2 e^{|x|}}{\sqrt{1 - T}} (t_2 - t_1), \quad \forall x \in \mathbb{R}, \quad (\text{II.3.4})$$

with $c_2 > 0$ as in (II.2.6). Moreover,

$$\lim_{t \rightarrow 1} v(t, x) = e^x, \quad \forall x \geq 0, \quad (\text{II.3.5})$$

$$\liminf_{t \rightarrow 1} v(t, x) \geq 1 > e^x, \quad \forall x < 0. \quad (\text{II.3.6})$$

Proof. For the proof of (II.3.4) we will refer to the problem formulation in (II.2.16). Fix $0 \leq t_1 < t_2 \leq T < 1$ and let $\theta_2 := \theta_{t_2, x}^*$ be the optimal stopping time for $v(t_2, x)$. Then, given that θ_2 is admissible and sub-optimal for $v(t_1, x)$, we have

$$\begin{aligned} & v(t_2, x) - v(t_1, x) \\ & \leq \mathbb{E} \left[e^{(1-\theta_2)x} \left(e^{\sqrt{(1-\theta_2)(1-t_2)}W_{\theta_2}} - e^{\sqrt{(1-\theta_2)(1-t_1)}W_{\theta_2}} \right) \right] \\ & \leq e^{|x|} \mathbb{E} \left[e^{\sqrt{(1-\theta_2)(1-t_1)}W_{\theta_2}} \sqrt{(1-\theta_2)} |W_{\theta_2}| \right] (\sqrt{1-t_1} - \sqrt{1-t_2}) \\ & \leq e^{|x|} \mathbb{E} \left[S_1 e^{S_1} \right] \frac{t_2 - t_1}{\sqrt{1-T}}, \end{aligned} \quad (\text{II.3.7})$$

where in the last inequality we have also used that $|W_{\theta_2}| \leq \sup_{0 \leq s \leq 1} W_s + \sup_{0 \leq s \leq 1} (-W_s)$ and $\mathcal{L}(W) = \mathcal{L}(-W)$. Now, setting $\theta_1 := \theta_{t_1, x}^*$ we notice that θ_1 is admissible and sub-optimal for the problem with value $v(t_2, x)$. Then, arguments as above give

$$v(t_2, x) - v(t_1, x) \geq -e^{|x|} \mathbb{E} \left[S_1 e^{S_1} \right] \frac{t_2 - t_1}{\sqrt{1-T}}, \quad (\text{II.3.8})$$

which, combined with (II.3.7), implies (II.3.4).

Finally, we show (II.3.5) and (II.3.6). Notice first that $v(1, x) = e^x$ and $v(t, x) \geq e^x$ for $t \in [0, 1)$. Let $x \geq 0$, then by (II.2.14) we have $e^x \leq v(t, x) \leq e^x \mathbb{E} [e^{S_{1-t}}]$ which implies (II.3.5) by dominated convergence and using that $S_{1-t} \rightarrow 0$ as $t \rightarrow 1$. If $x < 0$, instead, the sub-optimal strategy $\tau = 1 - t$ gives $v(t, x) \geq 1$. Hence, $\liminf_{t \rightarrow 1} v(t, x) \geq 1 > e^x = v(1, x)$ as in (II.3.6). $\square$

As a consequence of Proposition II.3.1 and Proposition II.3.2, we obtain that the value function $(t, x) \mapsto v(t, x)$ is continuous on $[0, 1) \times \mathbb{R}$. Moreover, notice that $\mathcal{C} = \mathcal{C} \cap [0, 1) \times \mathbb{R}$ and thus $\mathcal{C}$ is an open set. Combining this fact with the martingale property (in $\mathcal{C}$) of the value function, we obtain the following classical result in optimal stopping theory.

Proposition II.3.3. *We have that $v \in C^{1,2}(\mathcal{C})$ and it satisfies the free-boundary problem*

$$\left(\partial_t + \frac{1}{2} \partial_{xx} - \frac{x}{1-t} \partial_x \right) v(t, x) = 0, \quad (t, x) \in \mathcal{C} \quad (\text{II.3.9})$$

$$v(t, x) = e^x, \quad (t, x) \in \partial \mathcal{C}, \quad (\text{II.3.10})$$

where ∂_t , ∂_x and ∂_{xx} denote the time derivative, the first spatial derivative and the second spatial derivative, respectively.

Proof. As $\mathcal{C}$ is an open set then $\partial \mathcal{C} \subseteq \mathcal{D}$. Thus, the boundary condition (II.3.10) is trivially satisfied and so we now prove that also (II.3.9) holds.

Let $(t, x) \in \mathcal{C}$. Since $\mathcal{C}$ is an open set, there exists an open rectangle $\mathcal{R} := (t_1, t_2) \times (x_1, x_2)$ such

that $(t, x) \in \mathcal{R} \subset \mathcal{C}$. Let $\partial_0 \mathcal{R} := \partial \mathcal{R} \setminus \{t_1\} \times (x_1, x_2)$ and let us define the boundary problem

$$\left(\partial_t + \frac{1}{2} \partial_{xx} - \frac{x}{1-t} \partial_x \right) f(t, x) = 0, \quad (t, x) \in \mathcal{R} \quad (\text{II.3.11})$$

$$f(t, x) = v(t, x), \quad (t, x) \in \partial_0 \mathcal{R}. \quad (\text{II.3.12})$$

The general theory for parabolic partial differential equations (see, e.g., [46, Ch. 6, Sec. 3]) guarantees the existence and uniqueness of a solution $f \in C^{1,2}(\mathcal{R}) \cap C(\bar{\mathcal{R}})$ to the problem (II.3.11)-(II.3.12). Then, let $M_s := f(s, X_s)$ and $\sigma := \inf\{s \geq t : (s, X_s) \in \partial_0 \mathcal{R}\}$. By applying Ito's lemma, we obtain that $(M_{s \wedge \sigma})_{s \in [t, 1]}$ is a bounded martingale. Thus, by the optional sampling theorem and using $f(\sigma, X_\sigma) = v(\sigma, X_\sigma)$, we have

$$f(t, x) = M_t = \mathbb{E}[M_\sigma] = \mathbb{E}[v(\sigma, X_\sigma)] = v(t, x), \quad (\text{II.3.13})$$

where the last equality follows from the fact that $\sigma \leq \tau^*$ and that $v(s \wedge \tau^*, X_{s \wedge \tau^*})$ is a martingale. Since $(t, x) \in \mathcal{C}$ was arbitrary, we obtain that $v \in C^{1,2}(\mathcal{C})$ and it satisfies (II.3.9). $\square$

For future reference, we also denote by $\mathcal{A}$ the second order differential operator associated with X . That is, for every $f \in C^{0,2}(\mathbb{R}^2)$,

$$(\mathcal{A}f)(t, x) := \frac{1}{2} \partial_{xx} f(t, x) - \frac{x}{1-t} \partial_x f(t, x). \quad (\text{II.3.14})$$

II.4 Existence of an optimal boundary

In this section we derive the existence of an optimal stopping boundary for our problem.

To this purpose, it is convenient to perform a change of measure in our problem formulation (II.2.7). In particular, using the integral form of (II.2.1) (upon setting $B_s := W_{t+s} - W_t$), we have

$$\begin{aligned} \mathbb{E} [\exp(X_{t+\tau}^{t,x})] &= \mathbb{E} \left[\exp \left(x + B_\tau - \int_0^\tau \frac{X_{t+s}^{t,x}}{1-(t+s)} ds \right) \right] \\ &= e^x \mathbb{E} \left[\exp \left(B_\tau - \frac{1}{2} \tau \right) \exp \left(\int_0^\tau \left(\frac{1}{2} - \frac{X_{t+s}^{t,x}}{1-(t+s)} \right) ds \right) \right] \\ &= e^x \tilde{\mathbb{E}} \left[\exp \left(\int_0^\tau \left(\frac{1}{2} - \frac{X_{t+s}^{t,x}}{1-(t+s)} \right) ds \right) \right], \end{aligned} \quad (\text{II.4.1})$$

where

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} \Big|_{\mathcal{F}_s} := \exp \left(B_s - \frac{1}{2} s \right), \quad s \in [0, 1-t], \quad (\text{II.4.2})$$

defines a new equivalent probability measure $\tilde{\mathbb{P}}$ on $(\Omega, \mathcal{F})$ and the associated expected value $\tilde{\mathbb{E}}$. Under $\tilde{\mathbb{P}}$, we have

$$dX_s^{t,x} = \left(1 - \frac{X_s^{t,x}}{1-s} \right) ds + d\tilde{W}_s, \quad s \in [t, 1], \quad (\text{II.4.3})$$

with $X_t^{t,x} = x$, and with $\tilde{W}_t := W_t - t$ defining a $\tilde{\mathbb{P}}$ -Brownian motion by Girsanov's theorem.

Thus, by (II.4.1), we obtain the equivalent stopping problem

$$\tilde{v}(t, x) := \sup_{0 \leq \tau \leq 1-t} \tilde{\mathbb{E}} \left[\exp \left(\int_0^\tau \left(\frac{1}{2} - \frac{X_{t+s}^{t,x}}{1-(t+s)} \right) ds \right) \right], \quad (\text{II.4.4})$$

where $v(t, x) = e^x \tilde{v}(t, x)$. Consequently, we also have that

$$\mathcal{C} = \{(t, x) \in [0, 1] \times \mathbb{R} : \tilde{v}(t, x) > 1\}. \quad (\text{II.4.5})$$

Moreover, since V is a $\mathbb{P}$ -supermartingale and V^* is a $\mathbb{P}$ -martingale then, as a consequence of Girsanov theorem, the process $\tilde{V} := (\tilde{V}_t)_{t \in [0, 1]}$ defined as

$$\tilde{V}_t := \exp \left(\int_0^t \left(\frac{1}{2} - \frac{X_s}{1-s} \right) ds \right) \tilde{v}(t, X_t) \quad (\text{II.4.6})$$

is a $\tilde{\mathbb{P}}$ -supermartingale and $\tilde{V}^* := (\tilde{V}_{t \wedge \tau^*})_{t \in [0, 1]}$ is a $\tilde{\mathbb{P}}$ -martingale, with τ^* as in (II.2.11).

In this formulation the set $\mathcal{C}$ is expressed in terms of the value function $\tilde{v}(t, x)$ only and not explicitly in terms of the exponential function e^x . This allows us to prove the existence of an optimal boundary for our problem.

Proposition II.4.1. *There exists a function $b : [0, 1] \rightarrow \mathbb{R}_+ \cup \{+\infty\}$ such that*

$$\mathcal{C} = \{(t, x) \in [0, 1] \times \mathbb{R} : x < b(t)\}. \quad (\text{II.4.7})$$

Proof. By the pathwise uniqueness of the Brownian bridge (II.2.2), it is clear that for any $x \leq x'$ we have, $\mathbb{P}$ -a.s. (hence, also $\tilde{\mathbb{P}}$ -a.s.)

$$X_s^{t,x} \leq X_s^{t,x'}, \quad \forall s \in [t, 1]. \quad (\text{II.4.8})$$

Using such comparison principle and (II.4.4), it is easy to show that $x \mapsto \tilde{v}(t, x)$ is non-increasing. It follows, in particular, that if $(t, x) \in \mathcal{D}$, then $(t, x') \in \mathcal{D}$ for all $x' \geq x$. Then, setting $b(1) := 0$, we define

$$\begin{aligned} b(t) &:= \sup\{x \in \mathbb{R} : \tilde{v}(t, x) > 1\} \\ &= \sup\{x \in \mathbb{R} : v(t, x) > e^x\}, \quad t \in [0, 1), \end{aligned} \quad (\text{II.4.9})$$

and (II.4.7) holds by continuity of the value function. For future reference, notice that (II.4.9) and (II.3.5)–(II.3.6) give also $b(1) = 0$.

It remains to show that $b(t) \geq 0$ for all $t \in [0, 1]$. By selecting the stopping rule $\tau = 1 - t$, we obtain $v(t, x) \geq 1 > e^x$ for every $x < 0$ and every $t \in [0, 1)$. Hence,

$$[0, 1) \times (-\infty, 0) \subset \mathcal{C}, \quad (\text{II.4.10})$$

and the claim follows. $\square$

As a straightforward consequence of this proposition and of (II.2.11), we have that

$$\tau_{t,x}^* = \inf\{s \in [0, 1-t] : X_{t+s}^{t,x} \geq b(t+s)\}, \quad \mathbb{P}\text{-a.s.} \quad (\text{II.4.11})$$

II.5 Regularity of the optimal boundary

In this section we show that the optimal boundary is monotonic, continuous and bounded. We will then use these properties to derive smoothness of the value function across the optimal boundary, in the next section.

We first show a simple preliminary result. Let

$$\mathcal{Q} := \{(t, x) \in [0, 1) \times \mathbb{R} : x < \tfrac{1}{2}(1 - t)\}, \quad (\text{II.5.1})$$

$$\mathcal{R} := \{(t, x) \in [0, 1) \times \mathbb{R} : x = \tfrac{1}{2}(1 - t)\}. \quad (\text{II.5.2})$$

Fix $(t, x) \in \mathcal{Q}$, and notice that

$$\delta := \inf_{(s, y) \in \mathcal{R}} |(t, x) - (s, y)| > 0. \quad (\text{II.5.3})$$

Then, let $B_\delta := \{(s, y) \in [0, 1) \times \mathbb{R} : |(t, x) - (s, y)| \leq \delta/2\}$ and $\tau_\delta := \inf\{s \geq t : X_s^{t, x} \notin B_\delta\} \wedge 1$. By continuity of X , we have that $\tau_\delta > t$, $\mathbb{P}$ -a.s. Hence, by Dynkin's formula, we obtain

$$\mathbb{E} \left[e^{X_{\tau_\delta}^{t, x}} \right] = e^x + \mathbb{E} \left[\int_t^{\tau_\delta} e^{X_s^{t, x}} \left(\frac{1}{2} - \frac{X_s^{t, x}}{1 - s} \right) ds \right] > e^x. \quad (\text{II.5.4})$$

Thus, if we start inside $\mathcal{Q}$ then stopping immediately is suboptimal (we would rather wait at least until τ_δ) and so $\mathcal{Q} \subseteq \mathcal{C}$.

The next result, concerning monotonicity of the optimal boundary, is crucial for the subsequent analysis of the stopping set and of the value function. Monotonicity of optimal boundaries is relatively easy to establish in optimal stopping problems when the underlying diffusion is time-homogeneous and the gain function is independent of time. In our case, the latter condition holds but our diffusion is time-dependent, hence new ideas are needed in the proof of the theorem below. We also remark that, while in some stopping problems of a Brownian bridge (see, e.g., [41]) it is possible to rely upon a time-change in order to formulate an auxiliary equivalent stopping problem for a time-homogeneous diffusion (see [77]), this is not the case here, due to the exponential nature of the gain function.

Theorem II.5.1. *The optimal boundary $t \mapsto b(t)$ is non-increasing on $[0, 1]$.*

Proof. It is sufficient to show that, for every fixed $x \in \mathbb{R}$, the map $t \mapsto v(t, x)$ is non-increasing on $[0, 1)$. Indeed, the latter implies monotonicity of the boundary on $[0, 1]$ by definition (II.4.9) and using that $b(t) \geq 0$ for all $t \in [0, 1)$ and $b(1) = 0$.

Recalling (II.3.9) and using convexity of $x \mapsto v(t, x)$, we obtain

$$\partial_t v(t, x) \leq \frac{x}{1 - t} \partial_x v(t, x), \quad \forall (t, x) \in \mathcal{C}. \quad (\text{II.5.5})$$

In particular, we have

$$\partial_t v(t, x) \leq 0, \quad \forall (t, x) \in [0, 1) \times (-\infty, 0], \quad (\text{II.5.6})$$

since $[0, 1) \times (-\infty, 0] \subseteq \mathcal{Q} \subseteq \mathcal{C}$ (see (II.5.1)) and since $\partial_x v \geq 0$ in $\mathcal{C}$ (by Proposition II.3.1).

Notice that if $(t, x) \in \mathcal{D} \setminus \partial \mathcal{C}$ then $v(t, x) = e^x$ and $\partial_t v(t, x) = 0$. Since $t \mapsto v(t, x)$ is continuous on

$[0, 1)$ (by Proposition II.3.2), it only remains to prove that $\partial_t v(t, x) \leq 0$ for $(t, x) \in \mathcal{C}$ with $x > 0$. To show this, we proceed in two steps.

Step 1. (Property of $t \mapsto X^{t,x}$). Consider $(t, x) \in \mathcal{C}$ with $x > 0$ and $0 < \varepsilon \leq t < 1$, for some $\varepsilon > 0$. For $s \in [0, 1 - t]$, we denote

$$Y_{t+s}^{t,x;\varepsilon} := X_{t+s}^{t,x} - X_{t-\varepsilon+s}^{t-\varepsilon,x}. \quad (\text{II.5.7})$$

Since (t, x) is fixed, we simplify the notation and set $Y_{t+s}^\varepsilon := Y_{t+s}^{t,x;\varepsilon}$, for $s \in [0, 1 - t]$. Next, for some small $\delta > 0$, let $t_\delta := (1 - t - \delta) > 0$ and $\rho_\delta := t_\delta \wedge \tau^0$, where

$$\tau^0 := \tau_{t,x}^0 := \inf\{u \in [0, 1 - t] : X_{t+u}^{t,x} \leq 0\}. \quad (\text{II.5.8})$$

Then, by (II.2.20), for an arbitrary $s \in [0, 1 - t]$ we have, $\mathbb{P}$ -a.s.

$$\begin{aligned} Y_{t+s \wedge \rho_\delta}^\varepsilon &= - \int_0^{s \wedge \rho_\delta} \frac{X_{t+u}^{t,x}}{1 - (t + u)} du + \int_0^{s \wedge \rho_\delta} \frac{X_{t-\varepsilon+u}^{t-\varepsilon,x}}{1 - (t - \varepsilon + u)} du \\ &= - \int_0^{s \wedge \rho_\delta} \left(\frac{\varepsilon X_{t+u}^{t,x}}{(1 - (t - \varepsilon + u))(1 - (t + u))} + \frac{Y_{t+u}^\varepsilon}{1 - (t - \varepsilon + u)} \right) du. \end{aligned} \quad (\text{II.5.9})$$

Since Y^ε is a continuous process of bounded variation then its local time is identically zero. Moreover $Y_t^\varepsilon = 0$ and thus, from Tanaka's formula, we obtain

$$[Y_{t+s \wedge \rho_\delta}^\varepsilon]^+ = \int_0^{s \wedge \rho_\delta} \mathbf{1}_{\{Y_{t+u}^\varepsilon \geq 0\}} dY_{t+u}^\varepsilon \leq 0, \quad (\text{II.5.10})$$

where the final inequality follows from (II.5.9), upon observing that $X_{t+u}^{t,x} \geq 0$ for all $u \leq \rho_\delta$. Then, $Y_{t+s \wedge \rho_\delta}^\varepsilon \leq 0$ for all $s \in [0, 1 - t]$. Furthermore, letting $\delta \rightarrow 0$, we obtain by continuity of paths

$$X_{t-\varepsilon+s \wedge \tau^0}^{t-\varepsilon,x} \geq X_{t+s \wedge \tau^0}^{t,x} \geq 0, \quad \forall s \in [0, 1 - t], \mathbb{P}\text{-a.s.} \quad (\text{II.5.11})$$

Hence, the process $X^{t,x}$ hits zero before the process $X^{t-\varepsilon,x}$ does.

Step 2. ($\partial_t v(t, x) \leq 0$). Fix $(t, x) \in \mathcal{C}$ with $x > 0$. Using the same notation as in Step 1 above, let $\sigma := \tau_{t,x}^* \wedge \tau_{t,x}^0$. By the (super)martingale property of the value function, noticing that τ^* is optimal for $v(t, x)$ and sub-optimal for $v(t - \varepsilon, x)$, we have

$$\begin{aligned} v(t, x) - v(t - \varepsilon, x) &\leq \mathbb{E} [v(t + \sigma, X_{t+\sigma}^{t,x}) - v(t - \varepsilon + \sigma, X_{t-\varepsilon+\sigma}^{t-\varepsilon,x})] \\ &\leq \mathbb{E} [\mathbf{1}_{\{\tau^* \leq \tau^0\} \cap \{\tau^* < 1-t\}} (\exp(X_{t+\tau^*}^{t,x}) - \exp(X_{t-\varepsilon+\tau^*}^{t-\varepsilon,x}))] \\ &\quad + \mathbb{E} [\mathbf{1}_{\{\sigma = 1-t\}} (\exp(X_1^{t,x}) - \exp(X_{1-\varepsilon}^{t-\varepsilon,x}))] \\ &\quad + \mathbb{E} [\mathbf{1}_{\{\tau^0 < \tau^*\} \cap \{\tau^0 < 1-t\}} (v(t + \tau^0, 0) - v(t - \varepsilon + \tau^0, X_{t-\varepsilon+\tau^0}^{t-\varepsilon,x}))]. \end{aligned} \quad (\text{II.5.12})$$

Recalling (II.5.11), on the event $\{\tau^* \leq \tau^0\} \cap \{\tau^* < 1 - t\}$ we have $X_{t-\varepsilon+\tau^*}^{t-\varepsilon,x} \geq X_{t+\tau^*}^{t,x}$ and on the event $\{\sigma = 1 - t\}$ we have that $X_{1-\varepsilon}^{t-\varepsilon,x} \geq X_1^{t,x}$. Moreover, $x \mapsto v(t, x)$ is non-decreasing (by Proposition II.3.1). Thus, combining these facts with (II.5.12), we obtain

$$v(t, x) - v(t - \varepsilon, x) \leq \mathbb{E} [\mathbf{1}_{\{\tau^0 < \tau^*\} \cap \{\tau^0 < 1-t\}} (v(t + \tau^0, 0) - v(t - \varepsilon + \tau^0, 0))] \leq 0, \quad (\text{II.5.13})$$

where the final inequality uses (II.5.6) and the fact that $\tau^0 < 1 - t$.

Finally, dividing both sides of (II.5.13) by ε and letting $\varepsilon \rightarrow 0$, we obtain $\partial_t v(t, x) \leq 0$ as needed. $\square$

By standard arguments in optimal stopping theory, monotonicity of the boundary leads to its right-continuity (or left-continuity). In our case we have a simple corollary.

Corollary II.5.2. *The optimal boundary $t \mapsto b(t)$ is right-continuous, whenever finite.*

Proof. Let $t \in [0, 1)$ be such that $b(t) < +\infty$. Consider a sequence $(t_n)_{n \in \mathbb{N}}$ such that $t_n \downarrow t$ as $n \rightarrow \infty$. By monotonicity of $t \mapsto b(t)$ (see Theorem II.5.1) and (II.4.7), we have that $b(t_n) < \infty$ and $(t_n, b(t_n)) \in \mathcal{D}$ for every $n \in \mathbb{N}$. Since $\mathcal{D}$ is a closed set and $(t_n, b(t_n)) \rightarrow (t, b(t+))$, then also $(t, b(t+)) \in \mathcal{D}$ (the right-limit $b(t+)$ exists by monotonicity). Hence, $b(t+) \geq b(t)$ (see (II.4.7)). However, by monotonicity $b(t) \geq b(t+)$, which leads to $b(t) = b(t+)$. $\square$

We can now show that the optimal boundary is continuous and bounded on $[0, 1]$.

Proposition II.5.3. *The optimal boundary $t \mapsto b(t)$ is continuous on $[0, 1]$ and we have*

$$\sup_{t \in [0, 1]} b(t) < +\infty. \quad (\text{II.5.14})$$

The proof of Proposition II.5.3 relies upon four lemmas. First we state and prove those lemmas and then we prove the proposition.

Lemma II.5.4. *For every $t \in [0, 1)$, we have $\mathcal{D} \cap ([t, 1) \times \mathbb{R}) \neq \emptyset$.*

Proof. Suppose by contradiction that this is not true and there exists $t \in [0, 1)$ such that $\mathcal{D} \cap ([t, 1) \times \mathbb{R}) = \emptyset$. Then $\tau_{t', x}^* = 1 - t'$, $\mathbb{P}$ -a.s. for all $(t', x) \in [t, 1) \times \mathbb{R}$, which implies $v(t', x) = 1$. This, however, contradicts the optimality of $\tau_{t', x}^*$ since immediate stopping gives $v(t', x) \geq e^x > 1$ for every $x > 0$ and $t' \in [t, 1)$. $\square$

Notice that the Corollary II.5.2 and Lemma II.5.4 imply that for every $t_1 \in [0, 1)$ there exists $t_2 \in (t_1, 1)$ such that $b(t_2) < +\infty$. This fact will be used in the next lemma.

Lemma II.5.5. *The boundary satisfies $b(t) < +\infty$ for every $t \in (0, 1]$.*

Proof. By contradiction, let us assume that there is $t \in (0, 1)$ such that $b(t) = +\infty$. Then, by Lemma II.5.4 and Corollary II.5.2, we can find $t' \in (t, 1)$ such that $0 \leq b(t') =: b_0 < +\infty$ and $(t, t') \times \mathbb{R} \subseteq \mathcal{C}$. Let $\sigma_0 := \inf\{s \in [0, 1 - t) : X_{t+s}^{t, x} \leq b_0\} \wedge (t' - t)$, then recalling $\tau_{t, x}^*$ as in (II.4.11), we immediately see that $\mathbb{P}(\tau_{t, x}^* \geq \sigma_0) = 1$. Using the martingale property of the value function (see (II.4.6)), we obtain

$$\begin{aligned} \tilde{v}(t, x) &= \tilde{\mathbb{E}} \left[\exp \left(\int_0^{\sigma_0} \left(\frac{1}{2} - \frac{X_{t+s}^{t, x}}{1 - (t + s)} \right) ds \right) \tilde{v}(t + \sigma_0, X_{t+\sigma_0}^{t, x}) \right] \\ &= \tilde{\mathbb{E}} \left[\mathbf{1}_{\{\sigma_0 < t' - t\}} \exp \left(\int_0^{\sigma_0} \left(\frac{1}{2} - \frac{X_{t+s}^{t, x}}{1 - (t + s)} \right) ds \right) \tilde{v}(t + \sigma_0, b_0) \right] \\ &\quad + \tilde{\mathbb{E}} \left[\mathbf{1}_{\{\sigma_0 = t' - t\}} \exp \left(\int_0^{t' - t} \left(\frac{1}{2} - \frac{X_{t+s}^{t, x}}{1 - (t + s)} \right) ds \right) \cdot 1 \right], \end{aligned} \quad (\text{II.5.15})$$

where we have used continuity of paths and the fact that on $\{\sigma_0 = t' - t\}$ it must be $X_{t'} \geq b(t') = b_0$, $\tilde{\mathbb{P}}$ -a.s.

Moreover, since $X_{t+s}^{t,x} \geq b_0$ for $s \leq \sigma_0$, we have

$$\begin{aligned}
 \tilde{v}(t, x) &\leq \tilde{\mathbb{E}} \left[\mathbf{1}_{\{\sigma_0 < t' - t\}} \exp \left(\int_0^{\sigma_0} \left(\frac{1}{2} - \frac{b_0}{1 - (t+s)} \right) ds \right) \tilde{v}(t + \sigma_0, b_0) \right] \\
 &\quad + \tilde{\mathbb{E}} \left[\mathbf{1}_{\{\sigma_0 = t' - t\}} \exp \left(\int_0^{t' - t} \left(\frac{1}{2} - \frac{X_{t+s}^{t,x} \vee b_0}{1 - (t+s)} \right) ds \right) \right] \\
 &\leq \tilde{\mathbb{E}} \left[\mathbf{1}_{\{\sigma_0 < t' - t\}} \left(\frac{1 - (t + \sigma_0)}{1 - t} \right)^{b_0} e^{\sigma_0/2} \right] \cdot c_1 \\
 &\quad + \tilde{\mathbb{E}} \left[\mathbf{1}_{\{\sigma_0 = t' - t\}} \exp \left(\int_0^{t' - t} \left(\frac{1}{2} - \frac{X_{t+s}^{t,x} \vee b_0}{1 - (t+s)} \right) ds \right) \right] \\
 &\leq c_1 e^{1/2} \tilde{\mathbb{P}}(\sigma_0 < t' - t) + \tilde{\mathbb{E}} \left[\exp \left(\int_0^{t' - t} \left(\frac{1}{2} - \frac{X_{t+s}^{t,x} \vee b_0}{1 - (t+s)} \right) ds \right) \right], \tag{II.5.16}
 \end{aligned}$$

where in the second inequality we have used (II.3.1) and $\tilde{v}(t, x) = e^{-x} v(t, x)$. Now, notice that

$$\tilde{\mathbb{P}}(\sigma_0 < t' - t) \leq \tilde{\mathbb{P}} \left(\inf_{s \in [t, t']} X_s^{t,x} \leq b_0 \right) \tag{II.5.17}$$

so that the first term on the right-hand side of (II.5.16) goes to zero as $x \rightarrow \infty$. Similarly, given that $\lim_{x \rightarrow \infty} X_{t+s}^{t,x} = +\infty$, $\mathbb{P}$ -a.s. for every $s \in [0, t' - t]$, the second term goes to zero as well by the reverse Fatou's lemma. Then, recalling that $\tilde{v} \geq 1$, we reach the contradiction

$$\limsup_{x \rightarrow +\infty} \tilde{v}(t, x) \leq 0. \tag{II.5.18}$$

It follows that $b(t) < +\infty$ for all $t \in (0, 1]$ since by definition $b(1) = 0$. $\square$

Lemma II.5.6. *We have $b(0) < +\infty$.*

Proof. Consider an auxiliary problem where the Brownian bridge is pinned at time $1 + h$, for some $h > 0$, and the time horizon of the optimisation is $1 + h$. That is, let us set

$$v^h(t, x) := \sup_{0 \leq \tau \leq 1+h-t} \mathbb{E} \left[e^{\hat{X}_{t+\tau}^{t,x}} \right], \tag{II.5.19}$$

where $\hat{X}$ is a Brownian bridge (II.2.2) pinned at time $1 + h$.

By the same argument as in Section II.2, it follows that $\mathcal{L}(\hat{X}^{t,x}) = \mathcal{L}(\hat{Z}^{t,x})$, where

$$\hat{Z}_s^{t,x} = \frac{1+h-s}{1+h-t} x + \sqrt{\frac{1+h-s}{1+h-t}} W_{s-t}, \quad s \in [t, 1+h]. \tag{II.5.20}$$

Thus,

$$v^h(t, x) = \sup_{0 \leq \tau \leq 1+h-t} \mathbb{E} \left[e^{\hat{Z}_{t+\tau}^{t,x}} \right] \tag{II.5.21}$$

and, since $\mathcal{L}(Z_s^{t,x}, s \in [t, 1]) = \mathcal{L}(\hat{Z}_{s+h}^{t+h,x}, s \in [t, 1])$ (compare (II.2.14) with (II.5.21)), we also have

that

$$v(t, x) = v^h(t + h, x), \quad \forall (t, x) \in [0, 1] \times \mathbb{R}. \quad (\text{II.5.22})$$

By the same arguments as for the original problem, we obtain that there exists a non-decreasing, right-continuous optimal boundary $t \mapsto b^h(t)$ such that

$$\mathcal{C}^h := \{(t, x) \in [0, 1 + h] \times \mathbb{R} : v^h(t, x) > e^x\} = \{(t, x) \in [0, 1 + h] \times \mathbb{R} : x < b^h(t)\}. \quad (\text{II.5.23})$$

Moreover, since the gain function e^x does not depend on time, using (II.5.22) we obtain

$$b(t) = b^h(t + h), \quad \forall t \in [0, 1]. \quad (\text{II.5.24})$$

In particular, $b(0) = b^h(h)$ and $b^h(h) < +\infty$ by applying the result in Lemma II.5.5 to the auxiliary problem. $\square$

Using ideas as in [28], we can also prove left-continuity of the optimal boundary.

Lemma II.5.7. *The optimal boundary $t \mapsto b(t)$ is left-continuous.*

Proof. We first prove that the boundary is left-continuous for every $t \in (0, 1)$ and then that its left limit at $t = 1$ is zero, that is $b(1-) = 0 = b(1)$.

Suppose, by contradiction, that there exists $t_0 \in (0, 1)$ such that $b(t_0-) > b(t_0)$ and consider an interval $[x_1, x_2] \subset (b(t_0), b(t_0-))$. By monotonicity of $t \mapsto b(t)$, we have that $[0, t_0) \times [x_1, x_2] \subset \mathcal{C}$. Now, pick an arbitrary, non-negative smooth function φ with compact support $[x_1, x_2]$. Since (II.3.9) holds in $[0, t_0) \times [x_1, x_2]$, then for every $t < t_0$ we have

$$\begin{aligned} 0 &= \int_{x_1}^{x_2} [\partial_t v(t, y) + \mathcal{A}v(t, y)] \varphi(y) dy \\ &\leq \int_{x_1}^{x_2} \mathcal{A}v(t, y) \varphi(y) dy = \int_{x_1}^{x_2} v(t, y) (\mathcal{A}^* \varphi)(t, y) dy, \end{aligned} \quad (\text{II.5.25})$$

where for the inequality we have used $\partial_t v \leq 0$ (see proof of Proposition II.5.1) and in the final equality we have applied integration by parts and used the adjoint operator

$$(\mathcal{A}^* \varphi)(t, y) := \frac{1}{2} \frac{d^2 \varphi}{dy^2}(y) + \frac{1}{1-t} \cdot \frac{d}{dy}(y \cdot \varphi(y)). \quad (\text{II.5.26})$$

Taking limits as $t \uparrow t_0$ and using dominated convergence theorem, we obtain

$$\begin{aligned} 0 &\leq \lim_{t \uparrow t_0} \int_{x_1}^{x_2} v(t, y) (\mathcal{A}^* \varphi)(t, y) dy = \int_{x_1}^{x_2} v(t_0, y) (\mathcal{A}^* \varphi)(t_0, y) dy \\ &= \int_{x_1}^{x_2} e^y (\mathcal{A}^* \varphi)(t_0, y) dy = \int_{x_1}^{x_2} e^y \left(\frac{1}{2} - \frac{y}{1-t_0} \right) \varphi(y) dy, \end{aligned} \quad (\text{II.5.27})$$

where we have used that $v(t_0, y) = e^y$ for $y \in [x_1, x_2]$ and integration by parts in the final equality.

Recalling that $x_1 \geq b(t_0) > \frac{1-t_0}{2}$, then (II.5.27) leads to a contradiction because the right-hand side of the expression is strictly negative.

In order to prove that $b(1-) = b(1) = 0$, we need a slight modification of the argument above. In particular, suppose by contradiction that $b(1-) > 0$ and consider an interval $[x_1, x_2] \subset (0, b(1-))$. Then, replacing φ in (II.5.25) with $\tilde{\varphi}(t, x) := (1 - t)\varphi(x)$, and using the same arguments with $t_0 = 1$ we reach a contradiction, i.e.,

$$0 \leq \lim_{t \uparrow 1} \int_{x_1}^{x_2} v(t, y) (\mathcal{A}^* \tilde{\varphi})(t, y) dy = \int_{x_1}^{x_2} e^y \frac{d}{dy} (y \cdot \varphi(y)) dy = - \int_{x_1}^{x_2} e^y y \varphi(y) dy < 0. \quad (\text{II.5.28})$$

□

We are now able to prove Proposition II.5.3.

Proof of Proposition II.5.3. The proof of (II.5.14) follows immediately from Lemma II.5.5 and Lemma II.5.6. Right-continuity of the boundary follows from Corollary II.5.2 and (II.5.14), whereas left-continuity follows from Lemma II.5.7. Thus, the optimal boundary is bounded and continuous on $[0, 1]$. □

II.6 Further regularity of the value function

In this section we show further regularity for the value function. In particular, we prove that the value function is smooth, i.e., $v \in C^1([0, 1] \times \mathbb{R})$ (see Theorem II.6.4) and that the second derivative $\partial_{xx} v$ is continuous but for a jump across the optimal boundary (see Corollary II.6.5).

In order to obtain these results, we first show that the optimal stopping time is continuous with respect to the starting point, i.e., that the map $(t, x) \mapsto \tau_{t,x}^*$ is continuous. The monotonicity of the optimal boundary is essential to obtain this result, which we prove by following similar arguments as in [29, Section 5].

For $(t, x) \in [0, 1] \times \mathbb{R}$, we define

$$\tau'_{t,x} := \inf\{s \geq 0 : X_{t+s}^{t,x} > b(t+s)\} \wedge (1-t). \quad (\text{II.6.1})$$

Then, we have the following results.

Lemma II.6.1. *For every $(t, x) \in [0, 1] \times \mathbb{R}$, we have $\tau'_{t,b(t)} = 0$ and $\tau'_{t,x} = \tau_{t,x}^*$, $\mathbb{P}$ -a.s.*

Proof. We first prove that $\tau'_{t,b(t)} = 0$, $\mathbb{P}$ -a.s. If $t = 1$, this is obvious by definition (II.6.1). If $t < 1$, we equivalently prove that for every $\varepsilon > 0$ there exists $t_\varepsilon \in [t, t + \varepsilon)$ such that $X_{t_\varepsilon}^{t,b(t)} > b(t_\varepsilon)$, $\mathbb{P}$ -a.s. Let us fix $\omega \in \Omega \setminus N$, where $\mathbb{P}(N) = 0$ and the law of iterated logarithm holds for W on $\Omega \setminus N$. Let $(h_n)_{n \in \mathbb{N}}$ be a sequence such that $h_n \downarrow 0$ and (without loss of generality) $h_n < 1 - t$ for every $n \in \mathbb{N}$. The law of iterated logarithm (see, e.g., [6, Theorem 3.2]) implies that, for every $\delta \in (0, 1)$,

$$W_{t+h_n} - W_t \geq (1 - \delta) \sqrt{2h_n \log(\log(1/h_n))}, \quad (\text{II.6.2})$$

for n sufficiently large. By (II.2.1), we have that

$$X_{t+h_n}^{t,b(t)} = b(t) - \int_t^{t+h_n} \frac{X_s^{t,b(t)}}{1-s} ds + W_{t+h_n} - W_t, \quad (\text{II.6.3})$$

Then, by (II.6.2), we obtain

$$X_{t+h_n}^{t,b(t)} \geq b(t) - \int_t^{t+h_n} \frac{X_s^{t,b(t)}}{1-s} ds + (1-\delta)\sqrt{2h_n \log(\log(1/h_n))}. \quad (\text{II.6.4})$$

By defining $M := \max_{s \in [t, t+h_1]} (X_s/(1-s))$ and by monotonicity of $t \mapsto b(t)$, we have that

$$X_{t+h_n}^{t,b(t)} \geq b(t+h_n) - Mh_n + (1-\delta)\sqrt{2h_n \log(\log(1/h_n))}, \quad (\text{II.6.5})$$

Finally, we can choose n sufficiently large such that

$$(1-\delta)\sqrt{2h_n \log(\log(1/h_n))} - Mh_n > 0 \quad \text{and} \quad h_n < \varepsilon, \quad (\text{II.6.6})$$

so that $X_{t_\varepsilon}^{t,b(t)} > b(t_\varepsilon)$ with $t_\varepsilon = t + h_n \in [t, t + \varepsilon)$, which is our desired result.

We now prove that $\tau'_{t,x} = \tau_{t,x}^*$, $\mathbb{P}$ -a.s. Let us assume that $(t, x) \in [0, 1] \times \mathbb{R}$ with $x \leq b(t)$, otherwise the claim is trivial. Since $(t, x) \in [0, 1] \times \mathbb{R}$ is fixed, we simply denote $\tau_{t,x}^*$ and $\tau'_{t,x}$ by τ^* and τ' , respectively. Clearly, we have that $\tau' \geq \tau^*$, $\mathbb{P}$ -a.s. Hence, we need to prove that $\tau' \leq \tau^*$, $\mathbb{P}$ -a.s.

Let $0 \leq s_1 < s_2 \leq 1 - t$ and define $Z_{s_1, s_2} := \sup_{s_1 \leq s \leq s_2} (X_{t+s}^{t,x} - b(t+s))$. We now want to show that

$$\mathbb{P}(Z_{s_1, s_2} = 0) = 0. \quad (\text{II.6.7})$$

To this purpose, let $\tau_1 := \inf\{s \in [s_1, 1-t] : X_{t+s}^{t,x} = b(t+s)\}$ and notice that

$$\mathbb{P}(Z_{s_1, s_2} = 0) = \mathbb{P}(Z_{s_1, s_2} = 0, \tau_1 \in [s_1, s_2]) + \mathbb{P}(Z_{s_1, s_2} = 0, \tau_1 = s_2), \quad (\text{II.6.8})$$

since on the event $\{Z_{s_1, s_2} = 0\}$ it must happen that $\tau_1 \in [s_1, s_2]$.

For the first term on the right-hand side of expression (II.6.8), we have that

$$\begin{aligned} \mathbb{P}(Z_{s_1, s_2} = 0, \tau_1 \in [s_1, s_2]) &= \mathbb{P}(Z_{s_1, s_2} = 0, \tau_1 \in [s_1, s_2], X_{t+s_1}^{t,x} \leq b(t+s_1)) \\ &= \mathbb{P}(Z_{s_1, \tau_1} \vee Z_{\tau_1, s_2} = 0, \tau_1 \in [s_1, s_2], X_{t+s_1}^{t,x} \leq b(t+s_1)) \\ &= \mathbb{P}(Z_{\tau_1, s_2} = 0, \tau_1 \in [s_1, s_2], X_{t+s_1}^{t,x} \leq b(t+s_1)) \\ &\leq \mathbb{P}(Z_{\tau_1, s_2} = 0, \tau_1 \in [s_1, s_2]) \\ &= 0, \end{aligned} \quad (\text{II.6.9})$$

where in the second equality we have used the continuity of $s \mapsto X_s^{t,x}$ and of $s \mapsto b(s)$ and in the last equality we have used that $\tau'_{t,b(t)} = 0$, $\mathbb{P}$ -a.s.

For the second term on the right-hand side of (II.6.8), we simply have that

$$\mathbb{P}(Z_{s_1, s_2} = 0, \tau_1 = s_2) \leq \mathbb{P}(X_{t+s_2}^{t,x} = b(t+s_2)) = 0, \quad (\text{II.6.10})$$

as $X_{t+s_2}^{t,x}$ is a continuous random variable. This completes the proof of (II.6.7).

Finally, if $\tau^* \in [s_1, s_2]$, then $Z_{s_1, s_2} \geq X_{t+\tau^*}^{t,x} - b(t+\tau^*) = 0$, and so

$$\begin{aligned} \mathbb{P}(\tau^* \in [s_1, s_2]) &= \mathbb{P}(\tau^* \in [s_1, s_2], Z_{s_1, s_2} \geq 0) \\ &= \mathbb{P}(\tau^* \in [s_1, s_2], Z_{s_1, s_2} > 0) \\ &= \mathbb{P}(\tau^* \in [s_1, s_2], \tau' \in [s_1, s_2]), \end{aligned} \quad (\text{II.6.11})$$

where in the second equality we used what we have proved in (II.6.7).

Therefore, for every $s_1, s_2 \in \mathbb{Q}$ such that $s_1 \leq s_2$, we have that

$$\mathbb{P}(s_1 \leq \tau^* \leq s_2, \tau' > s_2) = 0, \quad (\text{II.6.12})$$

which leads to our desired result, i.e., $\tau' \leq \tau^*$, $\mathbb{P}$ -a.s. and the proof is complete. $\square$

As a consequence of Lemma II.6.1, we deduce

$$\tau_{t,x}^* = \inf\{s \in [0, 1-t] : (t+s, X_{t+s}^{t,x}) \in \text{int}(\mathcal{D})\}, \quad \mathbb{P}\text{-a.s.}, \quad (\text{II.6.13})$$

where $\text{int}(\mathcal{D}) = \mathcal{D} \setminus \partial\mathcal{C}$ is the interior of the stopping set. In particular, since $\tau_{t,x}^* = 0$, $\mathbb{P}$ -a.s. for every $(t, x) \in \partial\mathcal{C}$, by its definition (II.2.11), this implies $\tau'_{t,x} = 0$, $\mathbb{P}$ -a.s. for every $(t, x) \in \partial\mathcal{C}$. This means that the boundary $\partial\mathcal{C}$ is regular for the interior of the stopping set in the sense of diffusions (see, e.g., [33]).

Lemma II.6.1 then leads to the continuity of the optimal stopping time.

Proposition II.6.2. *The function $(t, x) \mapsto \tau_{t,x}^*$ is continuous, $\mathbb{P}$ -a.s., on $[0, 1] \times \mathbb{R}$.*

Proof. Let $(t, x) \in [0, 1] \times \mathbb{R}$ and $\{(t_n, x_n)\}_{n \in \mathbb{N}} \subseteq [0, 1] \times \mathbb{R}$ such that $(t_n, x_n) \rightarrow (t, x)$ as $n \rightarrow \infty$. If $t = 1$, the claim is easy as $\tau_{t,x}^* = 0$, $\mathbb{P}$ -a.s. and $|\tau_{t_n, x_n}^* - \tau_{t,x}^*| \leq 1 - t_n \rightarrow 0$ as $n \rightarrow \infty$, $\mathbb{P}$ -a.s.

For $t < 1$, we will show that $\lim_{n \rightarrow \infty} \tau_{t_n, x_n}^* = \tau_{t,x}^*$ by first showing that $\limsup_{n \rightarrow \infty} \tau_{t_n, x_n}^* \leq \tau_{t,x}^*$ and then showing that $\liminf_{n \rightarrow \infty} \tau_{t_n, x_n}^* \geq \tau_{t,x}^*$. We consider $\omega \in \Omega$ fixed but we do not specify it every time for the sake of simplicity.

To prove the first inequality, we have that either $X_{t+s}^{t,x} < b(t+s)$ for every $s \in [0, 1-t]$ or there exists $s \in [0, 1-t]$ such that $X_{t+s}^{t,x} \geq b(t+s)$. In the former case, we obtain that $\tau_{t,x}^* = 1-t$ and so

$$\limsup_{n \rightarrow \infty} \tau_{t_n, x_n}^* \leq \limsup_{n \rightarrow \infty} (1 - t_n) = 1 - t = \tau_{t,x}^*. \quad (\text{II.6.14})$$

In the latter case, by Lemma II.6.1, there exists $s' \in [0, 1-t]$ such that $X_{t+s'}^{t,x} > b(t+s')$. Then, continuity of $(t, x) \mapsto X_{t+s'}^{t,x}$ and $t \mapsto b(t+s')$ yields $X_{t_n+s'}^{t_n, x_n} > b(t_n+s')$ for n sufficiently large. Hence, $\tau'_{t_n, x_n} \leq s'$ and so $\limsup_{n \rightarrow \infty} \tau'_{t_n, x_n} \leq s'$. Since $s' \in [0, 1-t]$ such that $X_{t+s'}^{t,x} > b(t+s')$ was arbitrary, we obtain that $\limsup_{n \rightarrow \infty} \tau'_{t_n, x_n} \leq \tau'_{t,x}$. Therefore, by Lemma II.6.1, we have that

$$\limsup_{n \rightarrow \infty} \tau_{t_n, x_n}^* = \limsup_{n \rightarrow \infty} \tau'_{t_n, x_n} \leq \tau'_{t,x} = \tau_{t,x}^*, \quad (\text{II.6.15})$$

and the first inequality is proved.

To show the second inequality, first assume that $x \geq b(t)$. Then, $\tau_{t,x}^* = 0$ and so

$$\liminf_{n \rightarrow \infty} \tau_{t_n, x_n}^* \geq 0 = \tau_{t,x}^*. \quad (\text{II.6.16})$$

If $x < b(t)$, then let $s \in [0, 1-t]$ be such that $X_{t+u}^{t,x} < b(t+u)$ for every $u \in [0, s]$. By continuity of $u \mapsto b(t+u)$ and $u \mapsto X_{t+u}^{t,x}$, we have that

$$\inf_{u \in [0, s]} (b(t+u) - X_{t+u}^{t,x}) =: \delta > 0. \quad (\text{II.6.17})$$

Then, by continuity of $x \mapsto X_{t+u}^{t,x}$ uniformly in $u \in [0, s]$, we obtain

$$\inf_{u \in [0, s]} (b(t+u) - X_{t+u}^{t, x_n}) \geq \frac{\delta}{2}, \quad (\text{II.6.18})$$

for n sufficiently large. Furthermore, continuity of $t \mapsto b(t+u)$ and $t \mapsto X_{t+u}^{t,x}$ implies

$$\inf_{u \in [0, s]} (b(t_n+u) - X_{t_n+u}^{t_n, x_n}) \geq \frac{\delta}{4} > 0, \quad (\text{II.6.19})$$

for n large enough. Hence, $\tau_{t_n, x_n}^* \geq s$ and so $\liminf_n \tau_{t_n, x_n}^* \geq s$. Since $s \in [0, 1-t)$ such that $X_{t+u}^{t,x} < b(t+u)$ for every $u \in [0, s]$ was arbitrary, we have that

$$\liminf_{n \rightarrow \infty} \tau_{t_n, x_n}^* \geq \tau_{t,x}^*, \quad (\text{II.6.20})$$

which concludes the proof. $\square$

Now we can use the continuity of the optimal stopping time and some related ideas from [33] to establish C^1 regularity of the value function.

First, we give a lemma concerning the spatial derivative of the value function.

Lemma II.6.3. *For every $(t, x) \in ([0, 1] \times \mathbb{R}) \setminus \partial \mathcal{C}$, we have*

$$\partial_x v(t, x) = \mathbb{E} \left[\frac{1-t-\tau^*}{1-t} e^{Z_{t+\tau^*}^{t,x}} \right]. \quad (\text{II.6.21})$$

Consequently, it also holds

$$\partial_x v(t, x) \leq v(t, x), \quad \forall (t, x) \in ([0, 1] \times \mathbb{R}) \setminus \partial \mathcal{C}. \quad (\text{II.6.22})$$

Proof. Recall that $v(t, x) = e^x$ on $\mathcal{D}$ and so $\partial_x v(t, x) = e^x$ on $\mathcal{D} \setminus \partial \mathcal{C}$. Since $\tau_{t,x}^* = 0$ for every $(t, x) \in \mathcal{D}$, then (II.6.21) holds on $\mathcal{D} \setminus \partial \mathcal{C}$. We now want to show that (II.6.21) holds also for every $(t, x) \in \mathcal{C}$.

Fix $(t, x) \in \mathcal{C}$ and take $\varepsilon > 0$. Recall problem formulation (II.2.14) with the explicit expression for Z (see (II.2.3)). Since τ^* is admissible but sub-optimal for the problem with value $v(t, x + \varepsilon)$, we have

$$\begin{aligned} v(t, x) - v(t, x + \varepsilon) &\leq \mathbb{E} [\exp(Z_{t+\tau^*}^{t,x}) - \exp(Z_{t+\tau^*}^{t,x+\varepsilon})] \\ &= \mathbb{E} \left[\left(1 - \exp \left(-\frac{1-t-\tau^*}{1-t} \varepsilon \right) \right) \exp(Z_{t+\tau^*}^{t,x}) \right]. \end{aligned} \quad (\text{II.6.23})$$

Hence, by dominated convergence theorem and recalling that $v \in C^{1,2}(\mathcal{C})$ (by Proposition II.3.3), we obtain

$$\partial_x v(t, x) = \lim_{\varepsilon \rightarrow 0} \frac{v(t, x + \varepsilon) - v(t, x)}{\varepsilon} \geq \mathbb{E} \left[\frac{1-t-\tau^*}{1-t} \exp(Z_{t+\tau^*}^{t,x}) \right]. \quad (\text{II.6.24})$$

By the same arguments, we also have that

$$v(t, x) - v(t, x - \varepsilon) \leq \mathbb{E} \left[\left(1 - \exp \left(-\frac{1-t-\tau^*}{1-t} \varepsilon \right) \right) \exp(Z_{t+\tau^*}^{t,x}) \right], \quad (\text{II.6.25})$$

which implies

$$\partial_x v(t, x) = \lim_{\varepsilon \rightarrow 0} \frac{v(t, x) - v(t, x - \varepsilon)}{\varepsilon} \leq \mathbb{E} \left[\frac{1 - t - \tau^*}{1 - t} \exp(Z_{t+\tau^*}^{t, x}) \right], \quad (\text{II.6.26})$$

Combining (II.6.24) and (II.6.26), we obtain that (II.6.21) holds also on $\mathcal{C}$.

Now, the inequality in (II.6.22) follows easily by comparison of (II.6.21) and (II.2.14). $\square$

Theorem II.6.4. *We have $v \in C^1([0, 1) \times \mathbb{R})$.*

Proof. We know, from Proposition II.3.3, that $\partial_x v$ and $\partial_t v$ exist and are continuous in $\mathcal{C}$. Moreover, $v(t, x) = e^x$ on $\mathcal{D}$ implies $\partial_x v(t, x) = e^x$ and $\partial_t v(t, x) = 0$ for $(t, x) \in \mathcal{D} \setminus \partial\mathcal{C}$. Then, it remains to prove that $\partial_x v$ and $\partial_t v$ are continuous across the boundary $\partial\mathcal{C}$. We do this in two steps.

Step 1. (Continuity of $\partial_x v$). Fix $(t_0, x_0) \in \partial\mathcal{C}$ with $t_0 < 1$ and recall (II.6.21). Then, for any sequence $(t_n, x_n)_{n \geq 1} \subseteq \mathcal{C}$ converging to (t_0, x_0) as $n \rightarrow \infty$, we can use the dominated convergence theorem, continuity of paths and Proposition II.6.2 to obtain

$$\lim_{n \rightarrow \infty} \partial_x v(t_n, x_n) = \mathbb{E} \left[\lim_{n \rightarrow \infty} \frac{1 - t_n - \tau_{t_n, x_n}^*}{1 - t_n} \exp(Z_{t_n + \tau_{t_n, x_n}^*}^{t_n, x_n}) \right] = e^{x_0}. \quad (\text{II.6.27})$$

Step 2. (Continuity of $\partial_t v$). Let $(t, x) \in \mathcal{C}$ and $0 < \varepsilon < 1 - t$. Then, repeating arguments as those used in (II.3.7) and recalling that $t \mapsto v(t, x)$ is non-increasing on $[0, 1)$ (see proof of Proposition II.5.1), we obtain

$$\begin{aligned} 0 &\geq v(t + \varepsilon, x) - v(t, x) \\ &\geq \mathbb{E} \left[e^{(1-\theta^*)x} \left(e^{\sqrt{(1-\theta^*)(1-t-\varepsilon)}W_{\theta^*}} - e^{\sqrt{(1-\theta^*)(1-t)}W_{\theta^*}} \right) \right] \\ &\geq -(\sqrt{1-t} - \sqrt{1-t-\varepsilon}) e^{|x|} \mathbb{E} \left[|W_{\theta^*}| e^{|W_{\theta^*}|} \right], \end{aligned} \quad (\text{II.6.28})$$

where $\theta^* := \theta_{t,x}^*$ is the optimal stopping time for $v(t, x)$ (see (II.2.16)).

Dividing all terms above by ε and letting $\varepsilon \rightarrow 0$, we obtain

$$0 \geq \partial_t v(t, x) \geq -\frac{1}{2\sqrt{1-t}} e^{|x|} \mathbb{E} \left[|W_{\theta^*}| e^{|W_{\theta^*}|} \right]. \quad (\text{II.6.29})$$

The inequalities in (II.6.29) hold if we replace (t, x) by (t_n, x_n) and θ^* by $\theta_n^* := \theta_{t_n, x_n}^*$, where the sequence $(t_n, x_n)_{n \geq 1} \subseteq \mathcal{C}$ converges to $(t_0, x_0) \in \partial\mathcal{C}$ as $n \rightarrow \infty$.

Now we aim at letting $n \rightarrow \infty$. Notice that Proposition II.6.2 and the definition of θ in (II.2.16) imply

$$\lim_{n \rightarrow \infty} \theta_{t_n, x_n}^* = \lim_{n \rightarrow \infty} \frac{\tau_{t_n, x_n}^*}{1 - t_n} = 0, \quad \mathbb{P}\text{-a.s.} \quad (\text{II.6.30})$$

Thus, using the dominated convergence theorem, we obtain

$$0 \geq \lim_{n \rightarrow \infty} \partial_t v(t_n, x_n) \geq -\frac{1}{2\sqrt{1-t_0}} e^{|x_0|} \mathbb{E} \left[\lim_{n \rightarrow \infty} |W_{\theta_n^*}| e^{|W_{\theta_n^*}|} \right] = 0. \quad (\text{II.6.31})$$

$\square$

Theorem II.6.4 has a simple corollary which shows the regularity of $\partial_{xx} v$. In particular, $\partial_{xx} v$ is

continuous but for a jump across the optimal boundary.

Corollary II.6.5. *The second derivative $\partial_{xx}v$ is continuous on $([0, 1] \times \mathbb{R}) \setminus \partial\mathcal{C}$. Moreover, for every $(t_0, x_0) \in \partial\mathcal{C}$ with $t_0 < 1$ and any sequence $(t_n, x_n)_{n \geq 1} \subseteq \mathcal{C}$ converging to (t_0, x_0) as $n \rightarrow \infty$, we have*

$$\lim_{n \rightarrow \infty} \partial_{xx}v(t_n, x_n) = \frac{2x_0}{1 - t_0} e^{x_0} > e^{x_0}. \quad (\text{II.6.32})$$

Proof. Since $v(t, x) = e^x$ in $\mathcal{D}$, then $\partial_{xx}v(t, x) = e^x$ in $\mathcal{D} \setminus \partial\mathcal{C}$ which is continuous. Moreover, by Proposition II.3.3, $\partial_{xx}v \in C(\mathcal{C})$ and so $\partial_{xx}v$ is continuous on $([0, 1] \times \mathbb{R}) \setminus \partial\mathcal{C}$.

To show (II.6.32), it is sufficient to take limits in (II.3.9), that is

$$\lim_{n \rightarrow \infty} \partial_{xx}v(t_n, x_n) = \lim_{n \rightarrow \infty} 2 \left(-\partial_t v(t_n, x_n) + \frac{x_n}{1 - t_n} \partial_x v(t_n, x_n) \right) = \frac{2x_0}{1 - t_0} e^{x_0}, \quad (\text{II.6.33})$$

where we used Theorem II.6.4 in the last equality. The inequality in (II.6.32) follows from the fact that $\mathcal{Q} \subseteq \mathcal{C}$ (see (II.5.1)). $\square$

II.7 Integral equations for the optimal boundary

In this section we show how the regularity of the value function proved in the previous section allows us to derive a family of integral equations that characterise the optimal boundary. This follows well-known steps (see, e.g., [79]) which we replicate below.

Theorem II.7.1. *For all $(t, x) \in [0, 1] \times \mathbb{R}$, the value function has the following representation*

$$v(t, x) = 1 + \mathbb{E} \left[\int_0^{1-t} e^{X_{t+s}^{t,x}} \left(\frac{X_{t+s}^{t,x}}{1-t-s} - \frac{1}{2} \right) \mathbf{1}_{\{X_{t+s}^{t,x} > b(t+s)\}} ds \right]. \quad (\text{II.7.1})$$

Moreover, the optimal boundary $t \mapsto b(t)$ is the unique continuous solution of the following family of nonlinear integral equations: for every $t \in [0, 1]$,

$$e^{b(t)} = 1 + \mathbb{E} \left[\int_0^{1-t} e^{X_{t+s}^{t,b(t)}} \left(\frac{X_{t+s}^{t,b(t)}}{1-t-s} - \frac{1}{2} \right) \mathbf{1}_{\{X_{t+s}^{t,b(t)} > b(t+s)\}} ds \right], \quad (\text{II.7.2})$$

with $b(1) = 0$ and $b(t) \geq (1-t)/2$.

Proof. In view of Theorem II.6.4 and Corollary II.6.5, we can find a mollifying sequence $(v_n)_{n \geq 0} \subseteq C^\infty([0, 1] \times \mathbb{R})$ for v such that (see, e.g., [47, Section 7.2])

$$(v_n, \partial_x v_n, \partial_t v_n) \rightarrow (v, \partial_x v, \partial_t v) \quad (\text{II.7.3})$$

as $n \rightarrow \infty$, uniformly on compact sets, and

$$\lim_{n \rightarrow \infty} \partial_{xx}v_n(t, x) = \partial_{xx}v(t, x), \quad \forall (t, x) \notin \partial\mathcal{C}. \quad (\text{II.7.4})$$

Let $\varepsilon > 0$ and let $(K_m)_{m \geq 0}$ be a sequence of compact sets increasing to $[0, 1 - \varepsilon] \times \mathbb{R}$. For $t < 1 - \varepsilon$,

define

$$\tau_m := \inf\{s \geq 0 : (t+s, X_{t+s}^{t,x}) \notin K_m\} \wedge (1-t-\varepsilon). \quad (\text{II.7.5})$$

By an application of Itô's formula to v_n and noticing that $\mathbb{P}(X_{t+s}^{t,x} = b(t+s)) = 0$ for $s \in [0, 1-t)$, we obtain

$$\begin{aligned} v_n(t, x) &= \mathbb{E} \left[v_n(t + \tau_m, X_{t+\tau_m}^{t,x}) \right. \\ &\quad \left. - \int_0^{\tau_m} \left(\partial_t v_n(t+s, X_{t+s}^{t,x}) + \mathcal{A}v_n(t+s, X_{t+s}^{t,x}) \right) \mathbf{1}_{\{X_{t+s}^{t,x} \neq b(t+s)\}} ds \right]. \end{aligned} \quad (\text{II.7.6})$$

Now, since $(t+s, X_{t+s}^{t,x})_{s \leq \tau_m}$ lives in a compact set, letting $n \rightarrow \infty$, applying the dominated convergence theorem and using (II.7.3)-(II.7.4), we obtain

$$\begin{aligned} v(t, x) &= \mathbb{E} \left[v(t + \tau_m, X_{t+\tau_m}^{t,x}) \right. \\ &\quad \left. - \int_0^{\tau_m} \left(\partial_t v(t+s, X_{t+s}^{t,x}) + \mathcal{A}v(t+s, X_{t+s}^{t,x}) \right) \mathbf{1}_{\{X_{t+s}^{t,x} \neq b(t+s)\}} ds \right] \\ &= \mathbb{E} \left[v(t + \tau_m, X_{t+\tau_m}^{t,x}) + \int_0^{\tau_m} e^{X_{t+s}^{t,x}} \left(\frac{X_{t+s}^{t,x}}{1-t-s} - \frac{1}{2} \right) \mathbf{1}_{\{X_{t+s}^{t,x} > b(t+s)\}} ds \right], \end{aligned} \quad (\text{II.7.7})$$

where in the second equality we have used (II.3.9) and the fact that $v(t, x) = e^x$ in $\mathcal{D}$.

Notice that $\tau_m \rightarrow 1-t-\varepsilon$ as $m \rightarrow \infty$ and the integrand on the right-hand side of (II.7.7) is non-negative. Recalling (II.3.1) and letting $m \rightarrow \infty$, we can apply the dominated convergence theorem and the monotone convergence theorem (for the integral term) in order to obtain

$$v(t, x) = \mathbb{E} \left[v(1-\varepsilon, X_{1-\varepsilon}^{t,x}) + \int_0^{1-t-\varepsilon} e^{X_{t+s}^{t,x}} \left(\frac{X_{t+s}^{t,x}}{1-t-s} - \frac{1}{2} \right) \mathbf{1}_{\{X_{t+s}^{t,x} > b(t+s)\}} ds \right]. \quad (\text{II.7.8})$$

By the same arguments, letting $\varepsilon \rightarrow 0$ we obtain (II.7.1), i.e.,

$$\begin{aligned} v(t, x) &= \mathbb{E} \left[v(1-, X_{1-}^{t,x}) + \int_0^{1-t} e^{X_{t+s}^{t,x}} \left(\frac{X_{t+s}^{t,x}}{1-t-s} - \frac{1}{2} \right) \mathbf{1}_{\{X_{t+s}^{t,x} > b(t+s)\}} ds \right] \\ &= 1 + \mathbb{E} \left[\int_0^{1-t} e^{X_{t+s}^{t,x}} \left(\frac{X_{t+s}^{t,x}}{1-t-s} - \frac{1}{2} \right) \mathbf{1}_{\{X_{t+s}^{t,x} > b(t+s)\}} ds \right], \end{aligned} \quad (\text{II.7.9})$$

where in the second line we have used that, for $t_n < 1$,

$$1 \leq \liminf_{(t_n, x_n) \rightarrow (1, 0)} v(t_n, x_n) \leq \limsup_{(t_n, x_n) \rightarrow (1, 0)} v(t_n, x_n) \leq \limsup_{(t_n, x_n) \rightarrow (1, 0)} e^{|x_n|} \mathbb{E}[e^{|W_{\tau_n^*}|}] = 1, \quad (\text{II.7.10})$$

which follows from problem formulation (II.2.14) with $\tau_n^* := \tau_{t_n, x_n}^*$ and from Proposition II.6.2.

Now the integral equation (II.7.2) is obtained by setting $(t, x) = (t, b(t))$ in (II.7.1). Uniqueness of the solution to such equation follows a standard proof in four steps that was originally developed in [78]. We repeat these steps here.

Suppose there exists another continuous function $c : [0, 1] \rightarrow \mathbb{R}_+$ with $c(1) = 0$ and such that, for

every $t \in [0, 1]$, it holds $c(t) \geq (1 - t)/2$ and

$$e^{c(t)} = 1 + \mathbb{E} \left[\int_0^{1-t} e^{X_{t+s}^{t, c(t)}} \left(\frac{X_{t+s}^{t, c(t)}}{1-t-s} - \frac{1}{2} \right) \mathbf{1}_{\{X_{t+s}^{t, c(t)} > c(t+s)\}} ds \right]. \quad (\text{II.7.11})$$

Then, define the function

$$v^c(t, x) := 1 + \mathbb{E} \left[\int_0^{1-t} e^{X_{t+s}^{t, x}} \left(\frac{X_{t+s}^{t, x}}{1-t-s} - \frac{1}{2} \right) \mathbf{1}_{\{X_{t+s}^{t, x} > c(t+s)\}} ds \right], \quad (\text{II.7.12})$$

i.e., the analogue of (II.7.1) but replacing $b(\cdot)$ therein with $c(\cdot)$. Since $t \mapsto c(t)$ is assumed continuous and the Brownian bridge admits a continuous transition density, it is not hard to show that $(t, x) \mapsto v^c(t, x)$ is continuous on $[0, 1] \times \mathbb{R}$ (see next section for more accurate computations). Moreover, it is clear that $v^c(1, x) = 1$ for every $x \in \mathbb{R}$ and, by (II.7.11), $v^c(t, c(t)) = e^{c(t)}$ for every $t \in [0, 1]$. By the Markov property, it is not difficult to show that

$$s \mapsto v(t+s, X_{t+s}^{t, x}) + \int_0^s e^{X_{t+u}^{t, x}} \left(\frac{X_{t+u}^{t, x}}{1-t-u} - \frac{1}{2} \right) \mathbf{1}_{\{X_{t+u}^{t, x} > b(t+u)\}} du, \quad (\text{II.7.13})$$

$$s \mapsto v^c(t+s, X_{t+s}^{t, x}) + \int_0^s e^{X_{t+u}^{t, x}} \left(\frac{X_{t+u}^{t, x}}{1-t-u} - \frac{1}{2} \right) \mathbf{1}_{\{X_{t+u}^{t, x} > c(t+u)\}} du, \quad (\text{II.7.14})$$

are $\{\mathcal{F}_{t+s}\}$ -martingales for every $(t, x) \in [0, 1] \times \mathbb{R}$ and, moreover, that they are continuous for $s \in [0, 1-t]$. Then, for fixed $(t, x) \in [0, 1] \times \mathbb{R}$, let us denote

$$\sigma_c := \inf\{s \in [0, 1-t] : X_{t+s}^{t, x} \leq c(t+s)\}, \quad (\text{II.7.15})$$

$$\sigma_b := \inf\{s \in [0, 1-t] : X_{t+s}^{t, x} \leq b(t+s)\}, \quad (\text{II.7.16})$$

$$\tau_c := \inf\{s \in [0, 1-t] : X_{t+s}^{t, x} \geq c(t+s)\}, \quad (\text{II.7.17})$$

$$\tau_b := \inf\{s \in [0, 1-t] : X_{t+s}^{t, x} \geq b(t+s)\}. \quad (\text{II.7.18})$$

We can now show the four steps of the proof.

Step 1. Here, we prove that $v^c(t, x) = e^x$ for every $x \geq c(t)$ and $t \in [0, 1]$. This was already observed for $x = c(t)$, thus let $x > c(t)$. By the martingale property (II.7.14) and the optional sampling theorem, we have

$$\begin{aligned} v^c(t, x) &= \mathbb{E} [v^c(t + \sigma_c, X_{t+\sigma_c}^{t, x})] + \mathbb{E} \left[\int_0^{\sigma_c} e^{X_{t+u}^{t, x}} \left(\frac{X_{t+u}^{t, x}}{1-t-u} - \frac{1}{2} \right) \mathbf{1}_{\{X_{t+u}^{t, x} > c(t+u)\}} du \right] \\ &= \mathbb{E} [e^{X_{t+\sigma_c}^{t, x}}] + \mathbb{E} \left[\int_0^{\sigma_c} e^{X_{t+u}^{t, x}} \left(\frac{X_{t+u}^{t, x}}{1-t-u} - \frac{1}{2} \right) du \right] \\ &= e^x - \mathbb{E} \left[\int_0^{\sigma_c} e^{X_{t+u}^{t, x}} \left(\frac{X_{t+u}^{t, x}}{1-t-u} - \frac{1}{2} \right) du \right] \\ &\quad + \mathbb{E} \left[\int_0^{\sigma_c} e^{X_{t+u}^{t, x}} \left(\frac{X_{t+u}^{t, x}}{1-t-u} - \frac{1}{2} \right) du \right] \\ &= e^x, \end{aligned} \quad (\text{II.7.19})$$

where in the second equality we have also used that $v^c(t + \sigma_c, X_{t+\sigma_c}^{t, x}) = e^{X_{t+\sigma_c}^{t, x}}$ as $X_{t+\sigma_c}^{t, x} = c(t + \sigma_c)$ by continuity of $s \mapsto X_s$ and $s \mapsto c(s)$; in the third equality we have used Dynkin's formula.

Step 2. Here, we show that $v(t, x) \geq v^c(t, x)$ for every $(t, x) \in [0, 1] \times \mathbb{R}$. For $x \geq c(t)$, by Step 1, we have $v^c(t, x) = e^x \leq v(t, x)$. For $x < c(t)$, by the martingale property (II.7.14) and the optional sampling theorem, we obtain

$$\begin{aligned} v^c(t, x) &= \mathbb{E} [v^c(t + \tau_c, X_{t+\tau_c}^{t,x})] + \mathbb{E} \left[\int_0^{\tau_c} e^{X_{t+u}^{t,x}} \left(\frac{X_{t+u}^{t,x}}{1-t-u} - \frac{1}{2} \right) \mathbf{1}_{\{X_{t+u}^{t,x} > c(t+u)\}} du \right] \\ &= \mathbb{E} [e^{X_{t+\tau_c}^{t,x}}] \\ &\leq v(t, x), \end{aligned} \quad (\text{II.7.20})$$

where the second equality follows from the definition (II.7.17) and the fact that $v^c(t + \tau_c, X_{t+\tau_c}) = e^{X_{t+\tau_c}}$.

Step 3. Here, we prove that $b(t) \geq c(t)$ for every $t \in [0, 1]$. Let $t \in [0, 1]$ and $x > b(t) \wedge c(t)$. By the martingale property (II.7.14) and the optional sampling theorem, we have

$$\begin{aligned} \mathbb{E} [v^c(t + \sigma_b, X_{t+\sigma_b}^{t,x})] &= v^c(t, x) - \mathbb{E} \left[\int_0^{\sigma_b} e^{X_{t+u}^{t,x}} \left(\frac{X_{t+u}^{t,x}}{1-t-u} - \frac{1}{2} \right) \mathbf{1}_{\{X_{t+u}^{t,x} > c(t+u)\}} du \right] \\ &= e^x - \mathbb{E} \left[\int_0^{\sigma_b} e^{X_{t+u}^{t,x}} \left(\frac{X_{t+u}^{t,x}}{1-t-u} - \frac{1}{2} \right) \mathbf{1}_{\{X_{t+u}^{t,x} > c(t+u)\}} du \right], \end{aligned} \quad (\text{II.7.21})$$

where in the second equality we have used Step 1. Similarly, by the martingale property (II.7.13) and the optional sampling theorem, we obtain

$$\begin{aligned} \mathbb{E} [v(t + \sigma_b, X_{t+\sigma_b}^{t,x})] &= v(t, x) - \mathbb{E} \left[\int_0^{\sigma_b} e^{X_{t+u}^{t,x}} \left(\frac{X_{t+u}^{t,x}}{1-t-u} - \frac{1}{2} \right) \mathbf{1}_{\{X_{t+u}^{t,x} > b(t+u)\}} du \right] \\ &= e^x - \mathbb{E} \left[\int_0^{\sigma_b} e^{X_{t+u}^{t,x}} \left(\frac{X_{t+u}^{t,x}}{1-t-u} - \frac{1}{2} \right) du \right]. \end{aligned} \quad (\text{II.7.22})$$

Hence, by comparing (II.7.21) and (II.7.22) and using Step 2, we obtain

$$\mathbb{E} \left[\int_0^{\sigma_b} e^{X_{t+u}^{t,x}} \left(\frac{X_{t+u}^{t,x}}{1-t-u} - \frac{1}{2} \right) \mathbf{1}_{\{X_{t+u}^{t,x} \leq c(t+u)\}} du \right] \leq 0. \quad (\text{II.7.23})$$

Since $X_{t+u} > b(t+u) \geq 1/2(1-t-u)$ for every $u \in [0, \sigma_b)$, by continuity of $t \mapsto b(t)$ and $t \mapsto c(t)$, we then obtain that $b(t) \geq c(t)$ (otherwise we would reach a contradiction).

Step 4. Here, we show that $b(t) \leq c(t)$ for every $t \in [0, 1]$ and thus, by Step 3, we conclude that $b(t) = c(t)$ for every $t \in [0, 1]$. By contradiction, let $t \in [0, 1]$ such that $b(t) > c(t)$ and let $x \in (c(t), b(t))$. Then, by the martingale property (II.7.14) and the optional sampling theorem, we have

$$\begin{aligned} v^c(t, x) &= \mathbb{E} [v^c(t + \tau_b, X_{t+\tau_b}^{t,x})] + \mathbb{E} \left[\int_0^{\tau_b} e^{X_{t+u}^{t,x}} \left(\frac{X_{t+u}^{t,x}}{1-t-u} - \frac{1}{2} \right) \mathbf{1}_{\{X_{t+u}^{t,x} > c(t+u)\}} du \right] \\ &= \mathbb{E} [e^{X_{t+\tau_b}^{t,x}}] + \mathbb{E} \left[\int_0^{\tau_b} e^{X_{t+u}^{t,x}} \left(\frac{X_{t+u}^{t,x}}{1-t-u} - \frac{1}{2} \right) \mathbf{1}_{\{X_{t+u}^{t,x} > c(t+u)\}} du \right], \end{aligned} \quad (\text{II.7.24})$$

where the second equality follows from Step 1 and Step 3. Similarly, by the martingale property

(II.7.13) and the optional sampling theorem, we have

$$\begin{aligned} v(t, x) &= \mathbb{E} [v(t + \tau_b, X_{t+\tau_b}^{t,x})] + \mathbb{E} \left[\int_0^{\tau_b} e^{X_{t+u}^{t,x}} \left(\frac{X_{t+u}^{t,x}}{1-t-u} - \frac{1}{2} \right) \mathbf{1}_{\{X_{t+u}^{t,x} > b(t+u)\}} du \right] \\ &= \mathbb{E} [e^{X_{t+\tau_b}^{t,x}}], \end{aligned} \quad (\text{II.7.25})$$

where the second equality follows by definition (II.7.18). By combining (II.7.24) and (II.7.25) and using Step 2, we then obtain

$$\mathbb{E} \left[\int_0^{\tau_b} e^{X_{t+u}^{t,x}} \left(\frac{X_{t+u}^{t,x}}{1-t-u} - \frac{1}{2} \right) \mathbf{1}_{\{X_{t+u}^{t,x} > c(t+u)\}} du \right] \leq 0, \quad (\text{II.7.26})$$

which, by continuity of $t \mapsto b(t)$ and $t \mapsto c(t)$, leads to a contradiction.

Finally, notice that it is shown in [31] that continuity of the boundaries can be relaxed to right/left-continuity. $\square$

II.8 Numerical results

In this section we conduct a numerical analysis of the non-linear Volterra integral equation (II.7.2) which characterises the optimal boundary. First, we describe a Picard scheme that we apply to find a numerical solution to equation (II.7.2). We learned this numerical algorithm for integral equations from [35].

Notice that equation (II.7.2) can be rewritten as

$$e^{b(t)} = 1 + \int_0^{1-t} \left(\int_{b(t+s)}^{\infty} e^y \left(\frac{y}{1-t-s} - \frac{1}{2} \right) p(t, b(t), t+s, y) dy \right) ds, \quad (\text{II.8.1})$$

where $p(t, x, t+s, y) := \partial_y \mathbb{P}(X_{t+s}^{t,x} \leq y)$ is the transition density of the Brownian bridge.

Let $\Pi := \{0 := t_0 < t_1 < \dots < t_n := 1\}$ be an equispaced partition of $[0, 1]$ with mesh $h = 1/n$. The algorithm is initialised by setting $b^{(0)}(t_j) := 0$ for all $j = 0, 1, \dots, n$. Now, let $b^{(k)}(t_j)$ denote, for $j = 0, 1, \dots, n$, the values of the boundary obtained after the k -th iteration. Then, the values for the $(k+1)$ -th iteration are computed, for all $j = 0, \dots, n$, as

$$e^{b^{(k+1)}(t_j)} = 1 + \int_0^{1-t_j} \left(\int_{b^{(k)}(t_j+s)}^{\infty} e^y \left(\frac{y}{1-t_j-s} - \frac{1}{2} \right) p(t_j, b^{(k)}(t_j), t_j+s, y) dy \right) ds. \quad (\text{II.8.2})$$

In particular, the inner integral with respect to dy can be computed explicitly. Indeed, noticing that

$$p(t, x, t+s, y) = \frac{1}{\sqrt{2\pi\alpha(t, s)}} \exp \left(-\frac{(y - \beta(x, t, s))^2}{2\alpha(t, s)} \right), \quad (\text{II.8.3})$$

with $\alpha(t, s) := s(1-t-s)/(1-t)$ and $\beta(x, t, s) := x(1-t-s)/(1-t)$, we can now substitute this expression inside the integral. Then, tedious but straightforward algebra allows to reduce the exponent of $e^y p(t_j, b^{(k)}(t_j), t_j+s, y)$ to an exact square plus a term independent of y . Thus,

properties of the Gaussian distribution give

$$\begin{aligned}
 & I(t_j, b^{(k)}(t_j), t_j + s, b^{(k)}(t_j + s)) \\
 &:= \int_{b^{(k)}(t_j + s)}^{\infty} e^y \left(\frac{y}{1 - t_j - s} - \frac{1}{2} \right) p(t_j, b^{(k)}(t_j), t_j + s, y) dy \\
 &= e^{\gamma^{(k)}(t_j, s)} \left[\frac{\zeta(t_j, s)}{\sqrt{2\pi}} e^{-\frac{\xi^{(k)}(t_j, s)^2}{2}} + \left(\eta^{(k)}(t_j, s) - \frac{1}{2} \right) \left(1 - \Phi(\xi^{(k)}(t_j, s)) \right) \right], \tag{II.8.4}
 \end{aligned}$$

where Φ is the cumulative density function of a standard normal distribution, and

$$\gamma^{(k)}(t_j, s) := \frac{(2b^{(k)}(t_j) + s)(1 - s - t_j)}{2(1 - t_j)}, \quad \eta^{(k)}(t_j, s) := \frac{b^{(k)}(t_j) + s}{1 - t_j}, \tag{II.8.5}$$

$$\zeta(t_j, s) := \frac{s}{(1 - s - t_j)(1 - t_j)}, \quad \xi^{(k)}(t_j, s) := \left(\frac{b^{(k)}(t_j + s)}{1 - s - t_j} - \eta^{(k)}(t_j, s) \right) \frac{1}{\zeta(t_j, s)}. \tag{II.8.6}$$

The integral with respect to the time variable (i.e., the one in ds) is then computed by a standard quadrature method. Hence, (II.8.2) reduces to

$$e^{b^{(k+1)}(t_j)} = 1 + h \sum_{m=0}^{n-1-j} I(t_j, b^{(k)}(t_j), t_j + mh + \frac{h}{2}, b^{(k)}(t_j + mh + \frac{h}{2})), \tag{II.8.7}$$

where each $b^{(k)}(t_j + mh + \frac{h}{2})$ is computed by interpolation and we use the convention $\sum_{m=0}^{-1} = 0$ for $j = n$. Finally,

$$b^{(k+1)}(t_j) = \log \left(1 + h \sum_{m=0}^{n-1-j} I(t_j, b^{(k)}(t_j), t_j + mh + \frac{h}{2}, b^{(k)}(t_j + mh + \frac{h}{2})) \right). \tag{II.8.8}$$

The algorithm stops when the numerical error $e_k := \max_{j=0, \dots, n} |b^{(k)}(t_j) - b^{(k-1)}(t_j)|$ fulfils the tolerance condition $e_k < \varepsilon$, for some $\varepsilon > 0$. A numerical approximation of the optimal boundary is presented in Figure II.1.

While a rigorous proof of the convergence of the scheme seems difficult and falls outside the scope of this work, in Figure II.2 we show that the numerical error e_k converges to zero as the number of iterations increases. Moreover, the convergence is monotone, which results in a good stability of the scheme.

Finally, in Figure II.3 we plot the value function as a surface in the (t, x) -plane using (II.7.1). It is interesting to observe that, as predicted by the theory, the value function exhibits a jump at $\{1\} \times (-\infty, 0)$.

Remark II.8.1. *As noted in Remark II.2.1, we could consider a Brownian bridge with a generic pinning time $T > t$ and our analysis would not be affected (except obvious adjustments). However, it may be interesting to observe that as $T \rightarrow \infty$ the Brownian bridge converges (in law) to a Brownian motion W . Thus, we also expect that the stopping problem (II.2.7) converges to the problem of stopping the exponential of a Brownian motion over an infinite time horizon. Since $t \mapsto \exp(x + W_t)$ is a submartingale, the optimal stopping rule is to never stop. This heuristics is confirmed by Figure II.4 where we observe numerically that the continuation set expands as T increases and, in the limit as $T \rightarrow +\infty$, the stopping set disappears.*

Remark II.8.2. *Traditionally, integral equations as in (II.7.2) are solved by discretisation of the*

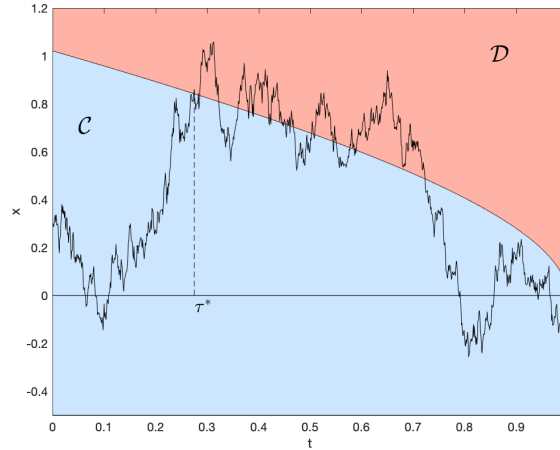


Figure II.1: A sample path of a Brownian bridge X starting at $X_0 = 0.3$ and pinned at $X_1 = 0$. The Brownian bridge hits the optimal boundary at $\tau^* \approx 0.3$. The boundary divides the state space into the continuation region $\mathcal{C}$ (in light blue) and the stopping region $\mathcal{D}$ (in red). The tolerance of the algorithm is set to $\varepsilon = 10^{-6}$ and the equispaced time step is $h = 10^{-3}$.

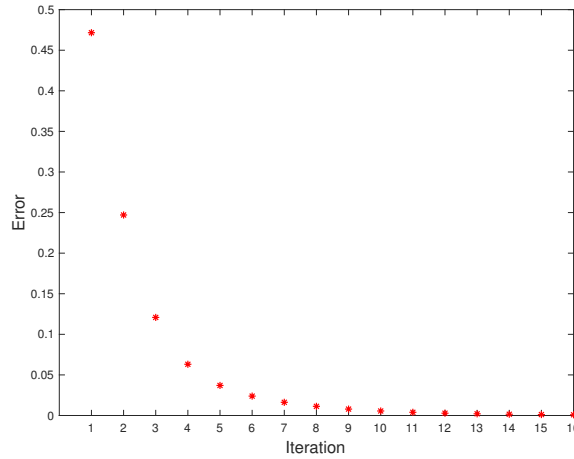


Figure II.2: The trajectory of the error e_k for the 16 iterations of the algorithm when $\varepsilon = 10^{-3}$.

integral with respect to time and a backward procedure, starting from the terminal time (see, e.g., [79, Ch. VII, Sec. 27, pp. 432-433] for details in the case of the Asian option or [79, Ch. VIII, Sec. 30, p. 475] for another example; this method has been developed in the seminal paper [67] and later extended). To the best of our knowledge, a rigorous proof of convergence for this “traditional” numerical scheme is not available.

At each time step, the scheme must find the root of a highly non-linear algebraic equation, making the procedure slower than the Picard scheme that we implement, which requires no root finding (see Figure II.5 for a comparison).

Another possibility is to use finite-difference methods to solve directly the free boundary problem in (II.3.9)–(II.3.10). The finite difference method, however, requires discretisation of both time and

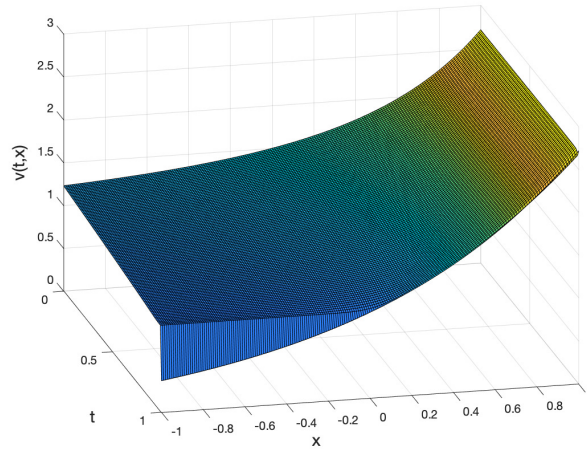


Figure II.3: The value function surface $v(t, x)$ plotted on a grid of points $(t, x) \in [0, 1] \times [-1, 1]$ with discretization step $h = 10^{-2}$.

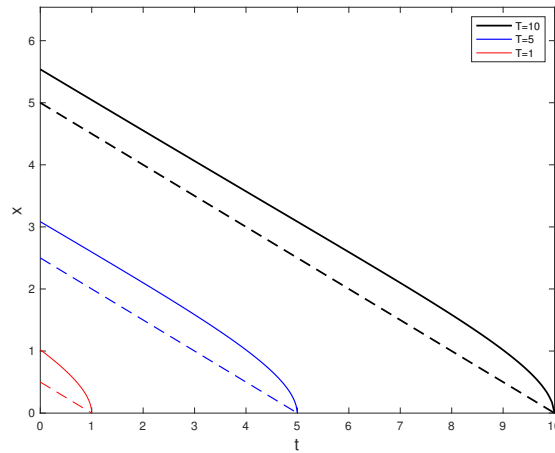


Figure II.4: Boundary functions with starting point $t = 0$ and increasing pinning times $T = 1$, $T = 5$ and $T = 10$ (plotted in red, blue and black, respectively, with increasing line thickness). Notice that every boundary lies above the corresponding line $x = \frac{1}{2}(T - t)$ (represented by the dashed lines with the same colour and same thickness), which generalises the line $\mathcal{R}$ from (II.5.2).

space (whereas we only discretise time) which leads to discretisation errors both in time and space and generally to a slower convergence. Moreover, in our case the coefficient associated to the first order partial derivative $\partial_x v$ is discontinuous at $t = 1$, which causes additional difficulties.

Remark II.8.3. After we solved numerically equation (II.7.2) and we were able to plot the boundary function as in Figure II.1, we began to investigate whether a suitable fit for the boundary could be found. It is known that in the stopping problem with linear payoff $\mathbb{E}[X_\tau]$ and pinning time $T = 1$, the optimal boundary can be found explicitly and it takes the form $b(t) = B\sqrt{1 - t}$ (see [82]). Motivated by this result, we considered candidate boundaries of the form $b_T(t) = A_T(1 - \exp(B_T\sqrt{T - t}))$, where T is the pinning time of the Brownian bridge, and A_T and B_T are parameters to be determined.

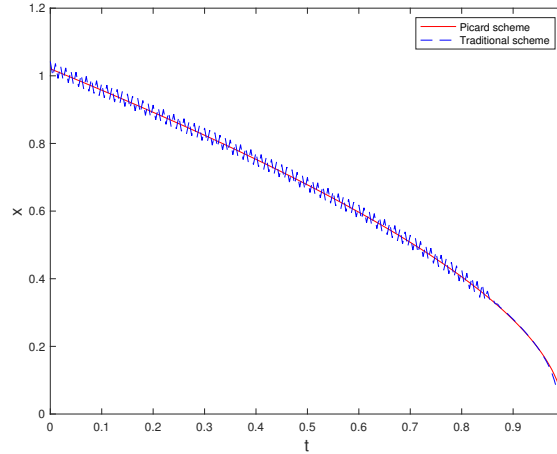


Figure II.5: The solution of the optimal boundary found via Picard scheme (continuous red line) and via traditional method (dashed blue line). The time step is $h = 5 \cdot 10^{-3}$ and the tolerance is $\varepsilon = 10^{-5}$. The Picard scheme stops after 0.1 seconds and 36 iterations, the traditional method stops after 2.9 seconds.

Using the ‘curve fitting toolbox’ in Matlab to fit our ansatz to the boundaries obtained from the integral equations, we obtain an excellent fit for $T = 1, 5, 10$. The quality of the fit slightly deteriorates for larger T ’s (e.g., for $T = 20$). Results are illustrated in Figure II.6. While these tests suggest that our problem might be amenable to an explicit solution, the question is more complex than in the linear payoff case (and its extensions in [41]) and remains open. The key difficulties are that (i) we must determine two parameters rather than one and (ii) we do not have a good guess for the value function that would allow us to transform the free-boundary problem (II.3.9)-(II.3.10) into a solvable ordinary differential equation. Indeed, in the linear case one uses the identity in law

$$\mathcal{L}(X_s^{t,x}, s \in [t, 1]) = \mathcal{L}(\hat{Z}_s^x, s \in [0, \infty)), \quad (\text{II.8.9})$$

with $\hat{Z}_s^x := (x + \sqrt{1-t}W_s)/(1+s)$, and obtains

$$U(t, x) := \sup_{0 \leq \tau \leq 1-t} \mathbb{E}[X_{t+\tau}^{t,x}] = \sup_{\tau \geq 0} \mathbb{E} \left[\frac{x + \sqrt{1-t}W_\tau}{1+\tau} \right] = \sqrt{1-t} U \left(0, \frac{x}{\sqrt{1-t}} \right). \quad (\text{II.8.10})$$

With the exponential payoff, the same identity does not provide any useful insight.

A different approach based on finding parameters for which the guessed boundary $b_T(t)$ solves the integral equation (II.7.2) seems even harder.

II.9 Conclusions

In this chapter we have studied the problem of optimally stopping the exponential of a Brownian bridge, which was posed by Ernst and Shepp in [42] and was motivated by financial applications. We have first shown the existence of an optimal boundary and proved its regularity, in particular its monotonicity. By means of this result, we have been able to show C^1 -regularity of the value function and to characterise the optimal boundary as the unique continuous solution to a family of

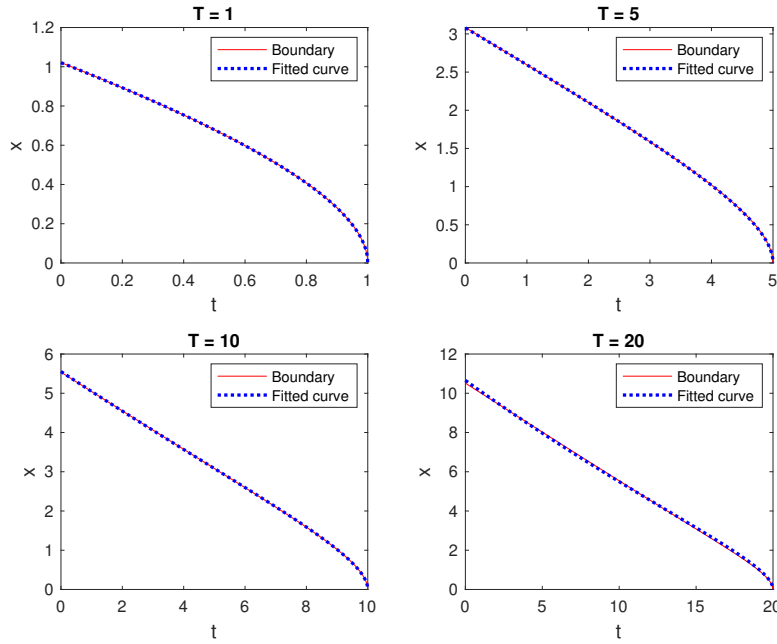


Figure II.6: Boundary functions (continuous lines) and the corresponding fitted curves (dashed lines) of the form $b_T(t) = A_T(1 - \exp(B_T\sqrt{T-t}))$ for the pinning times $T = 1, 5, 10, 20$. The values of the parameters are, respectively, $A_1 = -2.09$, $B_1 = 0.4$; $A_5 = -1.85$, $B_5 = 0.43$; $A_{10} = -1.86$, $B_{10} = 0.44$ and $A_{20} = -2.27$, $B_{20} = 0.39$.

integral equations. We have solved these integral equations numerically to obtain an approximation of the optimal boundary and we have carried some numerical analysis of the algorithm.

Our work is one of the very few fully-solved time-inhomogeneous optimal stopping problems. As expected, the time-inhomogeneous feature of the problem leads to mathematical challenges, in particular when proving the monotonicity of the optimal boundary which is essential for the subsequent analysis of the problem. To obtain this result, we have thus developed novel techniques based on pathwise properties of the Brownian bridge. We believe this method could be extended to more general time-inhomogeneous optimal stopping problems and our aim is now to study this extension. This would provide a new approach in the study of the monotonicity of the boundary for a class of time-inhomogeneous optimal stopping problems.

Chapter III

Singular stochastic control with discretionary stopping

In this chapter I present my doctoral research on singular stochastic control with discretionary stopping. The main result of this work is the detailed derivation of the dynamic programming principle for this class of problems. We also show, under some technical assumptions, how to obtain an optimal stopping rule for these problems.

III.1 Introduction

In stochastic optimal control theory, singular control problems are characterised by two main features. Firstly, the controls are of “singular” nature, i.e., they are, in general, not absolutely continuous with respect to the Lebesgue measure as a function of time. Secondly, the controls act additively on the underlying stochastic dynamics. The optimiser pays a cost which is proportional to the total variation of the control and represents, for instance, the cost of fuel used to activate the control or a transaction cost in mathematical finance.

In a stochastic control problem, if the controller can also choose when to end the optimisation, we then refer to stochastic control with discretionary stopping. The main result of this work is to provide a rigorous proof of the dynamic programming principle for singular control with discretionary stopping problems.

The dynamic programming principle, which was formulated by Bellman in [11], is a classical result in control theory. Although simple and intuitive in its statement, a rigorous proof is very technical. For this reason, it was studied in various articles and textbooks where it was established for different control problems and by different methods. Among several extended works on this topic, we cite, e.g., [68], [74], [17], [44], [85] and [43]. The dynamic programming principle has been obtained also for general stochastic control problems that can include multiple settings. In [21] a weak dynamic programming principle is derived, under some abstract assumptions, for general control problems. The same result is obtained in [19] for problems with state constraints. The purpose of [25] is to provide a rigorous technical proof of a pseudo-Markov property (which leads to the dynamic programming principle) for general control problems with continuous underlying diffusion.

One of the first examples of singular control problems dates back to [8]. Since the 1980's a wide literature has been developed for these problems, starting with the works [12], [60] and [61] where some specific one-dimensional singular control problems were studied. For general multi-dimensional singular control problems, [51]-[52] proved the existence of optimal controls and derived a version of the dynamic programming principle. In [4] the dynamic programming principle was also obtained for multi-dimensional singular control problems under state constraints. In contrast with [51] and [52] where the cost considered is strictly positive, in [24] the dynamic programming principle was established for a one-dimensional singular control problem with null cost. An equivalence between singular control problems and optimal stopping problems was established in [63]-[64]. Since then several authors have obtained a similar result for different versions of singular control problems. Among various work, we cite, e.g., [7], [16], [2], [49] and [30].

Problems of singular stochastic control with discretionary stopping arise in two main applications. The first one is target tracking problems where the optimiser's goal is to arrive as close as possible to a target by using fuel, and decide to stop when the target is sufficiently close. Some specific target problems were studied in [62] and [18]. The second application is in mathematical finance. It was shown in [26] that trading strategies in consumption/investment problems including transaction cost are of singular nature. In this context, the stopping decision variable acts as an early exit from the investment or the exercise time of an American option held in the portfolio. Several versions of singular control with discretionary stopping problems have been studied. In [27] the authors solve a singular control with discretionary stopping problem with quadratic cost functions, in infinite horizon and whose underlying diffusion is a controlled Brownian motion. A version of the same problem with finite-fuel controls is studied in [62]. In [72] the author applies a penalisation method to solve a problem of singular control with discretionary stopping in infinite horizon and whose underlying diffusion is a controlled geometric Brownian motion. Penalisation techniques are also used in [87] to obtain the dynamic programming principle for a more general singular control problem. Partial differential equations methods are applied in [23] to solve the problem in finite horizon, with quadratic costs and underlying controlled Brownian motion.

Problems of combined control and stopping have been studied also in the context of classical (i.e., non-singular) controls. In [65] the authors consider infinite horizon and apply martingale methods to derive the dynamic programming principle and obtain optimal solutions. Similar results are found also in finite horizon in [3], where backward stochastic differential equations methods are applied.

The purpose of this work is to develop a detailed technical study of general singular stochastic control with discretionary stopping problems in finite horizon. Then main result of our study is the proof of the dynamic programming principle for this class of problems. As highlighted in [25], the essential tool to prove this result is provided by the concept of regular conditional probability. To this extent, we follow closely the analysis of [43, Chapter 2] where the dynamic programming principle is obtained, by means of regular conditional probabilities, for classical stochastic control problems. For the generality of our setting in the context of singular control problems, our work mostly compares with [52]. Although we do not consider classical controls as in [52], the novelty of our work is to additionally include

- (i) discretionary stopping,
- (ii) a terminal cost/gain function,
- (iii) an exit time from a fixed domain.

We also differ from [18], as we aim to provide the dynamic programming principle specifically for general problems of singular controls with discretionary stopping, whereas it does not seem straightforward to adapt their results to our setting. Notice also that, however we follow a similar approach, the work in [25] does not apply directly to our problem as their underlying diffusion is continuous.

In addition to the dynamic programming principle, we also provide, under some technical assumptions, an optimal stopping strategy for our problem. The main technical assumption for this second result can be viewed as the uniform (left-)continuity assumption with respect to the class of non-linear expectations considered in [10].

The chapter is organised as follows. In Section III.2 we introduce some notation and terminology that will be used throughout the chapter. In Section III.3 we formulate the control problem. In Section III.4 we show that our problem is independent on the underlying reference probability system, result which will be a key to prove the dynamic programming principle. In Section III.5 we introduce the notion of regular conditional probabilities and some related results which will set the framework for the dynamic programming principle. In Section III.6 we define the value process which will lead to an easier proof of the dynamic programming principle and will also be used to derive the optimal stopping rule. In Section III.7 we prove the dynamic programming principle. In Section III.8 we provide an optimal stopping rule for our problem. In Section III.9 we prove the regularity of the value function which is necessary to obtain the previous results. Finally, in Section III.10 we show some miscellaneous results.

III.2 Notation and terminology

In this section we list some notation and terminology that will be used throughout this chapter.

- $T > 0$ is the time horizon and $d, d' \in \mathbb{N}$ are fixed dimensions.
- For $t \in [0, T]$, a reference probability system starting at time t is a 5-tuple

$$\nu := (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_s^t\}_{s \in [t, T]}, W), \quad (\text{III.2.1})$$

where $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space; $W := (W_s)_{s \in [t, T]}$ is a d' -dimensional Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$ starting at time t , i.e., $\mathbb{P}(W_t = 0) = 1$; $\mathcal{F}_s^{t,0} := \sigma(W_u : u \in [t, s])$ and $\mathcal{F}_s^t$ is the augmentation of $\mathcal{F}_s^{t,0}$ with the $\mathbb{P}$ -null sets. Without loss of generality, we assume that W has everywhere continuous paths. We denote by $\mathcal{V}_t$ the collection of reference probability systems starting at time t .

- We say that a reference probability system $\nu \in \mathcal{V}_t$ is standard if there exists a σ -algebra $\mathcal{F}^0$ such that $\mathcal{F}_T^{t,0} \subseteq \mathcal{F}^0 \subseteq \mathcal{F}$, where $\mathcal{F}$ is the completion of $\mathcal{F}^0$ with the $\mathbb{P}$ -null sets and $(\Omega, \mathcal{F}^0)$ is a standard measurable space, i.e., it is Borel isomorphic to $(\mathbb{N}, \mathcal{B}(\mathbb{N}))$.
- For $t \in [0, T]$, the canonical reference probability system (starting at time t) is the 5-tuple

$$\nu^* := (\Omega^*, \mathcal{F}^*, \mathbb{P}^*, \{\mathcal{B}_s^t\}_{s \in [t, T]}, W^*), \quad (\text{III.2.2})$$

where $\Omega^* := \{\omega \in C([t, T]; \mathbb{R}^{d'}) : \omega(t) = 0\}$; $\mathbb{P}^*$ is the Wiener measure on $(\Omega^*, \mathcal{B}(\Omega^*))$ that makes the process $W_s^*(\omega) := \omega(s)$ for $(s, \omega) \in [t, T] \times \Omega^*$ a Brownian motion starting at time

t ; $\mathcal{F}^*$ is the completion of $\mathcal{B}(\Omega^*)$ with the $\mathbb{P}^*$ -null sets; $\mathcal{B}_s^{t,0} := \sigma(W_u^* : u \in [t, s])$ and $\mathcal{B}_s^t$ is the augmentation of $\mathcal{B}_s^{t,0}$ with the $\mathbb{P}^*$ -null sets.

- For a fixed $\nu \in \mathcal{V}_t$ and $u \in [t, T]$, we denote by $\mathcal{A}_u^\nu$ the class of admissible singular controls, i.e., the collection of processes $\xi = (\xi_s)_{s \in [t, T]}$ such that

- (i) ξ is $\mathbb{R}^d$ -valued and $\{\mathcal{F}_s^t\}$ -adapted.
- (ii) ξ is left-continuous and of bounded variation $\mathbb{P}$ -a.s.
- (iii) $\xi_s = 0$ for every $s \in [t, u]$, $\mathbb{P}$ -a.s.
- (iv) $\mathbb{E}[|\xi_T^\pm|^2] < \infty$.

Here, $\xi^\pm$ denote the positive and negative parts of the Jordan decomposition $\xi = \xi^+ - \xi^-$, $\mathbb{P}$ -a.s. with

$$\xi^\pm := \frac{\check{\xi} \pm \xi}{2}, \quad \mathbb{P}\text{-a.s.}, \quad (\text{III.2.3})$$

where $\check{\xi}$ denotes the total variation of ξ . Notice that $\xi^\pm$ satisfy the same properties as ξ and are non-decreasing.

- For a fixed $\nu \in \mathcal{V}_t$, $\xi \in \mathcal{A}_t^\nu$ and $u \in [t, T]$, we define

$$\mathcal{A}_{u,\xi}^\nu := \{\zeta \in \mathcal{A}_t^\nu : \zeta_s = \xi_s, \forall s \in [t, u], \mathbb{P}\text{-a.s.}\}. \quad (\text{III.2.4})$$

- For a fixed $\nu \in \mathcal{V}_t$ and $u \in [t, T]$, we denote by $\mathcal{T}_u^\nu$ the class of admissible stopping times (greater than u), i.e., the collection of $\{\mathcal{F}_s^t\}$ -stopping times such that $u \leq \tau \leq T$, $\mathbb{P}$ -a.s.
- For a fixed $\nu \in \mathcal{V}_t$ and $u \in [t, T]$, we denote by $\mathcal{E}_u^\nu$ the collection of processes $\eta = (\eta_s)_{s \in [t, T]}$ such that
 - (i) η is $\{\mathcal{F}_s^t\}$ -adapted.
 - (ii) η is left-continuous and non-decreasing $\mathbb{P}$ -a.s.
 - (iii) $\eta_s \in \{0, 1\}$ for every $s \in [t, T]$ and $\eta_s = 0$ for every $s \in [t, u]$, $\mathbb{P}$ -a.s.
- For $\tau \in \mathcal{T}_t^\nu$, a Borel-measurable function $h : [t, T] \mapsto \mathbb{R}^d$ with $h = (h^1, h^2, \dots, h^d)$ and $\xi = (\xi^1, \xi^2, \dots, \xi^d) \in \mathcal{A}_t^\nu$, we use the notation

$$\int_{[t,\tau)} \langle h(s), d\xi_s \rangle := \sum_{i=1}^d \int_{[t,\tau)} h^i(s) d\xi_s^i, \quad (\text{III.2.5})$$

where every term of the sum on the right-hand side is a Lebesgue-Stieltjes integral.

- For a fixed $\nu \in \mathcal{V}_t$, we denote by $\mathcal{P}_\Omega$ the σ -algebra of $\{\mathcal{F}_s^{t,0}\}$ -predictable sets, i.e., the σ -algebra on $[t, T] \times \Omega$ generated by the sets of the form $(s, u] \times A$ with $t \leq s < u \leq T$ and $A \in \mathcal{F}_{s-}^{t,0}$, and $\{t\} \times A$ with $A \in \mathcal{F}_t^{t,0}$. Notice that, by the left-continuity of the filtration generated by a Brownian motion, we have $\mathcal{F}_{s-}^{t,0} = \mathcal{F}_s^{t,0}$ for every $s \in [t, T]$. A process $Y : [t, T] \times \Omega \rightarrow \mathbb{R}^d$ that is $\mathcal{P}_\Omega / \mathcal{B}(\mathbb{R}^d)$ -measurable is called $\{\mathcal{F}_s^{t,0}\}$ -predictable.

III.3 Problem formulation

Let $t \in [0, T]$, $\nu \in \mathcal{V}_t$ and let $\xi := (\xi_s)_{s \in [t, T]} \in \mathcal{A}_t^\nu$ be an admissible control. The controlled state process for our problem follows the stochastic differential equation (SDE)

$$\begin{cases} dX_s = \mu(s, X_s)ds + \sigma(s, X_s)dW_s + d\xi_s, & s \in [t, T], \\ X_t = x \in \mathbb{R}^d, \end{cases} \quad (\text{III.3.1})$$

where $\mu : [t, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$, $\sigma : [t, T] \times \mathbb{R}^d \rightarrow M(\mathbb{R}^d, \mathbb{R}^{d'})$. We assume that the SDE (III.3.1) admits a unique (up to indistinguishability) $\{\mathcal{F}_s^t\}$ -adapted solution which is $\mathbb{P}$ -a.s. left-continuous with finite right limits. This holds, for instance, under the standard conditions of sub-linear growth and Lipschitz continuity for the coefficients μ and σ , which we discuss in Section III.9. We denote such a solution by $X^{t,x;\xi,\nu}$ to highlight the dependence on the fixed system $\nu \in \mathcal{V}_t$ and the fixed admissible control $\xi \in \mathcal{A}_t^\nu$. When no confusion shall arise, we may denote the solution by $X^{t,x}$, $X^{\xi,\nu}$, X^ξ or simply X depending on the circumstances. Similarly, we denote by $\rho^{t,x;\xi,\nu}$ the first exit time of the process $X^{t,x;\xi,\nu}$ from a domain $\mathcal{O} \in \mathcal{B}(\mathbb{R}^d)$, i.e., $\rho^{t,x;\xi,\nu} := \inf\{s \geq t : X_s^{t,x;\xi,\nu} \notin \mathcal{O}\} \wedge T$. When no confusion shall arise, we may denote $\rho^{t,x;\xi,\nu}$ by $\rho^{\xi,\nu}$, ρ^ξ or simply ρ .

For every $(\tau, \xi) \in \mathcal{T}_t^\nu \times \mathcal{A}_t^\nu$, the gain functional for our problem is defined as

$$\begin{aligned} J_{t,x}^\nu(\tau, \xi) := & \mathbb{E} \left[\int_t^{\tau \wedge \rho^\xi} f(s, X_s^\xi) ds - \int_{[t, \tau \wedge \rho^\xi)} \langle c_+(s), d\xi_s^+ \rangle \right. \\ & \left. - \int_{[t, \tau \wedge \rho^\xi)} \langle c_-(s), d\xi_s^- \rangle + g_1(\rho^\xi, X_{\rho^\xi}^\xi) \mathbf{1}_{\{\rho^\xi \leq \tau\}} + g_2(\tau, X_\tau^\xi) \mathbf{1}_{\{\rho^\xi > \tau\}} \right], \end{aligned} \quad (\text{III.3.2})$$

where $f : [t, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ is the running gain; $g_1 : [t, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ is the terminal gain when ρ^ξ occurs before τ ; $g_2 : [t, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ is the terminal gain when ρ^ξ occurs after τ ; $c_\pm : [t, T] \rightarrow \mathbb{R}^d$ is the running costs that have to be paid when activating the control ξ . In our formulation, we consider the two separate costs $c_\pm$ in order to allow for a (possible) different cost to be paid when activating the control ξ in the two opposite directions $\xi^\pm$. For the sake of generality, $c_\pm$ take value in $\mathbb{R}^d$ so that negative cost may be allowed. We assume that the functions $(s, y) \mapsto f(s, y)$, $s \mapsto c_\pm(s)$, $(s, y) \mapsto g_1(s, y)$ and $(s, y) \mapsto g_2(s, y)$ are Borel-measurable.

The objective is to maximise the gain functional $J_{t,x}^\nu(\tau, \xi)$ over $(\tau, \xi) \in \mathcal{T}_t^\nu \times \mathcal{A}_t^\nu$. To accomplish this goal, we define two formulations of the problem. The problem in the strong formulation is defined for fixed $t \in [0, T]$ and $\nu \in \mathcal{V}_t$ as

$$v_t^\nu(x) := \sup_{(\tau, \xi) \in \mathcal{T}_t^\nu \times \mathcal{A}_t^\nu} J_{t,x}^\nu(\tau, \xi), \quad x \in \mathbb{R}^d. \quad (\text{III.3.3})$$

The problem in the weak formulation is defined as

$$v(t, x) := \sup_{\nu \in \mathcal{V}_t} \sup_{(\tau, \xi) \in \mathcal{T}_t^\nu \times \mathcal{A}_t^\nu} J_{t,x}^\nu(\tau, \xi) = \sup_{\nu \in \mathcal{V}_t} v_t^\nu(x), \quad (t, x) \in [0, T] \times \mathbb{R}^d. \quad (\text{III.3.4})$$

As it will be shown, the two formulations are in fact identical but this dual approach is essential in order to prove the dynamic programming principle. Throughout this work, we also assume that the function $(s, y) \mapsto v(s, y)$ is Borel-measurable and that $|v(t, x)| < \infty$ for every $(t, x) \in [0, T] \times \mathbb{R}^d$.

Remark III.3.1. We adopt the same notation and terminology as in [43, Section 2.1]. The term

“weak” in the weak formulation (III.3.4) refers to the fact that the reference probability system can vary together with the controls. The different notation for $t \in [0, T]$ in the two formulations of the problem is chosen to highlight that in the strong formulation if t changes also the fixed system $\nu \in \mathcal{V}_t$ has to change accordingly, whereas in the weak formulation t can vary freely with x .

III.4 Independence of the reference probability system

In this section we show that the problem is independent of the choice of the reference probability system $\nu \in \mathcal{V}_t$ and thus that the strong formulation (III.3.3) and the weak formulation (III.3.4) coincide, i.e., for every $(t, x) \in [0, T] \times \mathbb{R}^d$ and $\nu \in \mathcal{V}_t$,

$$v_t^\nu(x) = v(t, x). \quad (\text{III.4.1})$$

To obtain this result, we follow closely the analysis as in [43, Chapter 2], where the result is obtained in the case of classical controls and no stopping, and we adapt their arguments to our setting. Notice that, as stated in [43, Remark 2.15], all the conditions imposed on the class of admissible controls that go beyond measurability (for instance, left-continuity, bounded variation and the integrability condition in our case) may lead to further complications and need to be treated in a specific way. Therefore, our study requires different technical lemmas.

In this section, we fix a starting time $t \in [0, T]$. We now show that in our optimisation the stopping time can equivalently be replaced by a further singular control. Fix $\nu \in \mathcal{V}_t$, let $\tau \in \mathcal{T}_t^\nu$ and define $\eta_s^\tau := \mathbf{1}_{\{s > \tau\}}$ for $s \in [t, T]$. Then, $\eta^\tau \in \mathcal{E}_t^\nu$ (recall the definition in Section III.2). On the other hand, let $\eta \in \mathcal{E}_t^\nu$ then $\tau^\eta := \inf\{s \geq t : \eta_s = 1\} \wedge T \in \mathcal{T}_t^\nu$. Moreover, notice that $\eta_{s+} = \mathbf{1}_{\{s \geq \tau\}}$. Thus, for every $(\tau, \xi) \in \mathcal{T}_t^\nu \times \mathcal{A}_t^\nu$, we have that

$$\begin{aligned} J_{t,x}^\nu(\tau, \xi) &= \mathbb{E} \left[\int_t^\rho \mathbf{1}_{\{s \leq \tau\}} f(s, X_s^\xi) ds - \int_{[t,\rho)} \mathbf{1}_{\{s < \tau\}} \langle c_+(s), d\xi_s^+ \rangle \right. \\ &\quad \left. - \int_{[t,\rho)} \mathbf{1}_{\{s < \tau\}} \langle c_-(s), d\xi_s^- \rangle + g_1(\rho, X_\rho^\xi) \mathbf{1}_{\{\rho \leq \tau\}} + g_2(\tau, X_\tau^\xi) \mathbf{1}_{\{\rho > \tau\}} \right] \\ &= \mathbb{E} \left[\int_t^\rho (1 - \eta_s^\tau) f(s, X_s^\xi) ds \right. \\ &\quad \left. - \int_{[t,\rho)} (1 - \eta_{s+}^\tau) \langle c_+(s), d\xi_s^+ \rangle - \int_{[t,\rho)} (1 - \eta_{s+}^\tau) \langle c_-(s), d\xi_s^- \rangle \right. \\ &\quad \left. + g_1(\rho, X_\rho^\xi) (1 - \eta_\rho^\tau) + \int_{[t,\rho)} g_2(s, X_s^\xi) d\eta_s^\tau \right] \\ &=: I_{t,x}^\nu(\eta^\tau, \xi). \end{aligned} \quad (\text{III.4.2})$$

Similarly, for every $(\eta, \xi) \in \mathcal{E}_t^\nu \times \mathcal{A}_t^\nu$, we have that

$$I_{t,x}^\nu(\eta, \xi) = J_{t,x}^\nu(\tau^\eta, \xi). \quad (\text{III.4.3})$$

Thus,

$$v_t^\nu(x) = \sup_{(\eta, \xi) \in \mathcal{E}_t^\nu \times \mathcal{A}_t^\nu} I_{t,x}^\nu(\eta, \xi), \quad x \in \mathbb{R}^d.$$

We now introduce a series of lemmas which will lead us to the desired result (III.4.1).

The next lemma and the subsequent corollary show how to generate, from two admissible controls

ξ and η that are $\{\mathcal{F}_s^t\}$ -adapted, two new admissible controls ξ^0 and η^0 that are $\{\mathcal{F}_s^{t,0}\}$ -predictable and indistinguishable from ξ and η , respectively. This is a key result for the whole analysis of the problem.

Lemma III.4.1. *Let $\nu = (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_s^t\}, W) \in \mathcal{V}_t$ and γ be an $\mathbb{R}^d$ -valued, $\{\mathcal{F}_s^t\}$ -adapted process which is $\mathbb{P}$ -a.s. left-continuous. Then, there exists an $\{\mathcal{F}_s^{t,0}\}$ -predictable process γ^0 which is indistinguishable from γ .*

Proof. Let $\Omega_0 \subseteq \Omega$ be such that $\mathbb{P}(\Omega_0) = 1$ and $s \mapsto \gamma_s(\omega)$ is left-continuous for every $\omega \in \Omega_0$. Let $n \in \mathbb{N}$ and let $\{t_i^n\}_{i=0}^{2^n}$ be the corresponding dyadic partition of $[t, T]$, i.e., $t_i^n = t + i(T-t)/2^n$. By defining, for $n \in \mathbb{N}$, the sequence of processes

$$\gamma_s^n := \gamma_t \mathbf{1}_{\{t\}}(s) + \sum_{i=0}^{2^n-1} \gamma_{t_i^n} \mathbf{1}_{(t_i^n, t_{i+1}^n]}(s), \quad s \in [t, T], \quad (\text{III.4.4})$$

we obtain that $\gamma_s^n(\omega) \rightarrow \gamma_s(\omega)$ for every $(s, \omega) \in [t, T] \times \Omega_0$ as $n \rightarrow \infty$. By applying [59, Lemma 1.25] to every $\gamma_{t_i^n}$ that is $\mathcal{F}_{t_i^n}^t$ -measurable, we obtain $\gamma_{t_i^n}^0$ that is $\mathcal{F}_{t_i^n}^{t,0}$ -measurable and such that $\mathbb{P}(\gamma_{t_i^n} = \gamma_{t_i^n}^0) = 1$. Namely, for every $n \in \mathbb{N}$, $i = 0, 1, \dots, 2^n$ there exists $\Omega^{n,i} \subseteq \Omega$ with $\mathbb{P}(\Omega^{n,i}) = 1$ such that for every $\omega \in \Omega^{n,i}$ we have $\gamma_{t_i^n}(\omega) = \gamma_{t_i^n}^0(\omega)$. Let $\hat{\Omega} := \Omega_0 \cap (\cap_{n,i} \Omega^{n,i})$, then $\mathbb{P}(\hat{\Omega}) = 1$ and for every $\omega \in \hat{\Omega}$ we have $\gamma_{t_i^n}(\omega) = \gamma_{t_i^n}^0(\omega)$ for every $n \in \mathbb{N}$, $i = 0, 1, \dots, 2^n$. By defining, for $n \in \mathbb{N}$, the new sequence of processes

$$\gamma_s^{0,n} := \gamma_t^0 \mathbf{1}_{\{t\}}(s) + \sum_{i=0}^{2^n-1} \gamma_{t_i^n}^0 \mathbf{1}_{(t_i^n, t_{i+1}^n]}(s), \quad s \in [t, T], \quad (\text{III.4.5})$$

we obtain that $\gamma_s^{0,n}$ is $\{\mathcal{F}_s^{t,0}\}$ -predictable and $\gamma_s^{0,n}(\omega) = \gamma_s^n(\omega)$ for every $(s, \omega) \in [t, T] \times \hat{\Omega}$, $n \in \mathbb{N}$. Let

$$\gamma_s^0 := \liminf_{n \rightarrow \infty} \gamma_s^{0,n}, \quad s \in [t, T]. \quad (\text{III.4.6})$$

Then, γ^0 is $\{\mathcal{F}_s^{t,0}\}$ -predictable and if $\omega \in \hat{\Omega}$ we have

$$\gamma_s^0(\omega) := \liminf_{n \rightarrow \infty} \gamma_s^{0,n}(\omega) = \liminf_{n \rightarrow \infty} \gamma_s^n(\omega) = \gamma_s(\omega), \quad \forall s \in [t, T], \quad (\text{III.4.7})$$

i.e., γ^0 and γ are indistinguishable. $\square$

Corollary III.4.2. *Let $\nu = (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_s^t\}, W) \in \mathcal{V}_t$, $\xi \in \mathcal{A}_t^\nu$ and $\eta \in \mathcal{E}_t^\nu$. Then, there exist two $\{\mathcal{F}_s^{t,0}\}$ -predictable process $\xi^0 \in \mathcal{A}_t^\nu$ and $\eta^0 \in \mathcal{E}_t^\nu$ that are, respectively, indistinguishable from ξ and η .*

Proof. Applying Lemma III.4.1 to ξ and η we obtain ξ^0 and η^0 that are $\{\mathcal{F}_s^{t,0}\}$ -predictable. By indistinguishability, it immediately follows that $\xi^0 \in \mathcal{A}_t^\nu$ and $\eta^0 \in \mathcal{E}_t^\nu$. $\square$

Now consider the canonical reference probability system $\nu^* \in \mathcal{V}_t$ (see (III.2.2)). Since Ω^* with the usual supremum norm is a Polish space, then $(\Omega^*, \mathcal{B}(\Omega^*))$ is a standard measurable space (see, e.g., [76, Ch. V, Th. 2.2]) and so ν^* is a standard reference probability system (as in Section III.2).

With the next lemma, we show that, since the filtration $\{\mathcal{F}_s^{t,0}\}$ is generated by the Brownian motion W , an $\{\mathcal{F}_s^{t,0}\}$ -predictable process γ can be viewed as a function $\alpha : [t, T] \times \Omega^* \rightarrow \mathbb{R}^d$ of time and the path of W .

Lemma III.4.3. *Let $\nu = (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_s^t\}, W) \in \mathcal{V}_t$ and γ be an $\mathbb{R}^d$ -valued, $\{\mathcal{F}_s^{t,0}\}$ -predictable process. There exists a $\mathcal{P}_{\Omega^*}/\mathcal{B}(\mathbb{R}^d)$ -measurable function $\alpha : [t, T] \times \Omega^* \rightarrow \mathbb{R}^d$ such that*

$$\gamma_s(\omega) = \alpha(s, W_s(\omega)), \quad (s, \omega) \in [t, T] \times \Omega. \quad (\text{III.4.8})$$

Proof. Define $\beta : [t, T] \times \Omega \rightarrow [t, T] \times \Omega^*$ by $\beta(s, \omega) := (s, W_s(\omega))$ and consider the sets of the form

$$B_1 = (t_1, t_2] \times \{\omega \in \Omega^* : W_s^*(\omega) \in B\}, \quad (\text{III.4.9})$$

$$B_2 = \{t\} \times \{\omega \in \Omega^* : W_t^*(\omega) \in B\}, \quad (\text{III.4.10})$$

where $t \leq s \leq t_1 < t_2 \leq T$ and $B \in \mathcal{B}(\mathbb{R}^d)$. Notice that $\beta^{-1}(B_1) = A_1$ and $\beta^{-1}(B_2) = A_2$ where

$$A_1 = (t_1, t_2] \times \{\omega \in \Omega : W_s(\omega) \in B\}, \quad (\text{III.4.11})$$

$$A_2 = \{t\} \times \{\omega \in \Omega : W_t(\omega) \in B\}. \quad (\text{III.4.12})$$

Since the sets of the form B_1 and B_2 generate $\mathcal{P}_{\Omega^*}$ and the sets of the form A_1 and A_2 generate $\mathcal{P}_{\Omega}$, we obtain that $\mathcal{P}_{\Omega} = \beta^{-1}(\mathcal{P}_{\Omega^*})$. Moreover, since by assumption γ is $\mathcal{P}_{\Omega}/\mathcal{B}(\mathbb{R}^d)$ -measurable (i.e., it is β -measurable), we obtain that there exists a $\mathcal{P}_{\Omega^*}/\mathcal{B}(\mathbb{R}^d)$ -measurable function $\alpha : [t, T] \times \Omega^* \rightarrow \mathbb{R}^d$ such that $\gamma = \alpha \circ \beta$ (see, e.g. [59, Lemma 1.13]) and so we reach the thesis (III.4.8). See Figure III.1 for a visualisation of this result. $\square$

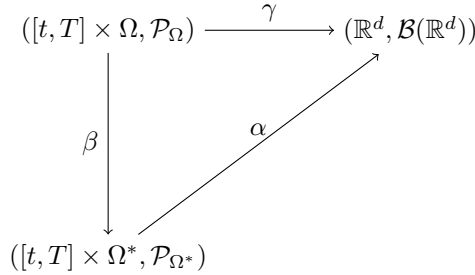


Figure III.1: The existence of a $\mathcal{P}_{\Omega^*}/\mathcal{B}(\mathbb{R}^d)$ -measurable function $\alpha : [t, T] \times \Omega^* \rightarrow \mathbb{R}^d$ such that $\gamma = \alpha \circ \beta$ as in the proof of Lemma III.4.3.

The next lemma shows how the map $\alpha : [t, T] \times \Omega^* \rightarrow \mathbb{R}^d$ in (III.4.8) can be applied to switch from controls that are admissible with respect to a reference probability system ν to controls that have the same joint law and are admissible with respect to a different reference probability system $\tilde{\nu}$.

Lemma III.4.4. *Let $\nu = (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_s^t\}, W) \in \mathcal{V}_t$, $\tilde{\nu} = (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}, \{\tilde{\mathcal{F}}_s^t\}, \tilde{W}) \in \mathcal{V}_t$ and let $\xi \in \mathcal{A}_t^{\nu}$, $\eta \in \mathcal{E}_t^{\nu}$ be $\{\mathcal{F}_s^{t,0}\}$ -predictable. Then, there exist $\tilde{\xi} \in \mathcal{A}_t^{\tilde{\nu}}$ and $\tilde{\eta} \in \mathcal{E}_t^{\tilde{\nu}}$ that are $\{\tilde{\mathcal{F}}_s^{t,0}\}$ -adapted and such that*

$$\mathcal{L}_{\mathbb{P}}(\xi^{\pm}, \eta, W) = \mathcal{L}_{\tilde{\mathbb{P}}}(\tilde{\xi}^{\pm}, \tilde{\eta}, \tilde{W}). \quad (\text{III.4.13})$$

Proof. Notice that in the decomposition $\xi = \xi^{+} - \xi^{-}$ the processes $\xi^{\pm}$ clearly preserve the measurability of ξ and thus $\xi^{\pm}$ are $\{\mathcal{F}_s^{t,0}\}$ -predictable. Since $\xi^{\pm}$ and η are $\{\mathcal{F}_s^{t,0}\}$ -predictable, by Lemma III.4.3, there exist two $\mathcal{P}_{\Omega^*}/\mathcal{B}(\mathbb{R}^d)$ -measurable functions $\alpha_{\xi}^{\pm} : [t, T] \times \Omega^* \rightarrow \mathbb{R}^d$ and a

$\mathcal{P}_{\Omega^*}/\mathcal{B}(\mathbb{R})$ -measurable function $\alpha_\eta : [t, T] \times \Omega^* \rightarrow \mathbb{R}$ such that

$$\xi_s^\pm(\omega) = \alpha_\xi^\pm(s, W(\omega)), \quad (s, \omega) \in [t, T] \times \Omega, \quad (\text{III.4.14})$$

$$\eta_s(\omega) = \alpha_\eta(s, W(\omega)), \quad (s, \omega) \in [t, T] \times \Omega. \quad (\text{III.4.15})$$

The same argument as in the proof of Lemma III.4.3 shows that the map $(s, \tilde{\omega}) \mapsto (s, W(\tilde{\omega}))$ is $\mathcal{P}_{\tilde{\Omega}}/\mathcal{P}_{\Omega^*}$ -measurable. Thus, by defining

$$\hat{\xi}_s^\pm(\tilde{\omega}) := \alpha_\xi^\pm(s, \tilde{W}(\tilde{\omega})), \quad (s, \tilde{\omega}) \in [t, T] \times \tilde{\Omega}, \quad (\text{III.4.16})$$

$$\hat{\eta}_s(\tilde{\omega}) := \alpha_\eta(s, \tilde{W}(\tilde{\omega})), \quad (s, \tilde{\omega}) \in [t, T] \times \tilde{\Omega}, \quad (\text{III.4.17})$$

we obtain, by composition, that $\hat{\xi}^\pm$ and $\hat{\eta}$ are $\{\tilde{\mathcal{F}}_s^{t,0}\}$ -predictable. Let $D := \{t_n\}_n \subseteq [t, T]$ be dense and countable such that $t =: t_0 < t_1 < \dots < t_n < \dots \leq T$. Then, for every $k, n, m \in \mathbb{N}$, we have that

$$\begin{aligned} 1 &= \mathbb{P}(\xi_{t_n}^\pm \leq \xi_{t_{n+k}}^\pm, \xi_{t_m}^\pm = \xi_{t_{m-}}^\pm) \\ &= \mathbb{P}(\alpha_\xi^\pm(t_n, W) \leq \alpha_\xi^\pm(t_{n+k}, W), \alpha_\xi^\pm(t_m, W) = \alpha_\xi^\pm(t_{m-}, W)) \\ &= \tilde{\mathbb{P}}(\alpha_\xi^\pm(t_n, \tilde{W}) \leq \alpha_\xi^\pm(t_{n+k}, \tilde{W}), \alpha_\xi^\pm(t_m, \tilde{W}) = \alpha_\xi^\pm(t_{m-}, \tilde{W})) \\ &= \tilde{\mathbb{P}}(\hat{\xi}_{t_n}^\pm \leq \hat{\xi}_{t_{n+k}}^\pm, \hat{\xi}_{t_m}^\pm = \hat{\xi}_{t_{m-}}^\pm), \end{aligned} \quad (\text{III.4.18})$$

where in the second equality we have used that $\mathcal{L}_{\mathbb{P}}(W) = \mathcal{L}_{\tilde{\mathbb{P}}}(\tilde{W})$. Namely, for every $k, n, m \in \mathbb{N}$ there exists $\tilde{\Omega}_m^{k,n} \subseteq \tilde{\Omega}$ with $\tilde{\mathbb{P}}(\tilde{\Omega}_m^{k,n}) = 1$ such that for every $\omega \in \tilde{\Omega}_m^{k,n}$ we have $\hat{\xi}_{t_n}^\pm(\omega) \leq \hat{\xi}_{t_{n+k}}^\pm(\omega)$ and $\hat{\xi}_{t_m}^\pm(\omega) = \hat{\xi}_{t_{m-}}^\pm(\omega)$. Now let $\hat{\Omega} := \cap_{k,n,m} \tilde{\Omega}_m^{k,n}$, then $\tilde{\mathbb{P}}(\hat{\Omega}) = 1$ and for every $\omega \in \hat{\Omega}$ we have $\hat{\xi}_s^\pm(\omega) \leq \hat{\xi}_u^\pm(\omega)$ and $\hat{\xi}_r^\pm(\omega) = \hat{\xi}_{r-}^\pm(\omega)$ for every $r, s, u \in D$ with $s \leq u$. Thus, we define

$$\tilde{\xi}_s^\pm := \begin{cases} \hat{\xi}_s^\pm & \text{if } s \in D \\ \lim_{\substack{t_n \in D \\ t_n \uparrow s}} \hat{\xi}_{t_n}^\pm & \text{if } s \in [t, T] \setminus D. \end{cases} \quad (\text{III.4.19})$$

Since $\hat{\xi}^\pm$ are $\{\tilde{\mathcal{F}}_s^{t,0}\}$ -predictable, then $\tilde{\xi}^\pm$ are $\{\tilde{\mathcal{F}}_s^{t,0}\}$ -predictable (and so $\{\tilde{\mathcal{F}}_s^{t,0}\}$ -adapted). Moreover, by construction, $\tilde{\xi}^\pm$ is $\tilde{\mathbb{P}}$ -a.s. left-continuous, non-decreasing and $\tilde{\xi}_t^\pm = 0$, $\tilde{\mathbb{P}}$ -a.s. Hence, $\tilde{\xi} := \tilde{\xi}^+ - \tilde{\xi}^-$ is $\tilde{\mathbb{P}}$ -a.s. left-continuous, of bounded variation and $\tilde{\xi}_t = 0$, $\tilde{\mathbb{P}}$ -a.s. To obtain that $\tilde{\xi} \in \mathcal{A}_t^\nu$, we are only left to show that $\tilde{\mathbb{E}}[|\tilde{\xi}_T^\pm|^2] < \infty$ where $\tilde{\mathbb{E}}$ denotes the expectation under the measure $\tilde{\mathbb{P}}$. This will follow when we prove (III.4.13) as $\xi \in \mathcal{A}_t^\nu$. With a very similar construction as for $\tilde{\xi}$, we can build $\tilde{\eta} \in \mathcal{E}_t^\nu$ such that $\tilde{\eta}$ is $\{\tilde{\mathcal{F}}_s^{t,0}\}$ -adapted.

Thus, we now need to show that the equality in law (III.4.13) holds. This is equivalent to show that the corresponding finite-dimensional distributions are the same (see, e.g., [59, Proposition 2.2]). First, let $\{s_k\}_{k=1}^n \subseteq D$ with $t \leq s_1 \leq \dots \leq s_n \leq T$, $N = (2d + d' + 1)n$ and $A \in \mathcal{B}(\mathbb{R}^N)$, then

$$\begin{aligned} \mathbb{P}((\xi_{s_k}^\pm, \eta_{s_k}, W_{s_k})_{k=1}^n \in A) &= \mathbb{P}((\alpha_\xi^\pm(s_k, W), \alpha_\eta(s_k, W), W_{s_k})_{k=1}^n \in A) \\ &= \tilde{\mathbb{P}}((\alpha_\xi^\pm(s_k, \tilde{W}), \alpha_\eta(s_k, \tilde{W}), \tilde{W}_{s_k})_{k=1}^n \in A) \\ &= \tilde{\mathbb{P}}((\hat{\xi}_{s_k}^\pm, \hat{\eta}_{s_k}, \tilde{W}_{s_k})_{k=1}^n \in A) \\ &= \tilde{\mathbb{P}}((\tilde{\xi}_{s_k}^\pm, \tilde{\eta}_{s_k}, \tilde{W}_{s_k})_{k=1}^n \in A), \end{aligned} \quad (\text{III.4.20})$$

where the second equality follows from $\mathcal{L}_{\mathbb{P}}(W) = \mathcal{L}_{\tilde{\mathbb{P}}}(\tilde{W})$; the third equality follows from (III.4.16)-(III.4.17); the fourth equality follows from (III.4.19). Now we want to prove that (III.4.20) still holds for every $\{s_k\}_{k=1}^n \subseteq [t, T]$ with $t \leq s_1 \leq \dots \leq s_n \leq T$ (not only on D) and every $A \in \mathcal{B}(\mathbb{R}^N)$. Let $Y_{s_k} := (\xi_{s_k}^{\pm}, \eta_{s_k}, W_{s_k})$, $\tilde{Y}_{s_k} := (\tilde{\xi}_{s_k}^{\pm}, \tilde{\eta}_{s_k}, \tilde{W}_{s_k})$ for every $k = 1, \dots, n$ and $Y := (Y_{s_1}, \dots, Y_{s_n})$, $\tilde{Y} := (\tilde{Y}_{s_1}, \dots, \tilde{Y}_{s_n})$. Then showing (III.4.20) is equivalent to showing that the characteristic functions of Y and $\tilde{Y}$ are the same, i.e.,

$$\mathbb{E}\left[e^{i\langle \lambda, Y \rangle}\right] = \tilde{\mathbb{E}}\left[e^{i\langle \lambda, \tilde{Y} \rangle}\right], \quad \forall \lambda \in \mathbb{R}^N \quad (\text{III.4.21})$$

For every $k = 1, \dots, n$ let $\{s_k^j\}_j \subseteq D$ be such that $s_k^j \uparrow s_k$ as $j \uparrow \infty$, then for every $\lambda \in \mathbb{R}^N$, we obtain

$$\begin{aligned} \mathbb{E}\left[e^{i\langle \lambda, Y \rangle}\right] &= \mathbb{E}\left[\lim_{j \rightarrow \infty} e^{i\langle \lambda, (Y_{s_1^j}, \dots, Y_{s_n^j}) \rangle}\right] \\ &= \lim_{j \rightarrow \infty} \mathbb{E}\left[e^{i\langle \lambda, (Y_{s_1^j}, \dots, Y_{s_n^j}) \rangle}\right] \\ &= \lim_{j \rightarrow \infty} \tilde{\mathbb{E}}\left[e^{i\langle \lambda, (\tilde{Y}_{s_1^j}, \dots, \tilde{Y}_{s_n^j}) \rangle}\right] \\ &= \tilde{\mathbb{E}}\left[e^{i\langle \lambda, \tilde{Y} \rangle}\right], \end{aligned} \quad (\text{III.4.22})$$

where we have used that the function $y \mapsto e^{i\langle \lambda, y \rangle}$ is continuous and bounded, that the processes are left-continuous and we have applied the dominated convergence theorem and (III.4.20) (which holds for $\{s_k^j\}_j \subseteq D$).

Thus, it follows that (III.4.20) holds also for every $\{s_k\}_{k=1}^n \subseteq [t, T]$ and so

$$\mathcal{L}_{\mathbb{P}}(\xi^{\pm}, \eta, W) = \mathcal{L}_{\tilde{\mathbb{P}}}(\tilde{\xi}^{\pm}, \tilde{\eta}, \tilde{W}), \quad (\text{III.4.23})$$

and the proof is complete. $\square$

Let us now consider the framework of Lemma III.4.4. As the process $X^{\xi, \nu}$ depends only on ξ and on W and $\rho^{\xi, \nu}$ is the first exit time of $X^{\xi, \nu}$ from $\mathcal{O}$, it is reasonable to assume that, under some conditions, $\mathcal{L}_{\mathbb{P}}(\xi^{\pm}, \eta, W) = \mathcal{L}_{\tilde{\mathbb{P}}}(\tilde{\xi}^{\pm}, \tilde{\eta}, \tilde{W})$ yields

$$\mathcal{L}_{\mathbb{P}}(\xi^{\pm}, \eta, X^{\xi, \nu}, \rho^{\xi, \nu}) = \mathcal{L}_{\tilde{\mathbb{P}}}(\tilde{\xi}^{\pm}, \tilde{\eta}, X^{\tilde{\xi}, \tilde{\nu}}, \rho^{\tilde{\xi}, \tilde{\nu}}). \quad (\text{III.4.24})$$

In our analysis, we assume that (III.4.24) holds and in Proposition III.10.1 we will show that, for instance, sub-linear growth and Lipschitz continuity of μ and σ are sufficient to obtain this result.

We are now ready to prove the independence of the reference probability system (III.4.1).

Proposition III.4.5. *Let $\nu = (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_s^t\}, W)$, $\tilde{\nu} = (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}, \{\tilde{\mathcal{F}}_s^t\}, \tilde{W}) \in \mathcal{V}_t$. For every $(\tau, \xi) \in \mathcal{T}_t^{\nu} \times \mathcal{A}_t^{\nu}$ there exists $(\tilde{\tau}, \tilde{\xi}) \in \mathcal{T}_t^{\tilde{\nu}} \times \mathcal{A}_t^{\tilde{\nu}}$ such that*

$$J_{t,x}^{\nu}(\tau, \xi) = J_{t,x}^{\tilde{\nu}}(\tilde{\tau}, \tilde{\xi}), \quad \forall x \in \mathbb{R}^d. \quad (\text{III.4.25})$$

Consequently, for every $(t, x) \in [0, T] \times \mathbb{R}^d$ and $\nu \in \mathcal{V}_t$, we obtain

$$v_t^{\nu}(x) = v(t, x). \quad (\text{III.4.26})$$

Proof. Let $(\tau, \xi) \in \mathcal{T}_t^{\nu} \times \mathcal{A}_t^{\nu}$ and, as in (III.4.2), let $\eta_s = \eta_s^{\tau} := \mathbf{1}_{\{\tau < s\}}$ be such that $\eta \in \mathcal{E}_t^{\nu}$

and $J_{t,x}^\nu(\tau, \xi) = I_{t,x}^\nu(\eta, \xi)$ for every $x \in \mathbb{R}^d$. By applying Corollary III.4.2 to both $\xi \in \mathcal{A}_t^\nu$ and $\eta \in \mathcal{E}_t^\nu$, we obtain that there exist two $\{\mathcal{F}_s^{t,0}\}$ -predictable controls $\xi^0 \in \mathcal{A}_t^\nu$ and $\eta^0 \in \mathcal{E}_t^\nu$ that are indistinguishable from ξ and η , respectively. By applying Lemma III.4.4 to both $\xi^0 \in \mathcal{A}_t^\nu$ and $\eta^0 \in \mathcal{E}_t^\nu$, we obtain that there exist $\tilde{\xi} \in \mathcal{A}_t^{\tilde{\nu}}$ and $\tilde{\eta} \in \mathcal{E}_t^{\tilde{\nu}}$ such that

$$\mathcal{L}_{\mathbb{P}}(\xi^\pm, \eta, W) = \mathcal{L}_{\mathbb{P}}(\xi^{0,\pm}, \eta^0, W) = \mathcal{L}_{\tilde{\mathbb{P}}}(\tilde{\xi}^\pm, \tilde{\eta}, \tilde{W}). \quad (\text{III.4.27})$$

Thus, by (III.4.24), we have that that

$$\mathcal{L}_{\mathbb{P}}(\xi^\pm, \eta, X^{t,x;\xi,\nu}, \rho^{t,x;\xi,\nu}) = \mathcal{L}_{\tilde{\mathbb{P}}}(\tilde{\xi}^\pm, \tilde{\eta}, X^{t,x;\tilde{\xi},\tilde{\nu}}, \rho^{t,x;\tilde{\xi},\tilde{\nu}}), \quad \forall x \in \mathbb{R}^d. \quad (\text{III.4.28})$$

This implies that $I_{t,x}^\nu(\eta, \xi) = I_{t,x}^{\tilde{\nu}}(\tilde{\eta}, \tilde{\xi})$ for every $x \in \mathbb{R}^d$. Moreover, as in (III.4.3), let $\tilde{\tau} = \tilde{\tau}^{\tilde{\eta}} := \inf\{s \geq t : \tilde{\eta}_s = 1\} \wedge T \in \mathcal{T}_t^{\tilde{\nu}}$ for which $I_{t,x}^{\tilde{\nu}}(\tilde{\eta}, \tilde{\xi}) = J_{t,x}^{\tilde{\nu}}(\tilde{\tau}, \tilde{\xi})$. Thus,

$$J_{t,x}^\nu(\tau, \xi) = I_{t,x}^\nu(\eta, \xi) = I_{t,x}^{\tilde{\nu}}(\tilde{\eta}, \tilde{\xi}) = J_{t,x}^{\tilde{\nu}}(\tilde{\tau}, \tilde{\xi}), \quad \forall x \in \mathbb{R}^d. \quad (\text{III.4.29})$$

and result (III.4.25) is proved. This also implies that

$$J_{t,x}^\nu(\tau, \xi) = J_{t,x}^{\tilde{\nu}}(\tilde{\tau}, \tilde{\xi}) \leq v_t^{\tilde{\nu}}(x), \quad \forall (\tau, \xi) \in \mathcal{T}_t^\nu \times \mathcal{A}_t^\nu \quad (\text{III.4.30})$$

Hence,

$$v_t^\nu(x) = \sup_{(\tau, \xi) \in \mathcal{T}_t^\nu \times \mathcal{A}_t^\nu} J_{t,x}^\nu(\tau, \xi) \leq v_t^{\tilde{\nu}}(x). \quad (\text{III.4.31})$$

Symmetrically, we obtain $v_t^{\tilde{\nu}}(x) \leq v_t^\nu(x)$ and so we have that $v_t^{\tilde{\nu}}(x) = v_t^\nu(x)$ for every $\nu, \tilde{\nu} \in \mathcal{V}_t$.

Finally, let $\nu \in \mathcal{V}_t$ and $\varepsilon > 0$. By definition (III.3.4), there exists $\nu^\varepsilon \in \mathcal{V}_t$ such that

$$v(t, x) < v_t^{\nu^\varepsilon}(x) + \varepsilon. \quad (\text{III.4.32})$$

Furthermore, from what we have just proved, we obtain that $v_t^{\nu^\varepsilon}(x) = v_t^\nu(x)$ and thus

$$v(t, x) < v_t^\nu(x) + \varepsilon. \quad (\text{III.4.33})$$

By letting $\varepsilon \rightarrow 0$, we have that $v(t, x) \leq v_t^\nu(x)$. By definition (III.3.4), the opposite inequality is trivial and so $v(t, x) = v_t^\nu(x)$. $\square$

Unless specified otherwise, in the following sections we shall fix a starting point $(t, x) \in [0, T] \times \mathbb{R}^d$ and without loss of generality - in view of Proposition III.4.5 - a standard reference probability system $\nu = (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_s^t\}, W) \in \mathcal{V}_t$.

III.5 Regular conditional probabilities

In this section, after recalling the definition of regular conditional probability (and some related results), we show how this tool allow us to construct reference probability systems starting at future times $u \in [t, T]$ that preserve the history of the problem up to time u (see Proposition III.5.4). This is an essential result in order to obtain the dynamic programming principle for our problem.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{G} \subseteq \mathcal{F}$ be a sub σ -algebra. A function $p : \Omega \times \mathcal{F} \rightarrow [0, 1]$ is a regular conditional probability (given $\mathcal{G}$) if

- (i) for every $\omega \in \Omega$, the function $\mathbb{P}_\omega := p(\omega, \cdot)$ is a probability measure on $(\Omega, \mathcal{F})$;
- (ii) for every $A \in \mathcal{F}$, the function $p(\cdot, A)$ is $\mathcal{G}$ -measurable;
- (iii) for every $A \in \mathcal{F}$, we have $\mathbb{P}_\omega(A) = \mathbb{E}^\mathbb{P}[\mathbf{1}_A | \mathcal{G}](\omega) =: \mathbb{P}(A | \mathcal{G})$ for $\mathbb{P}$ -a.e. ω .

It thus follows that, if $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, then

$$\mathbb{E}^\mathbb{P}[X | \mathcal{G}](\omega) = \mathbb{E}^\mathbb{P}_\omega[X] := \int_\Omega X(\omega') d\mathbb{P}_\omega(\omega'), \quad \mathbb{P}\text{-a.e. } \omega. \quad (\text{III.5.1})$$

This can be easily shown by a standard procedure, i.e., noticing that (III.5.1) holds trivially for simple random variables (by (iii) above), thus it also holds for non-negative random variables (by monotone convergence theorem) and finally for $X = [X]^+ - [X]^- \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ (by linearity).

If $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, $(\Omega, \mathcal{F})$ is a standard measurable space and $\mathcal{G} \subseteq \mathcal{F}$ is a sub σ -algebra, then there exists a regular conditional probability $p : \Omega \times \mathcal{F} \rightarrow [0, 1]$ which is $\mathbb{P}$ -a.s. unique. Namely, if $p' : \Omega \times \mathcal{F} \rightarrow [0, 1]$ is another regular conditional probability then there exists $N \in \mathcal{G}$ with $\mathbb{P}(N) = 0$ such that if $\omega \in \Omega \setminus N$ then $p(\omega, A) = p'(\omega, A)$ for every $A \in \mathcal{F}$. Moreover, if Γ is a Polish space and $\gamma : \Omega \rightarrow \Gamma$ is $\mathcal{G}/\mathcal{B}(\Gamma)$ measurable then

$$\mathbb{P}_\omega(\{\omega' \in \Omega : \gamma(\omega') = \gamma(\omega)\}) = 1, \quad \mathbb{P}\text{-a.e. } \omega, \quad (\text{III.5.2})$$

i.e., γ is $\mathbb{P}_\omega$ -a.s. equal to the constant $\gamma(\omega)$ for $\mathbb{P}$ -a.e. ω . For a proof of these results see, e.g., Theorem 3.1, Theorem 3.2 and the subsequent Corollary of Chapter I in [53].

Let $\nu = (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_s^t\}, W) \in \mathcal{V}_t$ be the fixed standard reference probability system. Recall that $\mathcal{F}$ is the completion of $\mathcal{F}^0$ with the $\mathbb{P}$ -null sets and that $(\Omega, \mathcal{F}^0)$ is a standard measurable space. Moreover, it is not difficult to show that the completion $\mathcal{F}$ of a σ -algebra $\mathcal{F}^0$ is determined by

$$\mathcal{F} = \{A \cup N_1 : A \in \mathcal{F}^0, N_1 \subseteq N \in \mathcal{F}^0, \mathbb{P}(N) = 0\}. \quad (\text{III.5.3})$$

In the following discussion we fix $u \in [t, T]$ and the subsequent notation. We denote by $\mathbb{P}_\omega$ the regular conditional probability given $\mathcal{F}_u^{t,0}$, which satisfies $\mathbb{P}_\omega(A) = \mathbb{P}(A | \mathcal{F}_u^{t,0})(\omega)$ for every $A \in \mathcal{F}$ for $\mathbb{P}$ -a.e. ω , and by $\mathcal{F}_\omega$ the completion of $\mathcal{F}^0$ with the $\mathbb{P}_\omega$ -null sets. Thus,

$$\mathcal{F}_\omega = \{A \cup N_1 : A \in \mathcal{F}^0, N_1 \subseteq N \in \mathcal{F}^0, \mathbb{P}_\omega(N) = 0\}. \quad (\text{III.5.4})$$

We define $W_s^u := W_s - W_u$ for $s \in [u, T]$, $\mathcal{F}_s^{u,0} := \sigma(W_r^u : r \in [u, s])$ and we denote by $\mathcal{F}_s^{u,\omega}$ the augmentation of $\mathcal{F}_s^{u,0}$ with the $\mathbb{P}_\omega$ -null sets. For the sake of simplicity, we also denote by $\mathbb{E}$ the expectation with respect to $\mathbb{P}$ and by $\mathbb{E}_\omega$ the expectation with respect to $\mathbb{P}_\omega$.

In this framework, we have the following results.

Lemma III.5.1. *Let $\Omega_1 \in \mathcal{F}$ be such that $\mathbb{P}(\Omega_1) = 1$ then, for $\mathbb{P}$ -a.e. ω , we have that $\Omega_1 \in \mathcal{F}_\omega$ and $\mathbb{P}_\omega(\Omega_1) = 1$. Conversely, let $\Omega' \in \mathcal{F}^0$ be such that $\mathbb{P}_\omega(\Omega') = 1$ for $\mathbb{P}$ -a.e. ω , then $\mathbb{P}(\Omega') = 1$.*

Proof. Since $\Omega_1 \in \mathcal{F}$, then $\Omega_1 = \Omega_2 \cup N_1$ where $\Omega_2 \in \mathcal{F}^0$ and $N_1 \subseteq N \in \mathcal{F}^0$ with $\mathbb{P}(N) = 0$. Therefore,

$$1 = \mathbb{P}(\Omega_2) = \mathbb{E}[\mathbb{E}[\mathbf{1}_{\Omega_2} | \mathcal{F}_u^{t,0}](\omega)] = \mathbb{E}[\mathbb{P}_\omega(\Omega_2)], \quad (\text{III.5.5})$$

and so $\mathbb{P}_\omega(\Omega_2) = 1$ for $\mathbb{P}$ -a.e. ω . Similarly, we obtain $\mathbb{P}_\omega(N) = 0$ for $\mathbb{P}$ -a.e. ω and thus $N_1 \in \mathcal{F}_\omega$

for $\mathbb{P}$ -a.e. ω . Hence, $\Omega_1 \in \mathcal{F}_\omega$ and $\mathbb{P}_\omega(\Omega_1) = 1$ for $\mathbb{P}$ -a.e. ω . The fact that $\mathbb{P}(\Omega') = 1$ follows in the same way as in (III.5.5). $\square$

Lemma III.5.2. *Let $Y : \Omega \rightarrow \mathbb{R}^d$ be such that $Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ then, for $\mathbb{P}$ -a.e. ω , we have that $Y \in L^1(\Omega, \mathcal{F}_\omega, \mathbb{P}_\omega)$ and $\mathbb{E}_\omega[Y] = \mathbb{E}[Y|\mathcal{F}_\omega^t](\omega)$.*

Proof. Since $Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ (in particular, it is $\mathcal{F}$ -measurable), then (see, e.g., [59, Lemma 1.25]) there exists $Y^0 \in L^1(\Omega, \mathcal{F}^0, \mathbb{P})$ such that $Y = Y^0$, $\mathbb{P}$ -a.s. Thus, by Lemma III.5.1, we also have that $Y = Y^0$, $\mathbb{P}_\omega$ -a.s. for $\mathbb{P}$ -a.e. ω . Namely, for $\mathbb{P}$ -a.e. ω , the set $N := \{Y \neq Y^0\} \in \mathcal{F}_\omega$, $\mathbb{P}_\omega(N) = 0$ and $Y = Y^0 \mathbf{1}_{\Omega \setminus N} + Y \mathbf{1}_N$. It is immediate to see that $Y^0 \mathbf{1}_{\Omega \setminus N}$ is $\mathcal{F}_\omega$ -measurable for $\mathbb{P}$ -a.e. ω . Now let $B \in \mathcal{B}(\mathbb{R}^d)$ such that $0 \in B$, then

$$\begin{aligned} (Y \mathbf{1}_N)^{-1}(B) &= \{\omega' \in \Omega \setminus N : Y(\omega') \mathbf{1}_N(\omega') \in B\} \cup \{\omega' \in N : Y(\omega') \mathbf{1}_N(\omega') \in B\} \\ &= (\Omega \setminus N) \cup \{\omega' \in N : Y(\omega') \mathbf{1}_N(\omega') \in B\} \in \mathcal{F}_\omega, \quad \mathbb{P}\text{-a.e. } \omega. \end{aligned} \quad (\text{III.5.6})$$

Similarly, if $B \in \mathcal{B}(\mathbb{R}^d)$ is such that $0 \notin B$, then

$$(Y \mathbf{1}_N)^{-1}(B) = \{\omega' \in N : Y(\omega') \mathbf{1}_N(\omega') \in B\} \in \mathcal{F}_\omega, \quad \mathbb{P}\text{-a.e. } \omega. \quad (\text{III.5.7})$$

Therefore, Y is $\mathcal{F}_\omega$ -measurable for $\mathbb{P}$ -a.e. ω . Furthermore, for $\mathbb{P}$ -a.e. ω , we have that $\mathbb{E}_\omega[|Y|] = \mathbb{E}[|Y||\mathcal{F}_\omega^t](\omega) < \infty$ and so $Y \in L^1(\Omega, \mathcal{F}_\omega, \mathbb{P}_\omega)$. Finally, if $A \in \mathcal{F}_\omega^t$ then $A = A_1 \cup N_1$ with $A_1 \in \mathcal{F}_\omega^{t,0}$ and $\mathbb{P}(N_1) = 0$. Thus,

$$\mathbb{E}[\mathbb{E}[Y|\mathcal{F}_\omega^t] \mathbf{1}_A] = \mathbb{E}[\mathbb{E}[Y|\mathcal{F}_\omega^{t,0}] \mathbf{1}_{A_1 \cup N_1}] = \mathbb{E}[\mathbb{E}[Y|\mathcal{F}_\omega^{t,0}] \mathbf{1}_{A_1}] = \mathbb{E}[Y \mathbf{1}_{A_1}] = \mathbb{E}[Y \mathbf{1}_A]. \quad (\text{III.5.8})$$

Hence, $\mathbb{E}[Y|\mathcal{F}_\omega^t] = \mathbb{E}[Y|\mathcal{F}_\omega^{t,0}] = \mathbb{E}[Y^0|\mathcal{F}_\omega^{t,0}]$, $\mathbb{P}$ -a.s. and so, for $\mathbb{P}$ -a.e. ω , we have

$$\mathbb{E}[Y|\mathcal{F}_\omega^t](\omega) = \mathbb{E}[Y^0|\mathcal{F}_\omega^{t,0}](\omega) = \int_\Omega Y^0(\omega') d\mathbb{P}_\omega(\omega') = \int_\Omega Y(\omega') d\mathbb{P}_\omega(\omega') = \mathbb{E}_\omega[Y], \quad (\text{III.5.9})$$

where we have used property (III.5.1), and the fact that Y is $\mathcal{F}_\omega$ -measurable and $Y = Y^0$, $\mathbb{P}_\omega$ -a.s. for $\mathbb{P}$ -a.e. ω . $\square$

In the upcoming Proposition III.5.4 we show how to generate, from the standard reference probability system $\nu \in \mathcal{V}_t$ starting at time $t \in [0, T]$, new reference probability systems ν_ω starting at time $u \in [t, T]$ such that $\nu_\omega \in \mathcal{V}_u$ for $\mathbb{P}$ -a.e. ω . Analogously, from the admissible controls $\xi \in \mathcal{A}_t^\nu$ and $\eta \in \mathcal{E}_t^\nu$, we can generate the admissible controls $\xi^{0,u} \in \mathcal{A}_u^{\nu_\omega}$ and $\eta^{0,u} \in \mathcal{E}_u^{\nu_\omega}$ for $\mathbb{P}$ -a.e. ω . The probability measure of the systems $\nu_\omega \in \mathcal{V}_u$ is the regular conditional probability $\mathbb{P}_\omega = \mathbb{P}(\cdot|\mathcal{F}_\omega^{t,0})(\omega)$ which preserves the history of the problem up to time u and is thus the key to obtain the dynamic programming principle. The proof of Proposition III.5.4 is supported by the following lemma.

Lemma III.5.3. *For $\mathbb{P}$ -a.e. ω , we have that $\mathcal{F}_s^{t,0} \subseteq \mathcal{F}_s^{u,\omega}$ for every $s \in [u, T]$. In particular, $\mathcal{F}_s^{u,\omega}$ is the augmentation of $\mathcal{F}_s^{t,0}$ with the $\mathbb{P}_\omega$ -null sets.*

Proof. Let $D := \{t_n\}_{n \in \mathbb{N}}$ be a dense subset of $[t, T]$. As W has continuous paths, we have that $\mathcal{F}_s^{t,0} = \sigma(W_r : r \in D \cap [t, s])$. To show that, for $\mathbb{P}$ -a.e. ω , we have $\mathcal{F}_s^{t,0} \subseteq \mathcal{F}_s^{u,\omega}$ for every $s \in [u, T]$, we proceed in two steps.

Step 1. Let $n, m \in \mathbb{N}$, $B \in \mathcal{B}(\mathbb{R}^{d'})$, $s_m \in D \cap [u, T]$ and $t_n \in D \cap [t, s_m]$. Here, we show that there exists $\Omega_{m,n} \subseteq \Omega$ such that $\mathbb{P}(\Omega_{m,n}) = 1$ and for every $\omega \in \Omega_{m,n}$ we have $W_{t_n}^{-1}(B) \in \mathcal{F}_{s_m}^{u,\omega}$. This

implies that $\mathcal{F}_{s_m}^{t,0} \subseteq \mathcal{F}_{s_m}^{u,\omega}$.

If $t_n \leq u$, by property (III.5.2), we obtain that $W_{t_n} = W_{t_n}(\omega)$, $\mathbb{P}_\omega$ -a.s. for $\mathbb{P}$ -a.e. ω . Thus, for $\mathbb{P}$ -a.e. ω , we have $N_\omega := \{\omega' \in \Omega : W_{t_n}(\omega') \neq W_{t_n}(\omega)\} \in \mathcal{F}^0$ and $\mathbb{P}_\omega(N_\omega) = 0$. Hence, for $\mathbb{P}$ -a.e. ω ,

$$\begin{aligned} W_{t_n}^{-1}(B) &= \{\omega' \in \Omega \setminus N_\omega : W_{t_n}(\omega') \in B\} \cup \{\omega' \in N_\omega : W_{t_n}(\omega') \in B\} \\ &= \{\omega' \in \Omega \setminus N_\omega : W_{t_n}(\omega) \in B\} \cup \{\omega' \in N_\omega : W_{t_n}(\omega') \in B\}. \end{aligned} \quad (\text{III.5.10})$$

Therefore, for $\mathbb{P}$ -a.e. ω ,

- (i) if $W_{t_n}(\omega) \notin B$ then $W_{t_n}^{-1}(B) \subseteq N_\omega \Rightarrow W_{t_n}^{-1}(B) \in \mathcal{F}_{s_m}^{u,\omega}$;
- (ii) if $W_{t_n}(\omega) \in B$ then $\Omega \setminus N_\omega \subseteq W_{t_n}^{-1}(B)$, i.e., $(W_{t_n}^{-1}(B))^c \subseteq N_\omega \Rightarrow W_{t_n}^{-1}(B) \in \mathcal{F}_{s_m}^{u,\omega}$.

If $u < t_n \leq s_m$, again by property (III.5.2), we have $N_\omega := \{\omega' \in \Omega : W_u(\omega') \neq W_u(\omega)\} \in \mathcal{F}^0$ and $\mathbb{P}_\omega(N_\omega) = 0$ for $\mathbb{P}$ -a.e. ω . Then, for $\mathbb{P}$ -a.e. ω ,

$$\begin{aligned} W_{t_n}^{-1}(B) &= \{\omega' \in \Omega \setminus N_\omega : W_{t_n}(\omega') \in B\} \cup \{\omega' \in N_\omega : W_{t_n}(\omega') \in B\} \\ &= \{\omega' \in \Omega \setminus N_\omega : W_{t_n}^u(\omega') + W_u(\omega') \in B\} \cup \{\omega' \in N_\omega : W_{t_n}(\omega') \in B\} \\ &= \{\omega' \in \Omega \setminus N_\omega : W_{t_n}^u(\omega') \in (B - W_u(\omega))\} \cup \{\omega' \in N_\omega : W_{t_n}(\omega') \in B\}. \end{aligned} \quad (\text{III.5.11})$$

Since $\{\omega' \in \Omega \setminus N_\omega : W_{t_n}^u(\omega') \in (B - W_u(\omega))\} \in \mathcal{F}_{s_m}^{u,\omega}$ and $\{\omega' \in N_\omega : W_{t_n}(\omega') \in B\} \subseteq N_\omega$ with $\mathbb{P}_\omega(N_\omega) = 0$ for $\mathbb{P}$ -a.e. ω , we still obtain that $W_{t_n}^{-1}(B) \in \mathcal{F}_{s_m}^{u,\omega}$.

Step 2. Let $\hat{\Omega} := \cap_{m,n \in \mathbb{N}} \Omega_{m,n}$. Thus, $\mathbb{P}(\hat{\Omega}) = 1$ and for every $\omega \in \hat{\Omega}$ we have that $W_r^{-1}(B) \in \mathcal{F}_s^{u,\omega}$ for every $s \in D \cap [u, T]$, $r \in D \cap [t, s]$ and $B \in \mathcal{B}(\mathbb{R}^{d'})$. Therefore, for $\mathbb{P}$ -a.e. ω , we obtain

$$\mathcal{F}_s^{t,0} = \sigma(W_r : r \in D \cap [t, s]) \subseteq \mathcal{F}_s^{u,\omega}, \quad \forall s \in D \cap [u, T]. \quad (\text{III.5.12})$$

If $s \in [u, T] \setminus D$, then by left-continuity of $\{\mathcal{F}_s^{t,0}\}$, we have

$$\mathcal{F}_s^{t,0} = \sigma(\cup_{r \in D \cap [t, s)} \mathcal{F}_r^{t,0}) \subseteq \mathcal{F}_s^{u,\omega}, \quad \forall \omega \in \hat{\Omega}. \quad (\text{III.5.13})$$

Thus, we have shown that, for $\mathbb{P}$ -a.e. ω , we have that $\mathcal{F}_s^{t,0} \subseteq \mathcal{F}_s^{u,\omega}$ for every $s \in [u, T]$. Finally, since $\mathcal{F}_s^{u,0} := \sigma(W_r^u : r \in [u, s]) \subseteq \mathcal{F}_s^{t,0}$, then, for $\mathbb{P}$ -a.e. ω , we obtain that $\mathcal{F}_s^{u,\omega}$ is the augmentation of $\mathcal{F}_s^{t,0}$ by the $\mathbb{P}_\omega$ -null sets. $\square$

Proposition III.5.4. *Let $\xi \in \mathcal{A}_t^\nu$ and $\eta \in \mathcal{E}_t^\nu$. Then, for $\mathbb{P}$ -a.e. ω , we have that*

$$\nu_\omega := (\Omega, \mathcal{F}_\omega, \mathbb{P}_\omega, \{\mathcal{F}_s^{u,\omega}\}, W^u) \in \mathcal{V}_u, \quad \xi^{0,u} \in \mathcal{A}_u^{\nu_\omega}, \quad \eta^{0,u} \in \mathcal{E}_u^{\nu_\omega}, \quad (\text{III.5.14})$$

where ξ^0 and η^0 are the $\{\mathcal{F}_s^{t,0}\}$ -adapted controls from Corollary III.4.2 and $\xi^{0,u}$ and $\eta^{0,u}$ are defined for $s \in [u, T]$ as $\xi_s^{0,u} := \xi_s^0 - \xi_u^0$, $\eta_s^{0,u} := \eta_s^0 - \eta_u^0$.

Proof. We first want to prove that $\nu_\omega \in \mathcal{V}_u$ for $\mathbb{P}$ -a.e. ω . By construction, $(\Omega, \mathcal{F}_\omega, \mathbb{P}_\omega)$ is a complete probability space. Thus, we just need to show that, for $\mathbb{P}$ -a.e. ω , the process W^u is a Brownian motion on $(\Omega, \mathcal{F}_\omega, \mathbb{P}_\omega)$ such that $\mathbb{P}_\omega(W_u^u = 0) = 1$. Obviously, $W_u^u = 0$ and W^u has continuous paths. Thus, showing that, for $\mathbb{P}$ -a.e. ω , the process W^u is a Brownian motion on $(\Omega, \mathcal{F}_\omega, \mathbb{P}_\omega)$ it is

equivalent to showing that, for $\mathbb{P}$ -a.e. ω , we have (see, e.g., [73, Proposition 1.2.7])

$$\mathbb{E}_\omega[e^{i\langle y, W_s^u - W_r^u \rangle} | \mathcal{F}_r^{u, \omega}] = e^{-\frac{|y|^2}{2}(s-r)}, \quad u \leq r < s \leq T, \quad y \in \mathbb{R}^{d'}. \quad (\text{III.5.15})$$

Since, for $\mathbb{P}$ -a.e. ω , $\mathcal{F}_r^{u, \omega}$ is the augmentation of $\mathcal{F}_r^{t, 0}$ with the $\mathbb{P}_\omega$ -null sets (see Lemma III.5.3), we obtain (with an argument as in the proof of Lemma III.5.2) that, for $\mathbb{P}$ -a.e. ω ,

$$\mathbb{E}_\omega[e^{i\langle y, W_s^u - W_r^u \rangle} | \mathcal{F}_r^{u, \omega}] = \mathbb{E}_\omega[e^{i\langle y, W_s^u - W_r^u \rangle} | \mathcal{F}_r^{t, 0}], \quad u \leq r < s \leq T, \quad y \in \mathbb{R}^{d'}. \quad (\text{III.5.16})$$

Thus, to prove (III.5.15) we need to show that, for $\mathbb{P}$ -a.e. ω , we have that

$$\mathbb{E}_\omega[e^{-\frac{|y|^2}{2}(s-r)} \mathbf{1}_A] = \mathbb{E}_\omega[e^{i\langle y, W_s^u - W_r^u \rangle} \mathbf{1}_A], \quad u \leq r < s \leq T, \quad y \in \mathbb{R}^{d'}, \quad A \in \mathcal{F}_r^{t, 0}. \quad (\text{III.5.17})$$

Now let $r \in [u, T] \cap \mathbb{Q}$, $D := [t, r] \cap \mathbb{Q}$ and $(B_k)_{k \in \mathbb{N}}$ be a countable generator of $\mathcal{B}(\mathbb{R}^{d'})$. Since W is continuous, we have that $\mathcal{F}_r^{t, 0} = \sigma(\mathcal{G})$ where $\mathcal{G} = \{W_q^{-1}(B_k), q \in D, k \in \mathbb{N}\}$ is a countable generator. Let $\pi(\mathcal{G})$ be the π -system generated by $\mathcal{G}$, then $\pi(\mathcal{G})$ is a countable π -system generating $\mathcal{F}_r^{t, 0}$. Thus, for $s \in (r, T] \cap \mathbb{Q}$, $y \in \mathbb{Q}^{d'}$ and $A \in \pi(\mathcal{G})$, we have

$$\begin{aligned} \mathbb{E}_\omega[e^{-\frac{|y|^2}{2}(s-r)} \mathbf{1}_A] &= \mathbb{E}_\omega[\mathbb{E}[e^{i\langle y, W_s - W_r \rangle} | \mathcal{F}_r^t] \mathbf{1}_A] \\ &= \mathbb{E}_\omega[\mathbb{E}[e^{i\langle y, W_s^u - W_r^u \rangle} \mathbf{1}_A | \mathcal{F}_r^t]] \\ &= \mathbb{E}[\mathbb{E}[e^{i\langle y, W_s^u - W_r^u \rangle} \mathbf{1}_A | \mathcal{F}_r^t] | \mathcal{F}_u^{t, 0}](\omega) \\ &= \mathbb{E}_\omega[e^{i\langle y, W_s^u - W_r^u \rangle} \mathbf{1}_A], \quad \mathbb{P}\text{-a.e. } \omega, \end{aligned} \quad (\text{III.5.18})$$

where in the first line we have used that W is a Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$, in the third line we have used the regular conditional probability property (III.5.1) and in the fourth line we have used tower property. Since r, s and y were taken in countable sets and $\pi(\mathcal{G})$ is a countable π -system generating $\mathcal{F}_r^{t, 0}$, then, for $\mathbb{P}$ -a.e. ω , we have that (see, e.g., [6, Theorem 1.1])

$$\mathbb{E}_\omega[e^{-\frac{|y|^2}{2}(s-r)} \mathbf{1}_A] = \mathbb{E}_\omega[e^{i\langle y, W_s^u - W_r^u \rangle} \mathbf{1}_A], \quad (\text{III.5.19})$$

for every $r \in [u, T] \cap \mathbb{Q}$, $s \in (r, T] \cap \mathbb{Q}$, $y \in \mathbb{Q}^{d'}$ and $A \in \mathcal{F}_r^{t, 0}$. Now let $r \in [u, T]$, $s \in (r, T]$, $y \in \mathbb{R}^{d'}$ and $(r_n)_{n \in \mathbb{N}} \subseteq [u, T] \cap \mathbb{Q}$, $(s_n)_{n \in \mathbb{N}} \subseteq (r, T] \cap \mathbb{Q}$ and $(y_n)_{n \in \mathbb{N}} \subseteq \mathbb{Q}^{d'}$ such that $r_n \rightarrow r$, $s_n \rightarrow s$ and $y_n \rightarrow y$ as $n \rightarrow \infty$. Then, for $\mathbb{P}$ -a.e. ω , we have

$$\begin{aligned} \mathbb{E}_\omega[e^{-\frac{|y|^2}{2}(s-r)} \mathbf{1}_A] &= \lim_{n \rightarrow \infty} \mathbb{E}_\omega[e^{-\frac{|y_n|^2}{2}(s_n - r_n)} \mathbf{1}_A] \\ &= \lim_{n \rightarrow \infty} \mathbb{E}_\omega[e^{i\langle y_n, W_{s_n}^u - W_{r_n}^u \rangle} \mathbf{1}_A] \\ &= \mathbb{E}_\omega[e^{i\langle y, W_s^u - W_r^u \rangle} \mathbf{1}_A], \\ &\quad \forall u \leq r < s \leq T, \quad y \in \mathbb{R}^d, \quad A \in \mathcal{F}_r^{t, 0}, \end{aligned} \quad (\text{III.5.20})$$

where in the first and in the third equality we have used the dominated convergence theorem and in the second equality we have used (III.5.19). Thus, (III.5.17) holds which in turns implies our desired result (III.5.15).

Finally, it remains to prove that $\xi^{0, u} \in \mathcal{A}_u^{\nu_\omega}$ and $\eta^{0, u} \in \mathcal{E}_u^{\nu_\omega}$ for $\mathbb{P}$ -a.e. ω . Since, for $\mathbb{P}$ -a.e. ω , we have $\mathcal{F}_s^{t, 0} \subseteq \mathcal{F}_s^{u, \omega}$ for every $s \in [u, T]$ (see Lemma III.5.3), it follows that $\xi^{0, u}$ and $\eta^{0, u}$ are $\{\mathcal{F}_s^{u, \omega}\}$ -adapted for $\mathbb{P}$ -a.e. ω . Since ξ^0 is left-continuous and of bounded variation $\mathbb{P}$ -a.s., then, by

Lemma III.5.1, $\xi^{0,u}$ is left-continuous and of bounded variation $\mathbb{P}_\omega$ -a.s. for $\mathbb{P}$ -a.e. ω . Moreover,

$$\mathbb{E}[\mathbb{E}_\omega[|\xi_T^{0,u,\pm}|^2]] \leq \mathbb{E}[\mathbb{E}[|\xi_T^{0,\pm}|^2 + |\xi_u^{0,\pm}|^2 | \mathcal{F}_u^t](\omega)] \leq 4\mathbb{E}[|\xi_T^{0,\pm}|^2] < \infty, \quad (\text{III.5.21})$$

where the first inequality follows from Lemma III.5.2 and the Jordan decomposition (III.2.3); the last inequality follows from the fact that ξ^0 and ξ are indistinguishable and that $\xi \in \mathcal{A}_t^\nu$. Thus, $\mathbb{E}_\omega[|\xi_T^{0,u,\pm}|^2] < \infty$ for $\mathbb{P}$ -a.e. ω . Trivially, $\xi_u^{0,u} := 0$ and thus $\xi^{0,u} \in \mathcal{A}_u^{\nu_\omega}$ for $\mathbb{P}$ -a.e. ω . Similar properties follow analogously for $\eta^{0,u}$ and so also $\eta^{0,u} \in \mathcal{E}_u^{\nu_\omega}$ for $\mathbb{P}$ -a.e. ω , which completes the proof. $\square$

To conclude this section we introduce two more lemmas that will be used to obtain the dynamic programming principle. The next lemma is an enhanced version of Lemma III.4.4 that, applied to the particular reference probability systems ν_ω (as in (III.5.14)), yields a stronger result. Namely, given a reference probability system $\bar{\nu} \in \mathcal{V}_u$ and two predictable controls ξ, η admissible on the system $\bar{\nu}$, we can obtain two controls $\tilde{\xi}, \tilde{\eta}$ that are admissible on every system ν_ω for $\mathbb{P}$ -a.e. ω and that maintain the equality in law. In particular, notice that $\tilde{\xi}$ and $\tilde{\eta}$ do not depend on any fixed ω . This follows from the fact that W^u is a Brownian motion on $(\Omega, \mathcal{F}_\omega, \mathbb{P}_\omega)$ for $\mathbb{P}$ -a.e. ω (see Proposition III.5.4). In the rest of the discussion, when we fix a measure $\bar{\mathbb{P}}$ we denote by $\bar{\mathbb{E}}$ the corresponding expectation.

Lemma III.5.5. *Let $\bar{\nu} = (\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}}, \{\bar{\mathcal{F}}_s^u\}, \bar{W}) \in \mathcal{V}_u$ and $\xi \in \mathcal{A}_u^{\bar{\nu}}$, $\eta \in \mathcal{E}_u^{\bar{\nu}}$ be $\{\bar{\mathcal{F}}_s^{u,0}\}$ -predictable. For $\mathbb{P}$ -a.e. ω , let ν_ω be defined as in (III.5.14). Then, there exist two controls $\tilde{\xi}$ and $\tilde{\eta}$ that are $\{\mathcal{F}_s^{u,0}\}$ -adapted and such that*

$$\tilde{\xi} \in \mathcal{A}_u^{\nu_\omega}, \quad \tilde{\eta} \in \mathcal{E}_u^{\nu_\omega} \quad \text{and} \quad \mathcal{L}_{\mathbb{P}_\omega}(\tilde{\xi}^\pm, \tilde{\eta}, W^u) = \mathcal{L}_{\bar{\mathbb{P}}}(\xi^\pm, \eta, \bar{W}), \quad \mathbb{P}\text{-a.e. } \omega. \quad (\text{III.5.22})$$

Moreover, the set $\hat{\Omega} \subseteq \Omega$ where $\tilde{\xi}$ is left-continuous, of bounded variation and $\tilde{\xi}_u = 0$ (and similarly for $\tilde{\eta}$) belongs to $\mathcal{F}^0$ and $\mathbb{P}_\omega(\hat{\Omega}) = 1$ for $\mathbb{P}$ -a.e. ω .

Proof. Fix $\omega_0 \in \Omega$ such that, by Proposition III.5.4, $\nu_{\omega_0} := (\Omega, \mathcal{F}_{\omega_0}, \mathbb{P}_{\omega_0}, \{\mathcal{F}_s^{u,\omega_0}\}, W^u) \in \mathcal{V}_u$. By applying Lemma III.4.4 between the systems $\bar{\nu}, \nu_{\omega_0} \in \mathcal{V}_u$ with the $\{\bar{\mathcal{F}}_s^{u,0}\}$ -predictable controls $\xi \in \mathcal{A}_u^{\bar{\nu}}$, $\eta \in \mathcal{E}_u^{\bar{\nu}}$ we obtain that there exist $\tilde{\xi} \in \mathcal{A}_u^{\nu_{\omega_0}}$ and $\tilde{\eta} \in \mathcal{E}_u^{\nu_{\omega_0}}$ that are $\{\mathcal{F}_s^{u,0}\}$ -adapted and such that $\mathcal{L}_{\mathbb{P}_{\omega_0}}(\tilde{\xi}^\pm, \tilde{\eta}, W^u) = \mathcal{L}_{\bar{\mathbb{P}}}(\xi^\pm, \eta, \bar{W})$. By Proposition III.5.4, we immediately have that since $\tilde{\xi}$ and $\tilde{\eta}$ are $\{\mathcal{F}_s^{u,0}\}$ -adapted then they are also $\{\mathcal{F}_s^{u,\omega}\}$ -adapted for $\mathbb{P}$ -a.e. ω . We now need to prove that (III.5.22) holds for $\mathbb{P}$ -a.e. ω (not only for the fixed ω_0) and that $\hat{\Omega} \in \mathcal{F}^0$ with $\mathbb{P}_\omega(\hat{\Omega}) = 1$ for $\mathbb{P}$ -a.e. ω . This follows from the construction of $\tilde{\xi}$ and $\tilde{\eta}$ in Lemma III.4.4. Indeed, recall that $\tilde{\xi} := \tilde{\xi}^+ - \tilde{\xi}^-$ and

$$\tilde{\xi}_s^\pm := \begin{cases} \hat{\xi}_s^\pm & \text{if } s \in D \\ \lim_{\substack{t_n \in D \\ t_n \uparrow s}} \hat{\xi}_{t_n}^\pm & \text{if } s \in [t, T] \setminus D. \end{cases} \quad (\text{III.5.23})$$

where $D := \{t_n\}_n \subseteq [u, T]$ is dense and countable and $\hat{\xi}^\pm$ are $\{\mathcal{F}_s^{u,0}\}$ -predictable. Recall also that we define $\tilde{\Omega}_m^{k,n} := \{\hat{\xi}_{t_n}^\pm \leq \hat{\xi}_{t_{n+k}}^\pm, \hat{\xi}_{t_m}^\pm = \hat{\xi}_{t_{m-}}^\pm\}$, $\hat{\Omega} := \cap_{k,n,m} \tilde{\Omega}_m^{k,n}$ such that $\mathbb{P}_{\omega_0}(\hat{\Omega}) = 1$ and $\tilde{\xi}$ is left-continuous, of bounded variation and $\tilde{\xi}_u = 0$ on $\hat{\Omega}$. Thus, since $\hat{\xi}^\pm$ are $\{\mathcal{F}_s^{u,0}\}$ -adapted and $\mathcal{F}_T^{u,0} \subseteq \mathcal{F}_T^{t,0} \subseteq \mathcal{F}^0$ (by definition of reference probability system), we obtain that $\tilde{\Omega}_m^{k,n} \in \mathcal{F}_T^{u,0} \subseteq \mathcal{F}^0$ and so also $\hat{\Omega} \in \mathcal{F}^0$. Furthermore, since W^u is a Brownian motion on $(\Omega, \mathcal{F}_\omega, \mathbb{P}_\omega)$ for $\mathbb{P}$ -a.e. ω (see Proposition III.5.4), it is clear from (III.4.18) (with $\bar{\mathbb{P}}$ substituted by $\mathbb{P}_\omega$ and $\bar{W}$ substituted by W^u) that $\mathbb{P}_\omega(\tilde{\Omega}_m^{k,n}) = 1$ for every $k, n, m \in \mathbb{N}$ and for $\mathbb{P}$ -a.e. ω and thus that $\mathbb{P}_\omega(\hat{\Omega}) = 1$ for

$\mathbb{P}$ -a.e. ω . Hence, we obtain that $\tilde{\xi}$ is $\mathbb{P}_\omega$ -a.s. left-continuous, of bounded variation and $\tilde{\xi}_u = 0$, $\mathbb{P}_\omega$ -a.s. for $\mathbb{P}$ -a.e. ω . In the same manner we also have that $\tilde{\eta} \in \mathcal{E}_u^{\nu_\omega}$ for $\mathbb{P}$ -a.e. ω . Again, since W^u is a Brownian motion on $(\Omega, \mathcal{F}_\omega, \mathbb{P}_\omega)$ for $\mathbb{P}$ -a.e. ω , by an argument as in (III.4.20) and then (III.4.22), we obtain that $\mathcal{L}_{\mathbb{P}_\omega}(\tilde{\xi}^\pm, \tilde{\eta}, W^u) = \mathcal{L}_{\mathbb{P}}(\xi^\pm, \eta, \bar{W})$ for $\mathbb{P}$ -a.e. ω . This also implies that

$$\mathbb{E}_\omega[|\tilde{\xi}_T^\pm|^2] = \mathbb{E}[|\xi_T^\pm|^2] < \infty, \quad \mathbb{P}\text{-a.e. } \omega, \quad (\text{III.5.24})$$

and so $\tilde{\xi} \in \mathcal{A}_u^{\nu_\omega}$ for $\mathbb{P}$ -a.e. ω , which concludes the proof. $\square$

In the last lemma, we show that the stochastic integral with respect to the underlying probability $\mathbb{P}$ is indistinguishable from the stochastic integral with respect to the regular conditional probability $\mathbb{P}_\omega$ for $\mathbb{P}$ -a.e. ω . Here, for a square integrable process H , we denote by

$$\int_{[u,s]}^{\mathbb{P}} H_r dW_r^u \quad (\text{III.5.25})$$

the stochastic integral of H against W^u when constructed with respect to the probability $\mathbb{P}$ and, similarly, by

$$\int_{[u,s]}^{\mathbb{P}_\omega} H_r dW_r^u \quad (\text{III.5.26})$$

the stochastic integral of H against W^u when constructed with respect to the regular conditional probability $\mathbb{P}_\omega$.

Lemma III.5.6. *Let H be an $\{\mathcal{F}_s^t\}$ -adapted process such that*

$$\int_u^s H_r^2 dr < \infty, \quad s \in [u, T], \quad \mathbb{P}\text{-a.s.} \quad (\text{III.5.27})$$

Then,

$$\mathbb{P}_\omega \left(\int_{[u,s]}^{\mathbb{P}} H_r dW_r^u = \int_{[u,s]}^{\mathbb{P}_\omega} H_r dW_r^u, \forall s \in [u, T] \right) = 1, \quad \mathbb{P}\text{-a.e. } \omega. \quad (\text{III.5.28})$$

Proof. This proof follows the same arguments as in [25, Lemma 11].

Notice that since the stochastic integrals in (III.5.28) are continuous, it is sufficient to prove that they are a modification of each other. Namely we want to prove that, for every $s \in [u, T]$, we have

$$\mathbb{P}_\omega \left(\int_{[u,s]}^{\mathbb{P}} H_r dW_r^u = \int_{[u,s]}^{\mathbb{P}_\omega} H_r dW_r^u \right) = 1, \quad \mathbb{P}\text{-a.e. } \omega, \quad (\text{III.5.29})$$

By a standard localisation procedure, we can assume that

$$\mathbb{E} \left[\int_u^s H_r^2 dr \right] < \infty, \quad (\text{III.5.30})$$

and so to prove our desired result (III.5.29) we want to show that

$$\mathbb{E}_\omega \left[\left| \int_{[u,s]}^{\mathbb{P}} H_r dW_r^u - \int_{[u,s]}^{\mathbb{P}_\omega} H_r dW_r^u \right|^2 \right] = 0, \quad \mathbb{P}\text{-a.e. } \omega. \quad (\text{III.5.31})$$

Notice that in (III.5.31) since $|\int_{[u,s]}^{\mathbb{P}} H_r dW_r^u|^2 \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ then, by Lemma III.5.2, we also have $|\int_{[u,s]}^{\mathbb{P}_\omega} H_r dW_r^u|^2 \in L^1(\Omega, \mathcal{F}_\omega, \mathbb{P}_\omega)$ for $\mathbb{P}$ -a.e. ω .

Let $(H^{(n)})_{n \in \mathbb{N}}$ be a sequence of approximating simple processes, i.e.,

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\int_u^s |H_r - H_r^{(n)}|^2 dr \right] = 0. \quad (\text{III.5.32})$$

Then, by Lemma III.5.2, we also have that

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E} \left[\mathbb{E}_\omega \left[\int_u^s |H_r - H_r^{(n)}|^2 dr \right] \right] &= \lim_{n \rightarrow \infty} \mathbb{E} \left[\mathbb{E} \left[\int_u^s |H_r - H_r^{(n)}|^2 dr \middle| \mathcal{F}_u^t \right] (\omega) \right] \\ &= \lim_{n \rightarrow \infty} \mathbb{E} \left[\int_u^s |H_r - H_r^{(n)}|^2 dr \right] \\ &= 0. \end{aligned} \quad (\text{III.5.33})$$

Therefore, by Markov's inequality, we obtain that

$$\mathbb{E}_\omega \left[\int_u^s |H_r - H_r^{(n)}|^2 dr \right] \xrightarrow{\mathbb{P}} 0, \quad (\text{III.5.34})$$

where $\xrightarrow{\mathbb{P}}$ denotes the limit in probability. Thus, up to a subsequence which we still denote by $(H^{(n)})_{n \in \mathbb{N}}$, we have that

$$\lim_{n \rightarrow \infty} \mathbb{E}_\omega \left[\int_u^s |H_r - H_r^{(n)}|^2 dr \right] = 0, \quad \mathbb{P}\text{-a.e. } \omega. \quad (\text{III.5.35})$$

Moreover, we have that

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E} \left[\mathbb{E}_\omega \left[\left| \int_{[u,s]}^\mathbb{P} (H_r^{(n)} - H_r) dW_r^u \right|^2 \right] \right] &= \lim_{n \rightarrow \infty} \mathbb{E} \left[\mathbb{E} \left[\left| \int_{[u,s]}^\mathbb{P} (H_r^{(n)} - H_r) dW_r^u \right|^2 \middle| \mathcal{F}_u^t \right] (\omega) \right] \\ &= \lim_{n \rightarrow \infty} \mathbb{E} \left[\left| \int_{[u,s]}^\mathbb{P} (H_r^{(n)} - H_r) dW_r^u \right|^2 \right] \\ &= \lim_{n \rightarrow \infty} \mathbb{E} \left[\int_u^s |H_r^{(n)} - H_r|^2 dr \right] \\ &= 0, \end{aligned} \quad (\text{III.5.36})$$

where in the first equality we have used Lemma III.5.2, in the third equality we have used Ito's isometry with respect to the probability $\mathbb{P}$, and in the fourth equality we have used (III.5.32). Thus, again by using Markov's inequality and up to a new subsequence which we still denote by $(H^{(n)})_{n \in \mathbb{N}}$, we obtain that

$$\lim_{n \rightarrow \infty} \mathbb{E}_\omega \left[\left| \int_{[u,s]}^\mathbb{P} (H_r^{(n)} - H_r) dW_r^u \right|^2 \right] = 0, \quad \mathbb{P}\text{-a.e. } \omega. \quad (\text{III.5.37})$$

Now notice that

$$\begin{aligned} \int_{[u,s]}^\mathbb{P} H_r dW_r^u - \int_{[u,s]}^{\mathbb{P}_\omega} H_r dW_r^u &= \int_{[u,s]}^\mathbb{P} H_r^{(n)} dW_r^u - \int_{[u,s]}^{\mathbb{P}_\omega} H_r^{(n)} dW_r^u \\ &\quad + \int_{[u,s]}^{\mathbb{P}_\omega} (H_r^{(n)} - H_r) dW_r^u \\ &\quad - \int_{[u,s]}^\mathbb{P} (H_r^{(n)} - H_r) dW_r^u, \end{aligned} \quad (\text{III.5.38})$$

where the difference of the first two integrals on the right-hand side is equal to zero as stochastic integrals of simple processes are defined pathwise. Hence, for $\mathbb{P}$ -a.e. ω ,

$$\begin{aligned} \mathbb{E}_\omega \left[\left| \int_{[u,s]}^\mathbb{P} H_r dW_r^u - \int_{[u,s]}^{\mathbb{P}_\omega} H_r dW_r^u \right|^2 \right] &\leq 2 \mathbb{E}_\omega \left[\left| \int_{[u,s]}^{\mathbb{P}_\omega} (H_r^{(n)} - H_r) dW_r^u \right|^2 \right] \\ &\quad + 2 \mathbb{E}_\omega \left[\left| \int_{[u,s]}^\mathbb{P} (H_r^{(n)} - H_r) dW_r^u \right|^2 \right] \end{aligned} \quad (\text{III.5.39})$$

By Ito's isometry with respect to the regular conditional probability $\mathbb{P}_\omega$ and (III.5.35), we have that

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E}_\omega \left[\left| \int_{[u,s]}^{\mathbb{P}_\omega} (H_r^{(n)} - H_r) dW_r^u \right|^2 \right] &= \lim_{n \rightarrow \infty} \mathbb{E}_\omega \left[\int_u^s |H_r^{(n)} - H_r|^2 dr \right] \\ &= 0, \quad \mathbb{P}\text{-a.e. } \omega. \end{aligned} \quad (\text{III.5.40})$$

Thus, by letting $n \uparrow \infty$ in (III.5.39) and using also (III.5.37), we obtain the desired result (III.5.31). $\square$

III.6 Value process

In this section, for a fixed initial control $\xi \in \mathcal{A}_t^\nu$, we introduce the concept of value process V^ξ for our problem. Then, we show that, for every $u \in [t, T]$, we have

$$V_u^\xi = v(u, X_u^\xi), \quad \mathbb{P}\text{-a.s.}, \quad (\text{III.6.1})$$

i.e., that the value process is a modification of the value function composed with the state process (see Proposition III.6.1). The duality between the strong and the weak formulations (III.3.3)-(III.3.4) (together with Proposition III.4.5) and the regular conditional probabilities play a key role for the proof of (III.6.1). By explicitly connecting the value process to the value function, this result leads to a smoother proof of the dynamic programming principle which will be presented in the next section.

Let $u \in [t, T]$ and $\xi \in \mathcal{A}_t^\nu$ and recall the definition of $\mathcal{A}_{u,\xi}^\nu$ in Section III.2. The value process is defined as

$$V_u^\xi := \operatorname{ess\,sup}_{(\tau, \zeta) \in \mathcal{T}_u^\nu \times \mathcal{A}_{u,\xi}^\nu} \mathbb{E}[Y(u, \tau, \zeta, X^\zeta, \rho^\zeta) \mathbf{1}_{\{\rho^\zeta \geq u\}} | \mathcal{F}_u^t], \quad (\text{III.6.2})$$

where, as usual, $X^\zeta := X^{t,x;\zeta,\nu}$ and $\rho^\zeta := \rho^{t,x;\zeta,\nu}$, and

$$\begin{aligned} Y(u, \tau, \zeta, X, \rho) &:= \int_u^{\tau \wedge \rho} f(s, X_s) ds - \int_{[u, \tau \wedge \rho)} \langle c_+(s), d\zeta_s^+ \rangle \\ &\quad - \int_{[u, \tau \wedge \rho)} \langle c_-(s), d\zeta_s^- \rangle + g_1(\rho, X_\rho) \mathbf{1}_{\{\rho \leq \tau\}} + g_2(\tau, X_\tau) \mathbf{1}_{\{\rho > \tau\}}. \end{aligned} \quad (\text{III.6.3})$$

When $\zeta \in \mathcal{A}_{u,\xi}^\nu$ is fixed and no confusion may arise, we will simply denote $Y(u, \tau, \zeta, X^\zeta, \rho^\zeta)$ by $Y(u, \tau, \zeta)$. Notice that

$$J_{t,x}^\nu(\tau, \zeta) = \mathbb{E}[Y(t, \tau, \zeta, X^\zeta, \rho^\zeta)]. \quad (\text{III.6.4})$$

Also, notice that on $\{\rho^\xi \geq u\}$ we have $\rho^\zeta \geq u$ for $\zeta \in \mathcal{A}_{u,\xi}^\nu$ and so

$$\rho^\zeta := \inf\{s \geq t : X_s^\zeta \notin \mathcal{O}\} = \inf\{s \geq u : X_s^\zeta \notin \mathcal{O}\}, \quad \mathbb{P}\text{-a.s. on } \{\rho^\xi \geq u\}. \quad (\text{III.6.5})$$

In the same way as for the equivalence between $\tau \in \mathcal{T}_t^\nu$ and $\eta \in \mathcal{E}_t^\nu$ in (III.4.2), for the value process the optimisation over $\tau \in \mathcal{T}_u^\nu$ is equivalent to the optimisation over $\eta \in \mathcal{E}_u^\nu$. Hence,

$$V_u^\xi = \operatorname{ess\,sup}_{(\eta,\zeta) \in \mathcal{E}_u^\nu \times \mathcal{A}_{u,\xi}^\nu} \mathbb{E}[Z(u, \eta, \zeta, X^\zeta, \rho^\zeta) \mathbf{1}_{\{\rho^\xi \geq u\}} | \mathcal{F}_u^t], \quad (\text{III.6.6})$$

where

$$\begin{aligned} Z(u, \eta, \zeta, X, \rho) &:= \int_u^\rho (1 - \eta_s) f(s, X_s) ds \\ &\quad - \int_{[u,\rho)} (1 - \eta_{s+}) \langle c_+(s), d\zeta_s^+ \rangle - \int_{[u,\rho)} (1 - \eta_{s+}) \langle c_-(s), d\zeta_s^- \rangle \\ &\quad + g_1(\rho, X_\rho)(1 - \eta_\rho) + \int_{[u,\rho)} g_2(s, X_s) d\eta_s. \end{aligned} \quad (\text{III.6.7})$$

Notice that

$$I_{t,x}^\nu(\eta, \zeta) = \mathbb{E}[Z(t, \eta, \zeta, X^\zeta, \rho^\zeta)], \quad (\text{III.6.8})$$

and that, in the same way as in (III.4.2), if $\eta \in \mathcal{E}_u^\nu$ and $\tau^\eta := \inf\{s \geq u : \eta_s = 1\} \wedge T$ then

$$Z(u, \eta, \zeta, X, \rho) = Y(u, \tau^\eta, \zeta, X, \rho). \quad (\text{III.6.9})$$

Proposition III.6.1. *Let $u \in [t, T]$ and $\xi \in \mathcal{A}_t^\nu$. Assume that for every $\bar{v} \in \mathcal{V}_u$, the function $y \mapsto J_{u,y}^\nu(\bar{\tau}, \bar{\xi})$ is continuous uniformly in $(\bar{\tau}, \bar{\xi}) \in \mathcal{T}_u^\nu \times \mathcal{A}_u^\nu$. Then, on $\{\rho^\xi \geq u\}$, we have*

$$V_u^\xi = v(u, X_u^\xi), \quad \mathbb{P}\text{-a.s.} \quad (\text{III.6.10})$$

Remark III.6.2. *In the following proof we will often use the expression $\mathbb{P}$ -a.e. ω to indicate - as it is well known - that a certain property Γ_i holds on a set $\Omega_i \subseteq \Omega$ with $\mathbb{P}(\Omega_i) = 1$. Notice that the sets $\Omega_i \subseteq \Omega$ will, in general, depend on the property Γ_i and so they will, in general, be different. However, we will always consider a finite number $n \in \mathbb{N}$ of properties $\Gamma_1, \Gamma_2, \dots, \Gamma_n$ and so we will be able to generate a set $\Omega' := \cap_{i=1}^n \Omega_i \subseteq \Omega$ with $\mathbb{P}(\Omega') = 1$ where all the properties $\Gamma_1, \Gamma_2, \dots, \Gamma_n$ hold.*

Proof. Recall that X^ξ denotes $X^{t,x;\xi,\nu}$. For $\mathbb{P}$ -a.e. ω , let $\nu_\omega = (\Omega, \mathcal{F}_\omega, \mathbb{P}_\omega, \{\mathcal{F}_s^{u,\omega}\}, W^u) \in \mathcal{V}_u$ be as in (III.5.14). We prove the equality in (III.6.10) by proving the two inequalities.

Part 1. Here we show that $V_u^\xi \leq v(u, X_u^\xi)$, $\mathbb{P}$ -a.s. on $\{\rho^\xi \geq u\}$. Let $(\eta, \zeta) \in \mathcal{E}_u^\nu \times \mathcal{A}_{u,\xi}^\nu$ and set $X^\zeta := X^{t,x;\zeta,\nu}$. By Corollary III.4.2, let ζ^0 be the $\{\mathcal{F}_s^{t,0}\}$ -adapted control which is indistinguishable from ζ . Thus, also $X^{\zeta^0} := X^{t,x;\zeta^0,\nu}$ and X^ζ are indistinguishable. By Lemma III.4.1, let X^0 be the $\{\mathcal{F}_s^{t,0}\}$ -adapted process which is indistinguishable from X^{ζ^0} (and so also indistinguishable from X^ζ). Thus, we also have $\rho^\zeta = \rho^0 := \inf\{s \geq u : X_s^0 \notin \mathcal{O}\}$, $\mathbb{P}$ -a.s. on $\{\rho^\xi \geq u\}$. Moreover, let

$\zeta_s^{0,u} := \zeta_s^0 - \zeta_u^0$ for $s \in [u, T]$. Then, $\mathbb{P}$ -a.s., for every $s \in [u, T]$ we have

$$\begin{aligned} X_s^0 &= X_s^{\zeta^0} \\ &= X_u^{\zeta^0} + \int_u^s \mu(r, X_r^{\zeta^0}) dr + \int_u^s \sigma(r, X_r^{\zeta^0}) dW_r + \zeta_s^0 - \zeta_u^0 \\ &= X_u^0 + \int_u^s \mu(r, X_r^0) dr + \int_u^s \sigma(r, X_r^0) dW_r^u + \zeta_s^{0,u}. \end{aligned} \quad (\text{III.6.11})$$

Since X_u^0 is $\mathcal{F}_u^{t,0}$ -measurable then, by property (III.5.2), $X_u^0 = X_u^0(\omega)$, $\mathbb{P}$ -a.s. for $\mathbb{P}$ -a.e. ω . Thus, by Lemma III.5.1, we obtain that, $\mathbb{P}$ -a.s. for $\mathbb{P}$ -a.e. ω , for every $s \in [u, T]$

$$X_s^0 = X_u^0(\omega) + \int_u^s \mu(r, X_r^0) dr + \int_u^s \sigma(r, X_r^0) dW_r^u + \zeta_s^{0,u}. \quad (\text{III.6.12})$$

Also notice that the stochastic integral in (III.6.12) is constructed with respect to the probability $\mathbb{P}$, but by Lemma III.5.6 it is $\mathbb{P}_\omega$ -indistinguishable, for $\mathbb{P}$ -a.e. ω , from the stochastic integral constructed with respect to the regular conditional probability $\mathbb{P}_\omega$. Therefore, the process X^0 satisfies the same SDE as the process $X^\omega := X^{u, X_u^0(\omega); \zeta^{0,u}, \nu_\omega}$ on ν_ω for $\mathbb{P}$ -a.e. ω . Furthermore, X^0 is $\{\mathcal{F}_s^{t,0}\}$ -adapted and so $\{\mathcal{F}_s^{u,\omega}\}$ -adapted for $\mathbb{P}$ -a.e. ω (see Lemma III.5.3). Thus, by uniqueness (up to indistinguishability) of the SDE (III.3.1), we obtain that

$$\mathbb{P}_\omega(X_s^0 = X_s^\omega, \forall s \in [u, T]) = 1, \quad \mathbb{P}\text{-a.e. } \omega. \quad (\text{III.6.13})$$

As a consequence, for $\mathbb{P}$ -a.e. ω , we also have that

$$\rho^0 = \rho^\omega := \inf\{s \geq u : X_s^\omega \notin \mathcal{O}\}, \quad \mathbb{P}\text{-a.s.} \quad (\text{III.6.14})$$

By Corollary III.4.2, let $\eta^0 \in \mathcal{E}_u^\nu$ be the $\{\mathcal{F}_s^{t,0}\}$ -adapted control which is indistinguishable from $\eta \in \mathcal{E}_u^\nu$. By Proposition III.5.4, we have that $\eta^0 = \eta^{0,u} \in \mathcal{E}_u^{\nu_\omega}$ and $\zeta^{0,u} \in \mathcal{A}_u^{\nu_\omega}$ for $\mathbb{P}$ -a.e. ω . Hence, for $\mathbb{P}$ -a.e. ω , we obtain

$$\begin{aligned} \mathbb{E}[Z(u, \eta, \zeta, X^\zeta, \rho^\zeta) \mathbf{1}_{\{\rho^\zeta \geq u\}} | \mathcal{F}_u^t](\omega) &= \mathbb{E}[Z(u, \eta^0, \zeta^{0,u}, X^0, \rho^0) \mathbf{1}_{\{\rho^0 \geq u\}} | \mathcal{F}_u^t](\omega) \\ &= \mathbb{E}[Z(u, \eta^0, \zeta^{0,u}, X^0, \rho^0) | \mathcal{F}_u^t](\omega) \mathbf{1}_{\{\rho^0 \geq u\}}(\omega) \\ &= \mathbb{E}_\omega[Z(u, \eta^0, \zeta^{0,u}, X^\omega, \rho^\omega) \mathbf{1}_{\{\rho^\omega \geq u\}}](\omega) \\ &= I_{u, X_u^0(\omega)}^{\nu_\omega}(\eta^0, \zeta^{0,u}) \mathbf{1}_{\{\rho^\omega \geq u\}}(\omega) \\ &\leq v(u, X_u^0(\omega)) \mathbf{1}_{\{\rho^\omega \geq u\}}(\omega), \end{aligned} \quad (\text{III.6.15})$$

where the third equality follows from Lemma III.5.2 and equalities (III.6.13)-(III.6.14), and the inequality follows from the fact that $\nu_\omega \in \mathcal{V}_u$, $(\eta^0, \zeta^{0,u}) \in \mathcal{E}_u^{\nu_\omega} \times \mathcal{A}_u^{\nu_\omega}$ and Proposition III.4.5. Thus, for every $(\eta, \zeta) \in \mathcal{E}_u^\nu \times \mathcal{A}_{u,\xi}^\nu$, we have

$$\mathbb{E}[Z(u, \eta, \zeta, X^\zeta, \rho^\zeta) \mathbf{1}_{\{\rho^\zeta \geq u\}} | \mathcal{F}_u^t] \leq v(u, X_u^0), \quad \mathbb{P}\text{-a.s. on } \{\rho^\zeta \geq u\}. \quad (\text{III.6.16})$$

and by taking the essential supremum over $(\eta, \zeta) \in \mathcal{E}_u^\nu \times \mathcal{A}_{u,\xi}^\nu$, we obtain

$$V_u^\zeta \leq v(u, X_u^0) = v(u, X_u^\zeta) = v(u, X_u^\xi), \quad \mathbb{P}\text{-a.s. on } \{\rho^\zeta \geq u\}, \quad (\text{III.6.17})$$

and the first inequality is proved.

Part 2. Here we show that $V_u^\xi \geq v(u, X_u^\xi)$, $\mathbb{P}$ -a.s. on $\{\rho^\xi \geq u\}$. Let $\varepsilon > 0$ and $\bar{\nu} = (\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}}, \{\bar{\mathcal{F}}_s^u\}, \bar{W}) \in \mathcal{V}_u$. Since $y \mapsto J_{u,y}^{\bar{\nu}}(\bar{\tau}, \bar{\xi})$ is uniformly continuous on compact sets uniformly in $(\bar{\tau}, \bar{\xi}) \in \mathcal{T}_u^{\bar{\nu}} \times \mathcal{A}_u^{\bar{\nu}}$, we can choose $\delta_1 > 0$ such that if $y, z \in B(0, 1) := \{x \in \mathbb{R}^d : |x| \leq 1\}$ and $|y - z| < \delta_1$, then

$$|J_{u,y}^{\bar{\nu}}(\bar{\tau}, \bar{\xi}) - J_{u,z}^{\bar{\nu}}(\bar{\tau}, \bar{\xi})| + |v(u, y) - v(u, z)| < \varepsilon, \quad \forall (\bar{\tau}, \bar{\xi}) \in \mathcal{T}_u^{\bar{\nu}} \times \mathcal{A}_u^{\bar{\nu}}. \quad (\text{III.6.18})$$

Since $\mathbb{R}^d$ is separable, we can choose a partition $\{B_n^1\}_{n \in \mathbb{N}}$ of $B(0, 1)$ into countable disjoint Borel sets with diameter $\text{diam}(B_n^1) < \delta_1$ for every $n \in \mathbb{N}$, where $\text{diam}(B) := \sup\{|x - y| : x, y \in B\}$. Similarly, we can choose $\delta_2 > 0$ such that if $y, z \in B(0, 2)$ and $|y - z| < \delta_2$, then

$$|J_{u,y}^{\bar{\nu}}(\bar{\tau}, \bar{\xi}) - J_{u,z}^{\bar{\nu}}(\bar{\tau}, \bar{\xi})| + |v(u, y) - v(u, z)| < \varepsilon, \quad \forall (\bar{\tau}, \bar{\xi}) \in \mathcal{T}_u^{\bar{\nu}} \times \mathcal{A}_u^{\bar{\nu}}. \quad (\text{III.6.19})$$

Thus, we can also choose a partition $\{B_n^2\}_{n \in \mathbb{N}}$ of $B(0, 2) \setminus B(0, 1)$ into countable disjoint Borel sets with $\text{diam}(B_n^2) < \delta_2$ for every $n \in \mathbb{N}$. By iterating this argument, we can find a partition $\{B_n\}_{n \in \mathbb{N}}$ of $\mathbb{R}^d$ into countable disjoint Borel sets such that for every B_n and every $y, z \in B_n$, we obtain

$$|J_{u,y}^{\bar{\nu}}(\bar{\tau}, \bar{\xi}) - J_{u,z}^{\bar{\nu}}(\bar{\tau}, \bar{\xi})| + |v(u, y) - v(u, z)| < \varepsilon, \quad \forall (\bar{\tau}, \bar{\xi}) \in \mathcal{T}_u^{\bar{\nu}} \times \mathcal{A}_u^{\bar{\nu}}. \quad (\text{III.6.20})$$

Let $n, m \in \mathbb{N}$ and $x_n \in B_n$. By Proposition III.4.5, there exist $(\tau^{n,m}, \xi^{n,m}) \in \mathcal{T}_u^{\bar{\nu}} \times \mathcal{A}_u^{\bar{\nu}}$ such that

$$J_{u,x_n}^{\bar{\nu}}(\tau^{n,m}, \xi^{n,m}) \geq v(u, x_n) - 1/m. \quad (\text{III.6.21})$$

Since $m \in \mathbb{N}$ is fixed, in the following we simply denote $\tau^{n,m}$ and $\xi^{n,m}$ by τ^n and ξ^n , respectively. By the equivalence (III.4.2) between τ and η , for every $n \in \mathbb{N}$ there exists $\eta^n \in \mathcal{E}_u^{\bar{\nu}}$ such that

$$J_{u,y}^{\bar{\nu}}(\tau^n, \xi^n) = I_{u,y}^{\bar{\nu}}(\eta^n, \xi^n), \quad \forall y \in \mathbb{R}^d. \quad (\text{III.6.22})$$

The main idea for this second part of the proof is to build a pair of controls $(\chi, \zeta) \in \mathcal{E}_u^{\nu} \times \mathcal{A}_{u,\xi}^{\nu}$ which is δ -optimal with $\delta := 2\varepsilon + 1/m$. We proceed as follows.

Let $O_n := \{X_u^\xi \in B_n\} \in \mathcal{F}_u^t$ and notice that, since $\{B_n\}_{n \in \mathbb{N}}$ is a partition of $\mathbb{R}^d$, $\{O_n\}_{n \in \mathbb{N}}$ is a partition of Ω . We now want to construct, for every $k \in \mathbb{N}$, a specific pair of controls $(\chi^k, \zeta^k) \in \mathcal{E}_u^{\nu} \times \mathcal{A}_{u,\xi}^{\nu}$ that agrees with (η^n, ξ^n) on O_n for every $n \leq k$. We apply the following argument for every $n \in \mathbb{N}$. Let $\eta^{n,0} \in \mathcal{E}_u^{\bar{\nu}}$ and $\xi^{n,0} \in \mathcal{A}_u^{\bar{\nu}}$ be the $\{\bar{\mathcal{F}}_s^{u,0}\}$ -predictable controls from Corollary III.4.2, which are indistinguishable from η^n and ξ^n , respectively. By applying Lemma III.5.5 to $\eta^{n,0}$ and $\xi^{n,0}$, we obtain $\tilde{\eta}^n$ and $\tilde{\xi}^n$ that are $\{\mathcal{F}_s^{u,0}\}$ -adapted and such that $(\tilde{\eta}^n, \tilde{\xi}^n) \in \mathcal{E}_u^{\nu_\omega} \times \mathcal{A}_u^{\nu_\omega}$ and

$$\mathcal{L}_{\mathbb{P}_\omega}(\tilde{\eta}^n, \tilde{\xi}^{n,\pm}, W^u) = \mathcal{L}_{\bar{\mathbb{P}}}(\eta^{n,0}, \xi^{n,0,\pm}, \bar{W}), \quad \mathbb{P}\text{-a.e. } \omega. \quad (\text{III.6.23})$$

Since the pair $(\eta^{n,0}, \xi^{n,0})$ is indistinguishable from the pair (η^n, ξ^n) , we also obtain that

$$\mathcal{L}_{\mathbb{P}_\omega}(\tilde{\eta}^n, \tilde{\xi}^{n,\pm}, W^u) = \mathcal{L}_{\bar{\mathbb{P}}}(\eta^n, \xi^{n,\pm}, \bar{W}), \quad \mathbb{P}\text{-a.e. } \omega, \quad (\text{III.6.24})$$

and thus, by (III.4.24), for every $y \in \mathbb{R}^d$ we have

$$\mathcal{L}_{\mathbb{P}_\omega}(\tilde{\eta}^n, \tilde{\xi}^{n,\pm}, X^{u,y;\tilde{\xi}^n,\nu_\omega}, \rho^{u,y;\tilde{\xi}^n,\nu_\omega}) = \mathcal{L}_{\bar{\mathbb{P}}}(\eta^n, \xi^{n,\pm}, X^{u,y;\xi^n,\bar{\nu}}, \rho^{u,y;\xi^n,\bar{\nu}}), \quad \mathbb{P}\text{-a.e. } \omega. \quad (\text{III.6.25})$$

Moreover, still by Lemma III.5.5, the set $\tilde{\Omega}^n$ where $\tilde{\xi}^n$ is left-continuous, of bounded variation and $\tilde{\xi}_u^n = 0$ belongs to $\mathcal{F}^0$ and $\mathbb{P}_\omega(\tilde{\Omega}^n) = 1$ for $\mathbb{P}$ -a.e. ω . Therefore, by Lemma III.5.1, we also have $\mathbb{P}(\tilde{\Omega}^n) = 1$. Hence, $\tilde{\Omega} := \cap_n \tilde{\Omega}^n \in \mathcal{F}^0$, $\mathbb{P}(\tilde{\Omega}) = 1$ and on $\tilde{\Omega}$ every $\tilde{\xi}^n$ is left-continuous, of bounded variation and such that $\tilde{\xi}_u^n = 0$. Analogously, the same argument holds for $\tilde{\eta}^n$. Moreover, let $U_k := \cup_{n=1}^k O_n$ for $k \in \mathbb{N}$ and notice that $U_k \uparrow \Omega$ as $k \rightarrow \infty$. Thus, for every $k \in \mathbb{N}$, by defining

$$\chi_s^k := \left[\sum_{n=1}^k \tilde{\eta}_s^n \mathbf{1}_{O_n} \right] \mathbf{1}_{[u, T]}(s) + \mathbf{1}_{U_k^c} \mathbf{1}_{(u, T]}(s), \quad s \in [t, T], \quad (\text{III.6.26})$$

$$\zeta_s^{k, \pm} := \xi_s^\pm \mathbf{1}_{[t, u)}(s) + \left[\sum_{n=1}^k (\tilde{\xi}_s^{n, \pm} + \xi_u^\pm) \mathbf{1}_{O_n} + \xi_s^\pm \mathbf{1}_{U_k^c} \right] \mathbf{1}_{[u, T]}(s), \quad s \in [t, T], \quad (\text{III.6.27})$$

$$\zeta^k := \zeta^{k, +} - \zeta^{k, -}, \quad (\text{III.6.28})$$

we obtain that $(\chi^k, \zeta^k) \in \mathcal{E}_u^\nu \times \mathcal{A}_{u, \xi}^\nu$. In particular, to show that $\mathbb{E}[|\zeta_T^{k, \pm}|^2] < \infty$, notice that

$$\begin{aligned} \mathbb{E}[|\tilde{\zeta}_T^{n, \pm}|^2] &= \mathbb{E}[\mathbb{E}[|\tilde{\zeta}_T^{n, \pm}|^2 | \mathcal{F}_u^t](\omega)] \\ &= \mathbb{E}[\mathbb{E}_\omega[|\tilde{\zeta}_T^{n, \pm}|^2]] \\ &= \mathbb{E}[\mathbb{E}[\xi_T^{n, \pm}|^2]] \\ &= \bar{\mathbb{E}}[|\xi_T^{n, \pm}|^2] < \infty, \end{aligned} \quad (\text{III.6.29})$$

where the second equality follows from Lemma III.5.2, the third equality follows from (III.6.24) and the inequality follows from the fact that $\xi^n \in \mathcal{A}_u^\nu$. Also, notice that $\mathbb{P}$ -a.s. left-continuity at time u is guaranteed by the fact that $\zeta_u^k = \xi_u$ and $\chi_u^k = 0$, $\mathbb{P}$ -a.s.

We now fix $n, k \in \mathbb{N}$ with $n \leq k$. The next step is to show that

$$\mathbb{P}_\omega(\zeta_s^k = \tilde{\xi}_s^n + \xi_u, \forall s \in [u, T]) = 1, \quad \mathbb{P}\text{-a.e. } \omega \in O_n. \quad (\text{III.6.30})$$

Notice that $O_n \subseteq \{\zeta_s^k = \tilde{\xi}_s^n + \xi_u, \forall s \in [u, T]\}$. Thus, if we prove that $\mathbb{P}_\omega(O_n) = 1$ for $\mathbb{P}$ -a.e. $\omega \in O_n$ then (III.6.30) follows. This is easy to show as, by Lemma III.5.2, we have

$$\mathbb{P}_\omega(O_n) = \mathbb{E}_\omega[\mathbf{1}_{O_n}] = \mathbb{E}[\mathbf{1}_{O_n} | \mathcal{F}_u^t](\omega) = \mathbf{1}_{O_n}(\omega) = 1, \quad \mathbb{P}\text{-a.e. } \omega \in O_n. \quad (\text{III.6.31})$$

Now, let $X^k := X^{t, x; \zeta^k, \nu}$ and, by Lemma III.4.1, let $X^{k, 0}$ be the $\{\mathcal{F}_s^{t, 0}\}$ -predictable process that is indistinguishable from X^k . Moreover, for $\mathbb{P}$ -a.e. ω such that $\nu_\omega \in \mathcal{V}_u$ (as in (III.5.14)), let $X^{\omega, n} := X^{u, X_u^{k, 0}(\omega); \xi^n, \nu_\omega}$. Next, we want to prove that

$$\mathbb{P}_\omega(X_s^{k, 0} = X_s^{\omega, n}, \forall s \in [u, T]) = 1, \quad \mathbb{P}\text{-a.e. } \omega \in O_n. \quad (\text{III.6.32})$$

We have, $\mathbb{P}$ -a.s. for every $s \in [u, T]$,

$$\begin{aligned} X_s^{k, 0} &= X_s^k \\ &= X_u^k + \int_u^s \mu(r, X_r^k) dr + \int_u^s \sigma(r, X_r^k) dW_r + \zeta_s^k - \zeta_u^k \\ &= X_u^{k, 0} + \int_u^s \mu(r, X_r^{k, 0}) dr + \int_u^s \sigma(r, X_r^{k, 0}) dW_r^u + \zeta_s^k - \zeta_u^k. \end{aligned} \quad (\text{III.6.33})$$

Since $X_u^{k, 0}$ is $\mathcal{F}_u^{t, 0}$ -measurable, by property (III.5.2), we have that $X_u^{k, 0} = X_u^{k, 0}(\omega)$, $\mathbb{P}_\omega$ -a.s. for

$\mathbb{P}$ -a.e. ω . Thus, by Lemma III.5.1 and (III.6.30), we obtain that $\mathbb{P}_\omega$ -a.s. for $\mathbb{P}$ -a.e. $\omega \in O_n$ and for every $s \in [u, T]$

$$X_s^{k,0} = X_u^{k,0}(\omega) + \int_u^s \mu(r, X_r^{k,0})dr + \int_u^s \sigma(r, X_r^{k,0})dW_r^u + \tilde{\xi}_s^n, \quad (\text{III.6.34})$$

where, by means of Lemma III.5.6, the stochastic integral is intended as constructed with respect to the probability $\mathbb{P}_\omega$. Therefore, for $\mathbb{P}$ -a.e. $\omega \in O_n$, the process $X^{k,0}$ satisfies the same SDE as the process $X^{\omega,n}$ on ν_ω . Furthermore, $X^{k,0}$ is $\{\mathcal{F}_s^{t,0}\}$ -adapted and so $\{\mathcal{F}_s^{u,\omega}\}$ -adapted for $\mathbb{P}$ -a.e. ω (see Lemma III.5.3). Thus, by uniqueness (up to indistinguishability) of the SDE (III.3.1), equation (III.6.32) holds true.

Accordingly, we denote ρ^{ζ^k} by ρ^k . Notice that $\rho^k = \rho^{k,0} := \inf\{s \geq u : X_u^{k,0} \notin \mathcal{O}\}$ $\mathbb{P}$ -a.s. on $\{\rho^\xi \geq u\}$ and that (III.6.32) also implies that, for $\mathbb{P}$ -a.e. $\omega \in O_n$,

$$\rho^{\omega,n} := \inf\{s \geq u : X_s^{\omega,n} \notin \mathcal{O}\} = \rho^{k,0}, \quad \mathbb{P}_\omega\text{-a.s.} \quad (\text{III.6.35})$$

Moreover, from the definition of $X^{k,0}$ and from (III.6.25), we obtain

$$I_{u, X_u^\xi(\omega)}^{\bar{\nu}}(\eta^n, \xi^n) = I_{u, X_u^{k,0}(\omega)}^{\bar{\nu}}(\eta^n, \xi^n) = I_{u, X_u^{k,0}(\omega)}^{\nu_\omega}(\tilde{\eta}^n, \tilde{\xi}^n), \quad \mathbb{P}\text{-a.e. } \omega. \quad (\text{III.6.36})$$

Since $(\chi^k, \zeta^k) \in \mathcal{E}_u^\nu \times \mathcal{A}_{u,\xi}^\nu$, for $\mathbb{P}$ -a.e. ω , we have that

$$\begin{aligned} V_u^\xi(\omega) &\geq \mathbb{E}[Z(u, \chi^k, \zeta^k, X^k, \rho^k) \mathbf{1}_{\{\rho^\xi \geq u\}} | \mathcal{F}_u^t](\omega) \mathbf{1}_{U_k}(\omega) \\ &\quad + \mathbb{E}[Z(u, \chi^k, \zeta^k, X^k, \rho^k) \mathbf{1}_{\{\rho^\xi \geq u\}} | \mathcal{F}_u^t](\omega) \mathbf{1}_{U_k^c}(\omega). \end{aligned} \quad (\text{III.6.37})$$

For the first term on the right-hand side of (III.6.37), for $\mathbb{P}$ -a.e. ω , we obtain

$$\begin{aligned} &\mathbb{E}[Z(u, \chi^k, \zeta^k, X^k, \rho^k) \mathbf{1}_{\{\rho^\xi \geq u\}} | \mathcal{F}_u^t](\omega) \mathbf{1}_{U_k}(\omega) \\ &= \sum_{n=1}^k \mathbb{E}[Z(u, \chi^k, \zeta^k, X^{k,0}, \rho^{k,0}) \mathbf{1}_{O_n} | \mathcal{F}_u^t](\omega) \mathbf{1}_{\{\rho^\xi \geq u\}}(\omega) \\ &= \sum_{n=1}^k \mathbb{E}[Z(u, \tilde{\eta}^n, \tilde{\xi}^n, X^{k,0}, \rho^{k,0}) | \mathcal{F}_u^t](\omega) \mathbf{1}_{O_n \cap \{\rho^\xi \geq u\}}(\omega) \\ &= \sum_{n=1}^k \mathbb{E}_\omega[Z(u, \tilde{\eta}^n, \tilde{\xi}^n, X^{\omega,n}, \rho^{\omega,n})] \mathbf{1}_{O_n \cap \{\rho^\xi \geq u\}}(\omega) \\ &= \sum_{n=1}^k I_{u, X_u^{k,0}(\omega)}^{\nu_\omega}(\tilde{\eta}^n, \tilde{\xi}^n) \mathbf{1}_{O_n \cap \{\rho^\xi \geq u\}}(\omega) \\ &= \sum_{n=1}^k I_{u, X_u^\xi(\omega)}^{\bar{\nu}}(\eta^n, \xi^n) \mathbf{1}_{O_n \cap \{\rho^\xi \geq u\}}(\omega), \end{aligned} \quad (\text{III.6.38})$$

where the equality on the third line follows from the construction of χ^k and ζ^k (see (III.6.26)-(III.6.28)); the equality on the fourth line follows from Lemma III.5.2 and equations (III.6.32) and (III.6.35); and the equality on the last line follows from (III.6.36). Thus, for $\mathbb{P}$ -a.e. $\omega \in \{\rho^\xi \geq u\}$,

equation (III.6.37) reads as

$$\begin{aligned} V_u^\xi(\omega) &\geq \sum_{n=1}^k I_{u, X_u^\xi(\omega)}^{\bar{\nu}}(\eta^n, \xi^n) \mathbf{1}_{O_n}(\omega) \\ &\quad + \mathbb{E}[Z(u, \bar{\eta}, \xi, X^\xi, \rho^\xi) | \mathcal{F}_u^t](\omega) \mathbf{1}_{U_k^c}(\omega), \end{aligned} \quad (\text{III.6.39})$$

where $\bar{\eta} := \mathbf{1}_{(u, T]} = \chi^k$ on U_k^c and we have also used that $\zeta^k = \xi$ on U_k^c and so $X^k = X^\xi$ and $\rho^k = \rho^\xi$ on U_k^c . Also notice that $\tau^{\bar{\eta}} := \inf\{s \geq u : \bar{\eta}_s = 1\} \wedge T = u$ and so, by (III.6.9),

$$\begin{aligned} Z(u, \bar{\eta}, \xi, X^\xi, \rho^\xi) \mathbf{1}_{\{\rho^\xi \geq u\}} &= Y(u, \tau^{\bar{\eta}}, \xi, X^\xi, \rho^\xi) \mathbf{1}_{\{\rho^\xi \geq u\}} \\ &= g_1(\rho^\xi, X_{\rho^\xi}^\xi) \mathbf{1}_{\{\rho^\xi = u\}} + g_2(u, X_u^\xi) \mathbf{1}_{\{\rho^\xi > u\}}. \end{aligned} \quad (\text{III.6.40})$$

Furthermore, for every $k \in \mathbb{N}$ and for every $\omega \in \Omega$, we have

$$\begin{aligned} \sum_{n=1}^k I_{u, X_u^\xi(\omega)}^{\bar{\nu}}(\eta^n, \xi^n) \mathbf{1}_{O_n}(\omega) &= \sum_{n=1}^k J_{u, X_u^\xi(\omega)}^{\bar{\nu}}(\tau^n, \xi^n) \mathbf{1}_{O_n}(\omega) \\ &\geq \sum_{n=1}^k \left(J_{u, x_n}^{\bar{\nu}}(\tau^n, \xi^n) - \varepsilon \right) \mathbf{1}_{O_n}(\omega) \\ &\geq \sum_{n=1}^k \left(v(u, x_n) - 1/m - \varepsilon \right) \mathbf{1}_{O_n}(\omega) \\ &\geq \sum_{n=1}^k \left(v(u, X_u^\xi(\omega)) - 1/m - 2\varepsilon \right) \mathbf{1}_{O_n}(\omega) \\ &= \left(v(u, X_u^\xi(\omega)) - 1/m - 2\varepsilon \right) \mathbf{1}_{U_k}(\omega), \end{aligned} \quad (\text{III.6.41})$$

where in the equality on the first line we have used (III.6.22), in the inequalities on the second and fourth lines we have used (III.6.20) and in the inequality on the third line we have used (III.6.21).

Combining (III.6.39), (III.6.40) and (III.6.41), we then obtain, $\mathbb{P}$ -a.s. on $\{\rho^\xi \geq u\}$,

$$\begin{aligned} V_u^\xi &\geq \left(v(u, X_u^\xi) - 1/m - 2\varepsilon \right) \mathbf{1}_{U_k} \\ &\quad \left(g_1(\rho^\xi, X_{\rho^\xi}^\xi) \mathbf{1}_{\{\rho^\xi = u\}} + g_2(u, X_u^\xi) \mathbf{1}_{\{\rho^\xi > u\}} \right) \mathbf{1}_{U_k^c}. \end{aligned} \quad (\text{III.6.42})$$

Letting $k \rightarrow \infty$, as $U_k \uparrow \Omega$ and $U_k^c \downarrow \emptyset$, we obtain

$$V_u^\xi \geq v(u, X_u^\xi) - 1/m - 2\varepsilon, \quad \mathbb{P}\text{-a.s. on } \{\rho^\xi \geq u\}, \quad (\text{III.6.43})$$

Letting $m \rightarrow \infty$ and $\varepsilon \rightarrow 0$ in (III.6.43), we reach the second inequality and the theorem is proved. $\square$

III.7 Dynamic programming principle

In this section we prove that the dynamic programming principle holds for our general problem of singular control with discretionary stopping (see Corollary III.7.3), i.e., for $u \in [t, T]$,

$$\begin{aligned} v(t, x) = \sup_{(\tau, \xi) \in \mathcal{T}_t^\nu \times \mathcal{A}_t^\nu} \mathbb{E} & \left[\int_t^{u \wedge \tau \wedge \rho^\xi} f(s, X_s^\xi) ds - \int_{[t, u \wedge \tau \wedge \rho^\xi)} \langle c_+(s), d\xi_s^+ \rangle \right. \\ & - \int_{[t, u \wedge \tau \wedge \rho^\xi)} \langle c_-(s), d\xi_s^- \rangle + g_1(\rho^\xi, X_{\rho^\xi}^\xi) \mathbf{1}_{\{\rho^\xi \leq \tau\} \cap \{\rho^\xi < u\}} \\ & \left. + g_2(\tau, X_\tau^\xi) \mathbf{1}_{\{\tau < u \wedge \rho^\xi\}} + v(u, X_u^\xi) \mathbf{1}_{\{\tau \wedge \rho^\xi \geq u\}} \right]. \end{aligned} \quad (\text{III.7.1})$$

To obtain this result we make explicit use of the link between the value function and the value process (see Proposition III.6.1). However, notice that, under the same assumptions of Proposition III.6.1, this result could have been directly proved for our value function by means of Proposition III.4.5 and the notion of regular conditional probability.

We first introduce a lemma which will be used in the proof of the dynamic programming principle.

Lemma III.7.1. *Let $u \in [t, T]$ and $\xi \in \mathcal{A}_t^\nu$. The collection of random variables*

$$\{\mathbb{E}[Y(u, \tau, \zeta) | \mathcal{F}_u^t] : (\tau, \zeta) \in \mathcal{T}_u^\nu \times \mathcal{A}_{u, \xi}^\nu\} \quad (\text{III.7.2})$$

is closed under pairwise maximisation. Namely, for every $(\tau^1, \zeta^1) \in \mathcal{T}_u^\nu \times \mathcal{A}_{u, \xi}^\nu$ and $(\tau^2, \zeta^2) \in \mathcal{T}_u^\nu \times \mathcal{A}_{u, \xi}^\nu$ there exists $(\bar{\tau}, \bar{\zeta}) \in \mathcal{T}_u^\nu \times \mathcal{A}_{u, \xi}^\nu$ such that

$$\mathbb{E}[Y(u, \bar{\tau}, \bar{\zeta}) | \mathcal{F}_u^t] = \mathbb{E}[Y(u, \tau^1, \zeta^1) | \mathcal{F}_u^t] \vee \mathbb{E}[Y(u, \tau^2, \zeta^2) | \mathcal{F}_u^t]. \quad (\text{III.7.3})$$

Proof. Let $(\tau^1, \zeta^1) \in \mathcal{T}_u^\nu \times \mathcal{A}_{u, \xi}^\nu$ and $(\tau^2, \zeta^2) \in \mathcal{T}_u^\nu \times \mathcal{A}_{u, \xi}^\nu$. Set

$$A := \left\{ \mathbb{E}[Y(u, \tau^1, \zeta^1) | \mathcal{F}_u^t] \geq \mathbb{E}[Y(u, \tau^2, \zeta^2) | \mathcal{F}_u^t] \right\} \in \mathcal{F}_u^t, \quad (\text{III.7.4})$$

and define

$$\bar{\tau} := \tau^1 \mathbf{1}_A + \tau^2 \mathbf{1}_{A^c}, \quad (\text{III.7.5})$$

$$\bar{\zeta}_s^\pm := \zeta_s^{1, \pm} \mathbf{1}_A + \zeta_s^{2, \pm} \mathbf{1}_{A^c}, \quad s \in [t, T], \quad (\text{III.7.6})$$

$$\bar{\zeta} := \bar{\zeta}^+ - \bar{\zeta}^-. \quad (\text{III.7.7})$$

Then, notice that $(\bar{\tau}, \bar{\zeta}) \in \mathcal{T}_u^\nu \times \mathcal{A}_{u, \xi}^\nu$ and that

$$\begin{aligned} \mathbb{E}[Y(u, \bar{\tau}, \bar{\zeta}) | \mathcal{F}_u^t] &= \mathbb{E}[Y(u, \tau^1, \zeta^1) | \mathcal{F}_u^t] \mathbf{1}_A + \mathbb{E}[Y(u, \tau^2, \zeta^2) | \mathcal{F}_u^t] \mathbf{1}_{A^c} \\ &= \mathbb{E}[Y(u, \tau^1, \zeta^1) | \mathcal{F}_u^t] \vee \mathbb{E}[Y(u, \tau^2, \zeta^2) | \mathcal{F}_u^t], \end{aligned} \quad (\text{III.7.8})$$

which proves the desired result. $\square$

Notice that for $\xi \in \mathcal{A}_t^\nu$, we have, by definition (III.6.2), that $V_t^\xi = v(t, x)$. Thus, we shall simply denote V_t^ξ by V_t .

Theorem III.7.2 (Dynamic programming principle). *Let $(t, x) \in [0, T] \times \mathbb{R}^d$, $u \in [t, T]$ and*

$\nu = (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_s^t\}, W) \in \mathcal{V}_t$. Assume that, for every $\xi \in \mathcal{A}_t^\nu$, we have

$$\mathbb{E}[g_1(\rho^\xi, X_{\rho^\xi}^\xi) \mathbf{1}_{\{\rho^\xi = u\}} + g_2(u, X_u^\xi) \mathbf{1}_{\{\rho^\xi > u\}}] > -\infty. \quad (\text{III.7.9})$$

Then,

$$\begin{aligned} V_t = \sup_{(\tau, \xi) \in \mathcal{T}_t^\nu \times \mathcal{A}_t^\nu} \mathbb{E} & \left[\int_t^{u \wedge \tau \wedge \rho^\xi} f(s, X_s^\xi) ds - \int_{[t, u \wedge \tau \wedge \rho^\xi)} \langle c_+(s), d\xi_s^+ \rangle \right. \\ & - \int_{[t, u \wedge \tau \wedge \rho^\xi)} \langle c_-(s), d\xi_s^- \rangle + g_1(\rho^\xi, X_{\rho^\xi}^\xi) \mathbf{1}_{\{\rho^\xi \leq \tau\} \cap \{\rho^\xi < u\}} \\ & \left. + g_2(\tau, X_\tau^\xi) \mathbf{1}_{\{\tau < u \wedge \rho^\xi\}} + V_u^\xi \mathbf{1}_{\{\tau \wedge \rho^\xi \geq u\}} \right]. \quad (\text{III.7.10}) \end{aligned}$$

Proof. We prove the equality in (III.7.10) by proving the two inequalities.

Part 1. (inequality $\leq$). Let $(\tau, \xi) \in \mathcal{T}_t^\nu \times \mathcal{A}_t^\nu$. Then,

$$\begin{aligned} J_{t,x}^\nu(\tau, \xi) = \mathbb{E} & \left[\int_t^{u \wedge \tau \wedge \rho^\xi} f(s, X_s^\xi) ds \right. \\ & - \int_{[t, u \wedge \tau \wedge \rho^\xi)} \langle c_+(s), d\xi_s^+ \rangle - \int_{[t, u \wedge \tau \wedge \rho^\xi)} \langle c_-(s), d\xi_s^- \rangle \\ & \left. + g_1(\rho^\xi, X_{\rho^\xi}^\xi) \mathbf{1}_{\{\rho^\xi \leq \tau\}} \mathbf{1}_{\{\tau \wedge \rho^\xi < u\}} + g_2(\tau, X_\tau^\xi) \mathbf{1}_{\{\rho^\xi > \tau\}} \mathbf{1}_{\{\tau \wedge \rho^\xi < u\}} \right] \\ & + \mathbb{E} \left[\left(\int_u^{\tau \wedge \rho^\xi} f(s, X_s^\xi) ds \right. \right. \\ & - \int_{[u, \tau \wedge \rho^\xi)} \langle c_+(s), d\xi_s^+ \rangle - \int_{[u, \tau \wedge \rho^\xi)} \langle c_-(s), d\xi_s^- \rangle \\ & \left. \left. + g_1(\rho^\xi, X_{\rho^\xi}^\xi) \mathbf{1}_{\{\rho^\xi \leq \tau\}} + g_2(\tau, X_\tau^\xi) \mathbf{1}_{\{\rho^\xi > \tau\}} \right) \mathbf{1}_{\{\tau \wedge \rho^\xi \geq u\}} \right]. \quad (\text{III.7.11}) \end{aligned}$$

Let $\sigma := u \mathbf{1}_{\{\tau < u\}} + \tau \mathbf{1}_{\{\tau \geq u\}}$ and notice that $\sigma \in \mathcal{T}_u^\nu$ and $\sigma = \tau$ on $\{\tau \geq u\}$. Hence, for the second expectation in (III.7.11), we have

$$\begin{aligned} & \mathbb{E} \left[\left(\int_u^{\tau \wedge \rho^\xi} f(s, X_s^\xi) ds - \int_{[u, \tau \wedge \rho^\xi)} \langle c_+(s), d\xi_s^+ \rangle - \int_{[u, \tau \wedge \rho^\xi)} \langle c_-(s), d\xi_s^- \rangle \right. \right. \\ & \quad \left. \left. + g_1(\rho^\xi, X_{\rho^\xi}^\xi) \mathbf{1}_{\{\rho^\xi \leq \tau\}} + g_2(\tau, X_\tau^\xi) \mathbf{1}_{\{\rho^\xi > \tau\}} \right) \mathbf{1}_{\{\tau \wedge \rho^\xi \geq u\}} \right] \\ & = \mathbb{E} \left[\mathbb{E} \left[Y(u, \sigma, \xi) \mathbf{1}_{\{\rho^\xi \geq u\}} \middle| \mathcal{F}_u^t \right] \mathbf{1}_{\{\tau \wedge \rho^\xi \geq u\}} \right] \\ & \leq \mathbb{E} \left[V_u^\xi \mathbf{1}_{\{\tau \wedge \rho^\xi \geq u\}} \right], \quad (\text{III.7.12}) \end{aligned}$$

where the last inequality follows from the fact that $(\sigma, \xi) \in \mathcal{T}_u^\nu \times \mathcal{A}_{u,\xi}^\nu$. Therefore, from (III.7.11)

and (III.7.12), we obtain

$$\begin{aligned}
 J_{t,x}^\nu(\tau, \xi) &\leq \mathbb{E} \left[\int_t^{u \wedge \tau \wedge \rho^\xi} f(s, X_s^\xi) ds - \int_{[t, u \wedge \tau \wedge \rho^\xi)} \langle c_+(s), d\xi_s^+ \rangle \right. \\
 &\quad - \int_{[t, u \wedge \tau \wedge \rho^\xi)} \langle c_-(s), d\xi_s^- \rangle + g_1(\rho^\xi, X_{\rho^\xi}^\xi) \mathbf{1}_{\{\rho^\xi \leq \tau\} \cap \{\rho^\xi < u\}} \\
 &\quad \left. + g_2(\tau, X_\tau^\xi) \mathbf{1}_{\{\tau < u \wedge \rho^\xi\}} + V_u^\xi \mathbf{1}_{\{\tau \wedge \rho^\xi \geq u\}} \right]. \tag{III.7.13}
 \end{aligned}$$

Thus, by taking the supremum over $(\tau, \xi) \in \mathcal{T}_t^\nu \times \mathcal{A}_t^\nu$, we have

$$\begin{aligned}
 V_t &= v(t, x) \\
 &\leq \sup_{(\tau, \xi) \in \mathcal{T}_t^\nu \times \mathcal{A}_t^\nu} \mathbb{E} \left[\int_t^{u \wedge \tau \wedge \rho^\xi} f(s, X_s^\xi) ds - \int_{[t, u \wedge \tau \wedge \rho^\xi)} \langle c_+(s), d\xi_s^+ \rangle \right. \\
 &\quad - \int_{[t, u \wedge \tau \wedge \rho^\xi)} \langle c_-(s), d\xi_s^- \rangle + g_1(\rho^\xi, X_{\rho^\xi}^\xi) \mathbf{1}_{\{\rho^\xi \leq \tau\} \cap \{\rho^\xi < u\}} \\
 &\quad \left. + g_2(\tau, X_\tau^\xi) \mathbf{1}_{\{\tau < u \wedge \rho^\xi\}} + V_u^\xi \mathbf{1}_{\{\tau \wedge \rho^\xi \geq u\}} \right], \tag{III.7.14}
 \end{aligned}$$

and the first inequality is proved.

Part 2. (inequality $\geq$). Let $(\tau, \xi) \in \mathcal{T}_t^\nu \times \mathcal{A}_t^\nu$. Since, by Lemma III.7.1, the family $\{\mathbb{E}[Y(u, \sigma, \zeta) | \mathcal{F}_u^t] : (\sigma, \zeta) \in \mathcal{T}_u^\nu \times \mathcal{A}_{u,\xi}^\nu\}$ is upward directed then there exists a sequence $(\tau^n, \zeta^n)_{n \in \mathbb{N}} \subseteq \mathcal{T}_u^\nu \times \mathcal{A}_{u,\xi}^\nu$ such that (see, e.g., [79, Lemma 1.3])

$$\lim_{n \rightarrow \infty} \mathbb{E}[Y(u, \tau^n, \zeta^n) | \mathcal{F}_u^t] \mathbf{1}_{\{\rho^\xi \geq u\}} = V_u^\xi, \quad \mathbb{P}\text{-a.s.}, \tag{III.7.15}$$

where the sequence $(\mathbb{E}[Y(u, \tau^n, \zeta^n) | \mathcal{F}_u^t])_{n \in \mathbb{N}}$ is non-decreasing $\mathbb{P}$ -a.s. Also, notice that, without loss of generality, we can assume that $(\tau^1, \zeta^1) := (u, \xi)$. Now, for every $n \in \mathbb{N}$, let $\sigma^n := \tau \mathbf{1}_{\{\tau < u\}} + \tau^n \mathbf{1}_{\{\tau \geq u\}} \in \mathcal{T}_t^\nu$ and set $\rho^n := \rho^{\zeta^n}$ and $X^n := X^{\zeta^n}$. Thus, for every $n \in \mathbb{N}$, we obtain

$$\begin{aligned}
 V_t &\geq \mathbb{E}[Y(t, \sigma^n, \zeta^n)] \\
 &= \mathbb{E} \left[\int_t^{u \wedge \sigma^n \wedge \rho^n} f(s, X_s^\xi) ds \right. \\
 &\quad - \int_{[t, u \wedge \sigma^n \wedge \rho^n)} \langle c_+(s), d\xi_s^+ \rangle - \int_{[t, u \wedge \sigma^n \wedge \rho^n)} \langle c_-(s), d\xi_s^- \rangle \\
 &\quad \left. + g_1(\rho^n, X_{\rho^n}^\xi) \mathbf{1}_{\{\rho^n \leq \sigma^n\}} \mathbf{1}_{\{\sigma^n \wedge \rho^n < u\}} + g_2(\sigma^n, X_{\sigma^n}^\xi) \mathbf{1}_{\{\rho^n > \sigma^n\}} \mathbf{1}_{\{\sigma^n \wedge \rho^n < u\}} \right] \\
 &\quad + \mathbb{E} \left[\left(\int_u^{\sigma^n \wedge \rho^n} f(s, X_s^n) ds \right. \right. \\
 &\quad - \int_{[u, \sigma^n \wedge \rho^n)} \langle c_+(s), d(\zeta^n)_s^+ \rangle - \int_{[u, \sigma^n \wedge \rho^n)} \langle c_-(s), d(\zeta^n)_s^- \rangle \\
 &\quad \left. \left. + g_1(\rho^n, X_{\rho^n}^n) \mathbf{1}_{\{\rho^n \leq \sigma^n\}} + g_2(\sigma^n, X_{\sigma^n}^n) \mathbf{1}_{\{\rho^n > \sigma^n\}} \right) \mathbf{1}_{\{\sigma^n \wedge \rho^n \geq u\}} \right]. \tag{III.7.16}
 \end{aligned}$$

By definition of σ^n , notice that $u \wedge \sigma^n = u \wedge \tau$ and that $\{\sigma^n \geq u\} = \{\tau \geq u\}$ (and so $\{\sigma^n < u\} = \{\tau < u\}$). Moreover, we have that $u \wedge \rho^n = u \wedge \rho^\xi$, $\rho^n = \rho^\xi$ on $\{\rho^n < u\}$ and that

$\{\rho^n < u\} = \{\rho^\xi < u\}$ (and so also $\{\rho^n \geq u\} = \{\rho^\xi \geq u\}$). Hence,

$$\begin{aligned}
 \{\rho^n \leq \sigma^n\} \cap \{\sigma^n \wedge \rho^n < u\} &= (\{\rho^n < \sigma^n\} \cup \{\rho^n = \sigma^n\}) \cap \{\rho^n < u\} \\
 &= \{\rho^n < \sigma^n \wedge u\} \cup (\{\rho^n = \sigma^n\} \cap \{\rho^n < u\}) \\
 &= \{\rho^\xi < \sigma^n \wedge u\} \cup (\{\rho^n = \sigma^n\} \cap \{\rho^n < u\} \cap \{\sigma^n < u\}) \\
 &= \{\rho^\xi < \tau \wedge u\} \cup (\{\rho^n = \sigma^n\} \cap \{\rho^\xi < u\} \cap \{\tau < u\}) \\
 &= (\{\rho^\xi < \tau\} \cap \{\rho^\xi < u\}) \cup (\{\rho^\xi = \tau\} \cap \{\rho^\xi < u\}) \\
 &= \{\rho^\xi < u\} \cap (\{\rho^\xi < \tau\} \cup \{\rho^\xi = \tau\}) \\
 &= \{\rho^\xi < u\} \cap \{\rho^\xi \leq \tau\},
 \end{aligned} \tag{III.7.17}$$

and

$$\{\rho^n > \sigma^n\} \cap \{\sigma^n \wedge \rho^n < u\} = \{\sigma^n < u \wedge \rho^n\} = \{\sigma^n < u \wedge \rho^\xi\} = \{\tau < u \wedge \rho^\xi\}. \tag{III.7.18}$$

Thus, for the first expectation in (III.7.16), we have

$$\begin{aligned}
 &\mathbb{E} \left[\int_t^{u \wedge \sigma^n \wedge \rho^n} f(s, X_s^\xi) ds - \int_{[t, u \wedge \sigma^n \wedge \rho^n)} \langle c_+(s), d\xi_s^+ \rangle - \int_{[t, u \wedge \sigma^n \wedge \rho^n)} \langle c_-(s), d\xi_s^- \rangle \right. \\
 &\quad \left. + g_1(\rho^n, X_{\rho^n}^\xi) \mathbf{1}_{\{\rho^n \leq \sigma^n\}} \mathbf{1}_{\{\sigma^n \wedge \rho^n < u\}} + g_2(\sigma^n, X_{\sigma^n}^\xi) \mathbf{1}_{\{\rho^n > \sigma^n\}} \mathbf{1}_{\{\sigma^n \wedge \rho^n < u\}} \right] \\
 &= \mathbb{E} \left[\int_t^{u \wedge \tau \wedge \rho^\xi} f(s, X_s^\xi) ds - \int_{[t, u \wedge \tau \wedge \rho^\xi)} \langle c_+(s), d\xi_s^+ \rangle - \int_{[t, u \wedge \tau \wedge \rho^\xi)} \langle c_-(s), d\xi_s^- \rangle \right. \\
 &\quad \left. + g_1(\rho^\xi, X_{\rho^\xi}^\xi) \mathbf{1}_{\{\rho^\xi \leq \tau\}} \mathbf{1}_{\{\rho^\xi < u\}} + g_2(\tau, X_\tau^\xi) \mathbf{1}_{\{\tau < u \wedge \rho^\xi\}} \right].
 \end{aligned} \tag{III.7.19}$$

Moreover,

$$\{\sigma^n \wedge \rho^n \geq u\} = \{\sigma^n \geq u\} \cap \{\rho^n \geq u\} = \{\tau \geq u\} \cap \{\rho^\xi \geq u\} = \{\tau \wedge \rho^\xi \geq u\}, \tag{III.7.20}$$

and so, for the second expectation in (III.7.16) recalling that $\sigma^n = \tau^n$ on $\{\tau \geq u\}$, we have

$$\begin{aligned}
 &\mathbb{E} \left[\left(\int_u^{\sigma^n \wedge \rho^n} f(s, X_s^n) ds - \int_{[u, \sigma^n \wedge \rho^n)} \langle c_+(s), d(\zeta^n)_s^+ \rangle - \int_{[u, \sigma^n \wedge \rho^n)} \langle c_-(s), d(\zeta^n)_s^- \rangle \right. \right. \\
 &\quad \left. \left. + g_1(\rho^n, X_{\rho^n}^n) \mathbf{1}_{\{\rho^n \leq \sigma^n\}} + g_2(\sigma^n, X_{\sigma^n}^n) \mathbf{1}_{\{\rho^n > \sigma^n\}} \right) \mathbf{1}_{\{\sigma^n \wedge \rho^n \geq u\}} \right] \\
 &= \mathbb{E} \left[\mathbb{E} \left[Y(u, \tau^n, \zeta^n) \middle| \mathcal{F}_u^t \right] \mathbf{1}_{\{\tau \wedge \rho^\xi \geq u\}} \right].
 \end{aligned} \tag{III.7.21}$$

Thus, by using (III.7.19) and (III.7.21), equation (III.7.16) reads as

$$\begin{aligned}
 V_t &\geq \mathbb{E} \left[\int_t^{u \wedge \tau \wedge \rho^\xi} f(s, X_s^\xi) ds - \int_{[t, u \wedge \tau \wedge \rho^\xi)} \langle c_+(s), d\xi_s^+ \rangle - \int_{[t, u \wedge \tau \wedge \rho^\xi)} \langle c_-(s), d\xi_s^- \rangle \right. \\
 &\quad \left. + g_1(\rho^\xi, X_{\rho^\xi}^\xi) \mathbf{1}_{\{\rho^\xi \leq \tau\}} \mathbf{1}_{\{\rho^\xi < u\}} + g_2(\tau, X_\tau^\xi) \mathbf{1}_{\{\tau < u \wedge \rho^\xi\}} \right] \\
 &\quad + \mathbb{E} \left[\mathbb{E} \left[Y(u, \tau^n, \zeta^n) \middle| \mathcal{F}_u^t \right] \mathbf{1}_{\{\tau \wedge \rho^\xi \geq u\}} \right].
 \end{aligned} \tag{III.7.22}$$

Letting $n \rightarrow \infty$, applying the monotone convergence theorem (which is justified by hypothesis

(III.7.9)) and using (III.7.15), we obtain

$$V_t \geq \mathbb{E} \left[\int_t^{u \wedge \tau \wedge \rho^\xi} f(s, X_s^\xi) ds - \int_{[t, u \wedge \tau \wedge \rho^\xi)} \langle c_+(s), d\xi_s^+ \rangle - \int_{[t, u \wedge \tau \wedge \rho^\xi)} \langle c_-(s), d\xi_s^- \rangle \right. \\ \left. + g_1(\rho^\xi, X_{\rho^\xi}^\xi) \mathbf{1}_{\{\rho^\xi \leq \tau\} \cap \{\rho^\xi < u\}} + g_2(\tau, X_\tau^\xi) \mathbf{1}_{\{\tau < u \wedge \rho^\xi\}} + V_u^\xi \mathbf{1}_{\{\tau \wedge \rho^\xi \geq u\}} \right]. \quad (\text{III.7.23})$$

Therefore, by taking the supremum over $(\tau, \xi) \in \mathcal{T}_t^\nu \times \mathcal{A}_t^\nu$, we reach the second inequality

$$V_t \geq \sup_{(\tau, \xi) \in \mathcal{T}_t^\nu \times \mathcal{A}_t^\nu} \mathbb{E} \left[\int_t^{u \wedge \tau \wedge \rho^\xi} f(s, X_s^\xi) ds - \int_{[t, u \wedge \tau \wedge \rho^\xi)} \langle c_+(s), d\xi_s^+ \rangle \right. \\ \left. - \int_{[t, u \wedge \tau \wedge \rho^\xi)} \langle c_-(s), d\xi_s^- \rangle + g_1(\rho^\xi, X_{\rho^\xi}^\xi) \mathbf{1}_{\{\rho^\xi \leq \tau\} \cap \{\rho^\xi < u\}} \right. \\ \left. + g_2(\tau, X_\tau^\xi) \mathbf{1}_{\{\tau < u \wedge \rho^\xi\}} + V_u^\xi \mathbf{1}_{\{\tau \wedge \rho^\xi \geq u\}} \right], \quad (\text{III.7.24})$$

and the theorem is proved. $\square$

Corollary III.7.3. *Let $(t, x) \in [0, T] \times \mathbb{R}^d$, $u \in [t, T]$ and $\nu = (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_s^t\}, W) \in \mathcal{V}_t$. Under the assumptions of Proposition III.6.1 and of Theorem III.7.2, we obtain that*

$$v(t, x) = \sup_{(\tau, \xi) \in \mathcal{T}_t^\nu \times \mathcal{A}_t^\nu} \mathbb{E} \left[\int_t^{u \wedge \tau \wedge \rho^\xi} f(s, X_s^\xi) ds - \int_{[t, u \wedge \tau \wedge \rho^\xi)} \langle c_+(s), d\xi_s^+ \rangle \right. \\ \left. - \int_{[t, u \wedge \tau \wedge \rho^\xi)} \langle c_-(s), d\xi_s^- \rangle + g_1(\rho^\xi, X_{\rho^\xi}^\xi) \mathbf{1}_{\{\rho^\xi \leq \tau\} \cap \{\rho^\xi < u\}} \right. \\ \left. + g_2(\tau, X_\tau^\xi) \mathbf{1}_{\{\tau < u \wedge \rho^\xi\}} + v(u, X_u^\xi) \mathbf{1}_{\{\tau \wedge \rho^\xi \geq u\}} \right]. \quad (\text{III.7.25})$$

Proof. Simply use (III.7.10) and the fact that, by Proposition III.6.1, $V_u^\xi = v(u, X_u^\xi)$, $\mathbb{P}$ -a.s. on $\{\rho^\xi \geq u\}$. $\square$

Remark III.7.4. *Notice that the dynamic programming principle for the value function (III.7.25) could have been proved directly by means of the regular conditional probabilities with a very similar method as in the proof of Proposition III.6.1. However, our approach leads to a much smoother proof of the dynamic programming principle which does not require the same technicalities as in the proof of Proposition III.6.1.*

Remark III.7.5. *Notice that, for the controls $\xi \in \mathcal{A}_t^\nu$, the condition $\mathbb{E}[|\xi_T^\pm|^2] < \infty$ has not been explicitly used in any of the previous proofs. Therefore, our study could have been developed without assuming that the class of admissible controls $\mathcal{A}_t^\nu$ satisfies this integrability condition and the same results would follow. In particular, notice that the existence and uniqueness (up to indistinguishability) of the SDE (III.3.1) is guaranteed by, e.g., [36, Theorem 1] without assuming this condition. However, we preferred this approach as the integrability assumption will be essential to provide some regularity properties for the value function in Section III.9.*

III.8 An optimal stopping rule

In this section we provide an optimal stopping rule for our problem when the terminal payoff is symmetric, i.e., $g := g_1 = g_2$. Namely, we show that

$$v(t, x) = \sup_{\xi \in \mathcal{A}_t^\nu} J_{t,x}^\nu(\tau^\xi, \xi), \quad (\text{III.8.1})$$

where

$$\tau^\xi := \inf\{s \geq t : v(s, X_s^\xi) = g(s, X_s^\xi)\}. \quad (\text{III.8.2})$$

Notice that this is, indeed, the optimal stopping rule we would expect from the standard theory of optimal stopping. However, we are able to show this result only under some technical assumptions which we will discuss in this section.

Since the standard reference probability system $\nu \in \mathcal{V}_t$ is fixed, in this section we avoid any explicit reference to ν for the sake of simplicity. For instance, we simply denote $\mathcal{A}_t^\nu$ by $\mathcal{A}_t$.

To obtain (III.8.1), we make explicit use of the formulation of the problem with respect to continuous controls and Lipschitz controls. Thus, we define the value function for continuous controls

$$v^c(t, x) := \sup_{(\tau, \xi) \in \mathcal{T}_t \times \mathcal{A}_t^c} J_{t,x}(\tau, \xi), \quad (\text{III.8.3})$$

where $\mathcal{A}_t^c := \{\xi \in \mathcal{A}_t : \xi \text{ is continuous } \mathbb{P}\text{-a.s.}\}$. Moreover, for $n \in \mathbb{N}$, we define the value function for n^2 -Lipschitz controls

$$v_n(t, x) := \sup_{(\tau, \xi) \in \mathcal{T}_t \times \mathcal{A}_t^n} J_{t,x}(\tau, \xi), \quad (\text{III.8.4})$$

where $\mathcal{A}_t^n := \{\xi \in \mathcal{A}_t : \xi \text{ is } n^2\text{-Lipschitz } \mathbb{P}\text{-a.s.}\}$.

In this section we make the following standing assumptions.

Assumption III.8.1.

(A) $g_1(s, y) = g_2(s, y) =: g(s, y)$ for every $(s, y) \in [t, T] \times \mathbb{R}^d$ and $(s, y) \mapsto g(s, y)$ is continuous.

(B) For every $n \in \mathbb{N}$, there exists $\bar{\xi}^n \in \mathcal{A}_t^n$ such that

$$v_n(t, x) = \sup_{\tau \in \mathcal{T}_t} J_{t,x}(\tau, \bar{\xi}^n). \quad (\text{III.8.5})$$

(C) It either holds

$$\mathbb{E} \left[\sup_{s \in [t, T]} |Y(t, s, \bar{\xi}^n)| \right] < \infty \quad \text{and} \quad Y(t, \cdot, \bar{\xi}^n) \geq 0, \quad (\text{III.8.6})$$

or there exists $\varepsilon > 0$ such that

$$\mathbb{E} \left[\sup_{s \in [t, T]} |Y(t, s, \bar{\xi}^n)|^{1+\varepsilon} \right] < \infty. \quad (\text{III.8.7})$$

Remark III.8.2. Notice that when we restrict the optimisation to n^2 -Lipschitz controls, the singular control problem turns into a classical control problem. Indeed, if $\xi \in \mathcal{A}_t^n$, then it admits a derivative α almost everywhere, i.e., $\xi_s = \int_t^s \alpha_r dr$ for $s \in [t, T]$. Thus,

$$X_s^\xi = \tilde{X}_s^\alpha := x + \int_t^s \tilde{\mu}(r, \tilde{X}_r^\alpha, \alpha_r) dr + \int_t^s \sigma(r, \tilde{X}_r^\alpha) dW_r, \quad (\text{III.8.8})$$

where $\tilde{\mu}(r, y, a) := \mu(r, y) + a$. Moreover, for $\tau \in \mathcal{T}_t$,

$$J_{t,x}(\tau, \xi) = \tilde{J}_{t,x}(\tau, \alpha) := \mathbb{E} \left[\int_t^\tau \tilde{f}(s, \tilde{X}_s^\alpha, \alpha_s) ds + g(\tau, \tilde{X}_\tau^\alpha) \right], \quad (\text{III.8.9})$$

where $\tilde{f}(s, y, a) := f(s, y) - \langle c_+(s), [a]^+ \rangle - \langle c_-(s), [a]^- \rangle$, and we have assumed that $\mathcal{O} = \mathbb{R}^d$ for simplicity. Hence,

$$v_n(t, x) = \sup_{(\tau, \alpha) \in \mathcal{T}_t \times \tilde{\mathcal{A}}^n} \tilde{J}_{t,x}(\tau, \alpha), \quad (\text{III.8.10})$$

where $\tilde{\mathcal{A}}^n$ is the collection of progressively measurable processes α such that $|\alpha| \leq n^2$. The existence of optimal controls for these problems have been proved by different methods (see, e.g., [66, Theorem 9], [10, Theorem 5.2] and [3, Theorem 3.2]). Therefore, we remark that although the proof of Assumption III.8.1 (B) goes beyond the scope of this work, it may be obtained.

We now fix $n \in \mathbb{N}$. In what follows, when no confusion may arise, we denote the process $\{Y(t, s, \bar{\xi}^n)\}_{s \in [t, T]}$ by $\bar{Y}^n$ and analogously the process $X^{\bar{\xi}^n}$ by $\bar{X}^n$. Consequently, we also denote $\rho^{\bar{\xi}^n}$ by $\bar{\rho}^n$. Notice that since $\bar{\xi}^n$ is n^2 -Lipschitz $\mathbb{P}$ -a.s., then $\bar{X}^n$ is continuous $\mathbb{P}$ -a.s. Thus, under Assumption III.8.1, we have that

$$\begin{aligned} \bar{Y}_u^n &= \int_t^{u \wedge \bar{\rho}^n} f(s, \bar{X}_s^n) ds - \int_{[t, u \wedge \bar{\rho}^n)} \langle c_+(s), d\bar{\xi}_s^{n,+} \rangle \\ &\quad - \int_{[t, u \wedge \bar{\rho}^n)} \langle c_-(s), d\bar{\xi}_s^{n,-} \rangle + g(u \wedge \bar{\rho}^n, \bar{X}_{u \wedge \bar{\rho}^n}^n), \quad u \in [t, T], \end{aligned} \quad (\text{III.8.11})$$

and so also the process $\bar{Y}^n$ is continuous $\mathbb{P}$ -a.s.

We now introduce the stopping problem corresponding to the optimal control $\bar{\xi}^n \in \mathcal{A}_t^n$, i.e.,

$$\bar{\Psi}_u^n := \text{ess sup}_{\tau \in \mathcal{T}_u} \mathbb{E}[\bar{Y}_\tau^n | \mathcal{F}_u^t], \quad u \in [t, T], \quad (\text{III.8.12})$$

and we denote by $\bar{S}^n$ the corresponding Snell Envelope (see, e.g., [65, Appendix D]). Recall that $J_{t,x}(\tau, \xi) = \mathbb{E}[Y(t, \tau, \xi)]$ and thus, by (III.8.5) and (III.8.12), we obtain

$$v_n(t, x) = \bar{\Psi}_t^n. \quad (\text{III.8.13})$$

We then present a lemma which is inspired by [65, Theorem D.12] and provides an optimal stopping rule for problem (III.8.12).

Lemma III.8.3. *Let $u \in [t, T]$. Under Assumption III.8.1, we have that*

$$\bar{\Psi}_u^n = \mathbb{E}[\bar{Y}_{\bar{\sigma}_u^n}^n | \mathcal{F}_u^t], \quad \mathbb{P}\text{-a.s.}, \quad (\text{III.8.14})$$

where

$$\bar{\sigma}_u^n := \inf\{s \geq u : \bar{S}_s^n = \bar{Y}_s^n\}. \quad (\text{III.8.15})$$

Proof. In order to simplify the notation, since $n \in \mathbb{N}$ is fixed, in this proof we avoid any explicit reference to n . Thus, for instance, we simply denote $\bar{Y}^n$ by $\bar{Y}$, $\bar{\Psi}^n$ by $\bar{\Psi}$, et cetera.

If (III.8.6) holds, then result (III.8.14) follows directly from [65, Theorem D.12] (and [65, Theorem D.7]) as the process $\bar{Y}$ is non-negative, continuous and verifies the integrability condition in (III.8.6).

If instead (III.8.7) holds then $\bar{Y}$ is not necessarily non-negative. Thus, define

$$H := \inf_{s \in [t, T]} \bar{Y}_s \quad \text{and} \quad M_s := \mathbb{E}[H | \mathcal{F}_s^t], \quad s \in [t, T]. \quad (\text{III.8.16})$$

Notice that, by (III.8.7), $H \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ and so M is a uniformly integrable, continuous (as $\{\mathcal{F}_s^t\}$ is both left-continuous and right-continuous) martingale. Then, also define $\tilde{Y} := \bar{Y} - M$ and the new related stopping problem

$$\tilde{\Psi}_u := \operatorname{ess\,sup}_{\tau \in \mathcal{T}_u} [\tilde{Y}_\tau | \mathcal{F}_u^t], \quad u \in [t, T]. \quad (\text{III.8.17})$$

Notice that $\tilde{Y}$ is non-negative and continuous. Moreover,

$$\mathbb{E} \left[\sup_{s \in [t, T]} |\tilde{Y}_s| \right] \leq \mathbb{E} \left[\sup_{s \in [t, T]} |\bar{Y}_s| \right] + \mathbb{E} \left[\sup_{s \in [t, T]} |M_s| \right], \quad (\text{III.8.18})$$

By using Jensen's inequality and Doob's maximal inequality, we have

$$\begin{aligned} \mathbb{E} \left[\sup_{s \in [t, T]} |M_s| \right] &\leq \mathbb{E} \left[\left(\sup_{s \in [t, T]} |M_s| \right)^{1+\varepsilon} \right]^{\frac{1}{1+\varepsilon}} \\ &\leq C_\varepsilon \mathbb{E} \left[|M_T|^{1+\varepsilon} \right]^{\frac{1}{1+\varepsilon}} \\ &= C_\varepsilon \mathbb{E} \left[|H|^{1+\varepsilon} \right]^{\frac{1}{1+\varepsilon}} \\ &\leq C_\varepsilon \mathbb{E} \left[\sup_{s \in [t, T]} |\bar{Y}_s|^{1+\varepsilon} \right]^{\frac{1}{1+\varepsilon}} < \infty. \end{aligned} \quad (\text{III.8.19})$$

Thus, $\mathbb{E}[\sup_{s \in [t, T]} |\tilde{Y}_s|] < \infty$. Hence, $\tilde{Y}$ satisfies the hypotheses of [65, Theorem D.12] and so

$$\tilde{\Psi}_u = \mathbb{E}[\tilde{Y}_{\tilde{\sigma}_u} | \mathcal{F}_u^t], \quad \mathbb{P}\text{-a.s.}, \quad (\text{III.8.20})$$

where $\tilde{\sigma}_u := \inf\{s \geq u : \tilde{S}_s = \tilde{Y}_s\}$ and $\tilde{S}$ is the Snell Envelope of $\tilde{Y}$. Now notice, that since M is a continuous martingale, by the optional sampling theorem, we have that

$$\tilde{\Psi}_u = \bar{\Psi}_u - M_u \quad \text{and} \quad \mathbb{E}[\tilde{Y}_{\tilde{\sigma}_u} | \mathcal{F}_u^t] = \mathbb{E}[\bar{Y}_{\tilde{\sigma}_u} | \mathcal{F}_u^t] - M_u, \quad \mathbb{P}\text{-a.s.} \quad (\text{III.8.21})$$

Therefore, from (III.8.20) and (III.8.21), we obtain

$$\bar{\Psi}_u = \mathbb{E}[\bar{Y}_{\tilde{\sigma}_u} | \mathcal{F}_u^t], \quad \mathbb{P}\text{-a.s.} \quad (\text{III.8.22})$$

To conclude, we only need to show that $\tilde{\sigma}_u = \bar{\sigma}_u$ $\mathbb{P}$ -a.s. To obtain this result, we prove that $\tilde{S} - \tilde{Y}$ is indistinguishable from $\bar{S} - \bar{Y}$. Recall that $\bar{S}$ is a right-continuous modification of $\bar{\Psi}$ and $\tilde{S}$ is a right-continuous modification of $\tilde{\Psi}$. Since both $\bar{Y}$ and $\tilde{Y}$ are continuous, it is sufficient to show that $\tilde{S} - \tilde{Y}$ is a modification $\bar{S} - \bar{Y}$. For every $s \in [t, T]$, we have that

$$\tilde{S}_s - \tilde{Y}_s = \tilde{\Psi}_s - \tilde{Y}_s = \bar{\Psi}_s - M_s - (\bar{Y}_s - M_s) = \bar{\Psi}_s - \bar{Y}_s = \bar{S}_s - \bar{Y}_s, \quad \mathbb{P}\text{-a.s.}, \quad (\text{III.8.23})$$

and the proof is complete. $\square$

Let us denote $\bar{\sigma}_t^n$ as in (III.8.15) simply by $\bar{\sigma}^n$ and notice that, by (III.8.13) and (III.8.14), we

obtain

$$v_n(t, x) = \mathbb{E}[\bar{Y}_{\bar{\sigma}^n}^n], \quad (\text{III.8.24})$$

i.e., $(\bar{\sigma}^n, \bar{\xi}^n) \in \mathcal{T}_t \times \mathcal{A}_t^n$ is optimal for $v_n(t, x)$.

Now, for $u \in [t, T]$ and $\xi \in \mathcal{A}_t^n$, we define the value process in the context of Lipschitz controls, i.e.,

$$V_u^{\xi, n} := \operatorname{ess\,sup}_{(\tau, \zeta) \in \mathcal{T}_u \times \mathcal{A}_{u, \xi}^n} \mathbb{E}[Y(u, \tau, \zeta) \mathbf{1}_{\{\rho^\xi \geq u\}} | \mathcal{F}_u^t], \quad (\text{III.8.25})$$

where $\mathcal{A}_{u, \xi}^n := \{\zeta \in \mathcal{A}_t^n : \zeta_s = \xi_s, \forall s \in [t, u], \mathbb{P}\text{-a.s.}\}$. Moreover, we define

$$U_u^\xi := \operatorname{ess\,sup}_{\tau \in \mathcal{T}_u} \mathbb{E}[Y(u, \tau, \xi) \mathbf{1}_{\{\rho^\xi \geq u\}} | \mathcal{F}_u^t], \quad (\text{III.8.26})$$

which represents the value process $V_u^{\xi, n}$ if it were optimal to stick with the control $\xi \in \mathcal{A}_t^n$ also on $[u, T]$. In particular, notice that $V_u^{\xi, n} \geq U_u^\xi$, $\mathbb{P}$ -a.s.

By applying the same simplification as above, we denote $V^{\bar{\xi}^n, n}$ by $\bar{V}^n$ and $U^{\bar{\xi}^n}$ by $\bar{U}^n$. In the following, also recall that $V_t^{\xi, n} = v_n(t, x)$ for every $\xi \in \mathcal{A}_t^n$ and so we simply denote $V_t^{\xi, n}$ by V_t^n . Furthermore, notice that all the results proved in the previous sections for left-continuous controls hold, by similar arguments, for n^2 -Lipschitz controls. In particular, Proposition III.6.1 and Theorem III.7.2 hold also when optimising over n^2 -Lipschitz controls. We now prove a lemma which is a direct consequence of the dynamic programming principle (Theorem III.7.2). This shows that the optimal control $\bar{\xi}^n \in \mathcal{A}_t^n$ for the (n^2 -Lipschitz version of the) problem starting at the initial time t is still optimal for the (n^2 -Lipschitz version of the) problem starting at any time $u \in [t, T]$, i.e., $\bar{V}_u^n = \bar{U}_u^n$, $\mathbb{P}$ -a.s.

Lemma III.8.4. *Let $u \in [t, T]$ and assume that $\mathbb{P}(\bar{\rho}^n \geq u) > 0$. Then, under the hypotheses of Theorem III.7.2 and Assumption III.8.1, we have that*

$$\bar{V}_u^n = \bar{U}_u^n, \quad \mathbb{P}\text{-a.s.} \quad (\text{III.8.27})$$

Proof. Notice that if $\mathbb{P}(\bar{\rho}^n \geq u) = 0$ then, by definitions (III.8.25) and (III.8.26), $\bar{V}_u^n = \bar{U}_u^n = 0$, $\mathbb{P}$ -a.s. Therefore, let $\mathbb{P}(\bar{\rho}^n \geq u) > 0$ and by contradiction assume that

$$\mathbb{P}(\bar{V}_u^n > \bar{U}_u^n) > 0. \quad (\text{III.8.28})$$

Since $V_t^n = v_n(t, x)$ and, by (III.8.24), $v_n(t, x) = \mathbb{E}[\bar{Y}_{\bar{\sigma}^n}^n]$ then also $V_t^n = \mathbb{E}[\bar{Y}_{\bar{\sigma}^n}^n]$, i.e., the pair $(\bar{\sigma}^n, \bar{\xi}^n)$ is optimal for V_t^n . Thus, by following the same steps as in the proof of Theorem III.7.2 with the general (τ, ξ) substituted by the optimal $(\bar{\sigma}^n, \bar{\xi}^n)$, the dynamic programming principle (III.7.10) reads as

$$\begin{aligned} V_t^n = \mathbb{E} \left[\int_t^{u \wedge \bar{\sigma}^n \wedge \bar{\rho}^n} f(s, \bar{X}_s^n) ds - \int_{[t, u \wedge \bar{\sigma}^n \wedge \bar{\rho}^n)} \langle c_+(s), d\bar{\xi}_s^{n, +} \rangle - \int_{[t, u \wedge \bar{\sigma}^n \wedge \bar{\rho}^n)} \langle c_-(s), d\bar{\xi}_s^{n, -} \rangle \right. \\ \left. + g(\bar{\sigma}^n \wedge \bar{\rho}^n, \bar{X}_{\bar{\sigma}^n \wedge \bar{\rho}^n}^n) \mathbf{1}_{\{\bar{\sigma}^n \wedge \bar{\rho}^n < u\}} + \bar{V}_u^n \mathbf{1}_{\{\bar{\sigma}^n \wedge \bar{\rho}^n \geq u\}} \right]. \end{aligned} \quad (\text{III.8.29})$$

Since $\mathbb{P}(\bar{\sigma}^n \geq u) > 0$, $\mathbb{P}(\bar{\rho}^n \geq u) > 0$ and $\mathbb{P}(\bar{V}_u^n > \bar{U}_u^n) > 0$, we obtain

$$\begin{aligned} V_t^n &> \mathbb{E} \left[\int_t^{u \wedge \bar{\sigma}^n \wedge \bar{\rho}^n} f(s, \bar{X}_s^n) ds - \int_{[t, u \wedge \bar{\sigma}^n \wedge \bar{\rho}^n)} \langle c_+(s), d\bar{\xi}_s^{n,+} \rangle - \int_{[t, u \wedge \bar{\sigma}^n \wedge \bar{\rho}^n)} \langle c_-(s), d\bar{\xi}_s^{n,-} \rangle \right. \\ &\quad \left. + g(\bar{\sigma}^n \wedge \bar{\rho}^n, \bar{X}_{\bar{\sigma}^n \wedge \bar{\rho}^n}^n) \mathbf{1}_{\{\bar{\sigma}^n \wedge \bar{\rho}^n < u\}} + \bar{U}_u^n \mathbf{1}_{\{\bar{\sigma}^n \wedge \bar{\rho}^n \geq u\}} \right]. \end{aligned} \quad (\text{III.8.30})$$

Moreover, by choosing $\tau = \bar{\sigma}^n \vee u \in \mathcal{T}_u$ in (III.8.26), we have that

$$\begin{aligned} \bar{U}_u^n \mathbf{1}_{\{\bar{\sigma}^n \wedge \bar{\rho}^n \geq u\}} &\geq \mathbb{E} \left[\left(\int_u^{\bar{\sigma}^n \wedge \bar{\rho}^n} f(s, \bar{X}_s^n) ds - \int_{[u, \bar{\sigma}^n \wedge \bar{\rho}^n)} \langle c_+(s), d\bar{\xi}_s^{n,+} \rangle \right. \right. \\ &\quad \left. \left. - \int_{[t, \bar{\sigma}^n \wedge \bar{\rho}^n)} \langle c_-(s), d\bar{\xi}_s^{n,-} \rangle + g(\bar{\sigma}^n \wedge \bar{\rho}^n, \bar{X}_{\bar{\sigma}^n \wedge \bar{\rho}^n}^n) \right) \mathbf{1}_{\{\bar{\sigma}^n \wedge \bar{\rho}^n \geq u\}} \middle| \mathcal{F}_u^t \right] \\ &= \mathbb{E} \left[\int_{u \wedge \bar{\sigma}^n \wedge \bar{\rho}^n}^{\bar{\sigma}^n \wedge \bar{\rho}^n} f(s, \bar{X}_s^n) ds - \int_{[u \wedge \bar{\sigma}^n \wedge \bar{\rho}^n, \bar{\sigma}^n \wedge \bar{\rho}^n)} \langle c_+(s), d\bar{\xi}_s^{n,+} \rangle \right. \\ &\quad \left. - \int_{[u \wedge \bar{\sigma}^n \wedge \bar{\rho}^n, \bar{\sigma}^n \wedge \bar{\rho}^n)} \langle c_-(s), d\bar{\xi}_s^{n,-} \rangle \right. \\ &\quad \left. + g(\bar{\sigma}^n \wedge \bar{\rho}^n, \bar{X}_{\bar{\sigma}^n \wedge \bar{\rho}^n}^n) \mathbf{1}_{\{\bar{\sigma}^n \wedge \bar{\rho}^n \geq u\}} \middle| \mathcal{F}_u^t \right]. \end{aligned} \quad (\text{III.8.31})$$

Thus, from (III.8.30) and (III.8.31), we obtain

$$\begin{aligned} V_t^n &> \mathbb{E} \left[\int_t^{\bar{\sigma}^n \wedge \bar{\rho}^n} f(s, \bar{X}_s^n) ds - \int_{[t, \bar{\sigma}^n \wedge \bar{\rho}^n)} \langle c_+(s), d\bar{\xi}_s^{n,+} \rangle \right. \\ &\quad \left. - \int_{[t, \bar{\sigma}^n \wedge \bar{\rho}^n)} \langle c_-(s), d\bar{\xi}_s^{n,-} \rangle + g(\bar{\sigma}^n \wedge \bar{\rho}^n, \bar{X}_{\bar{\sigma}^n \wedge \bar{\rho}^n}^n) \right] \\ &= \mathbb{E} [\bar{Y}_{\bar{\sigma}^n}^n] \\ &= V_t^n, \end{aligned} \quad (\text{III.8.32})$$

which leads to a contradiction $\square$

We now prove a lemma which provides an optimal stopping rule for the problem with n^2 -Lipschitz controls.

Lemma III.8.5. *Assume that the function $(s, y) \mapsto v_n(s, y)$ is continuous on $[t, T] \times \mathbb{R}^d$. Under the hypotheses of Proposition III.6.1 and of Theorem III.7.2 and under Assumption III.8.1, we have that*

$$v_n(t, x) = \sup_{\xi \in \mathcal{A}_t^n} J_{t,x}(\tau^{\xi,n}, \xi), \quad (\text{III.8.33})$$

where

$$\tau^{\xi,n} := \inf\{s \geq t : v_n(s, X_s^\xi) = g(s, X_s^\xi)\}. \quad (\text{III.8.34})$$

Proof. It is immediate to see that $v_n(t, x) \geq \sup_{\xi \in \mathcal{A}_t^n} J_{t,x}(\tau^{\xi,n}, \xi)$ as $\tau^{\xi,n} \in \mathcal{T}_t$ for every $\xi \in \mathcal{A}_t^n$. Thus, we only need to prove the opposite inequality, which we first show under the assumption that $\mathbb{P}(\bar{\sigma}^n \geq u) > 0$ for every $u \in [t, T]$.

Let $u \in [t, T]$. For $\tau \in \mathcal{T}_u$, we have

$$\begin{aligned}
 Y(t, \tau, \bar{\xi}^n) &= \int_t^{u \wedge \bar{\rho}^n} f(s, \bar{X}_s^n) ds - \int_{[t, u \wedge \bar{\rho}^n)} \langle c_+(s), d\bar{\xi}_s^{n,+} \rangle - \int_{[t, u \wedge \bar{\rho}^n)} \langle c_-(s), d\bar{\xi}_s^{n,-} \rangle \\
 &\quad + g(\bar{\rho}^n, \bar{X}_{\bar{\rho}^n}^n) \mathbf{1}_{\{\bar{\rho}^n \leq u\}} + \int_{u \wedge \bar{\rho}^n}^{\tau \wedge \bar{\rho}^n} f(s, \bar{X}_s^n) ds - \int_{[u \wedge \bar{\rho}^n, \tau \wedge \bar{\rho}^n)} \langle c_+(s), d\bar{\xi}_s^{n,+} \rangle \\
 &\quad - \int_{[u \wedge \bar{\rho}^n, \tau \wedge \bar{\rho}^n)} \langle c_-(s), d\bar{\xi}_s^{n,-} \rangle + g(\tau \wedge \bar{\rho}^n, \bar{X}_{\tau \wedge \bar{\rho}^n}^n) \mathbf{1}_{\{\bar{\rho}^n > u\}}. \tag{III.8.35}
 \end{aligned}$$

Furthermore,

$$\begin{aligned}
 &\int_{u \wedge \bar{\rho}^n}^{\tau \wedge \bar{\rho}^n} f(s, \bar{X}_s^n) ds - \int_{[u \wedge \bar{\rho}^n, \tau \wedge \bar{\rho}^n)} \langle c_+(s), d\bar{\xi}_s^{n,+} \rangle \\
 &- \int_{[u \wedge \bar{\rho}^n, \tau \wedge \bar{\rho}^n)} \langle c_-(s), d\bar{\xi}_s^{n,-} \rangle + g(\tau \wedge \bar{\rho}^n, \bar{X}_{\tau \wedge \bar{\rho}^n}^n) \mathbf{1}_{\{\bar{\rho}^n > u\}} \\
 &= \left(\int_u^{\tau \wedge \bar{\rho}^n} f(s, \bar{X}_s^n) ds - \int_{[u, \tau \wedge \bar{\rho}^n)} \langle c_+(s), d\bar{\xi}_s^{n,+} \rangle \right. \\
 &\quad \left. - \int_{[u, \tau \wedge \bar{\rho}^n)} \langle c_-(s), d\bar{\xi}_s^{n,-} \rangle + g(\tau \wedge \bar{\rho}^n, \bar{X}_{\tau \wedge \bar{\rho}^n}^n) \right) \mathbf{1}_{\{\bar{\rho}^n > u\}} \\
 &= Y(u, \tau, \bar{\xi}^n) \mathbf{1}_{\{\bar{\rho}^n > u\}}. \tag{III.8.36}
 \end{aligned}$$

Therefore,

$$\begin{aligned}
 Y(t, \tau, \bar{\xi}^n) &= \int_t^{u \wedge \bar{\rho}^n} f(s, \bar{X}_s^n) ds - \int_{[t, u \wedge \bar{\rho}^n)} \langle c_+(s), d\bar{\xi}_s^{n,+} \rangle - \int_{[t, u \wedge \bar{\rho}^n)} \langle c_-(s), d\bar{\xi}_s^{n,-} \rangle \\
 &\quad + g(\bar{\rho}^n, \bar{X}_{\bar{\rho}^n}^n) \mathbf{1}_{\{\bar{\rho}^n \leq u\}} + Y(u, \tau, \bar{\xi}^n) \mathbf{1}_{\{\bar{\rho}^n \geq u\}} \mathbf{1}_{\{\bar{\rho}^n > u\}}. \tag{III.8.37}
 \end{aligned}$$

Thus, from definitions (III.8.12) and (III.8.26) and by (III.8.37), we obtain

$$\begin{aligned}
 \bar{\Psi}_u^n &= \int_t^{u \wedge \bar{\rho}^n} f(s, \bar{X}_s^n) ds - \int_{[t, u \wedge \bar{\rho}^n)} \langle c_+(s), d\bar{\xi}_s^{n,+} \rangle - \int_{[t, u \wedge \bar{\rho}^n)} \langle c_-(s), d\bar{\xi}_s^{n,-} \rangle \\
 &\quad + g(\bar{\rho}^n, \bar{X}_{\bar{\rho}^n}^n) \mathbf{1}_{\{\bar{\rho}^n \leq u\}} + \bar{U}_u^n \mathbf{1}_{\{\bar{\rho}^n > u\}}, \quad \mathbb{P}\text{-a.s.} \tag{III.8.38}
 \end{aligned}$$

Recall that $\bar{S}^n$ is the Snell Envelope of $\bar{Y}^n$ and so it is a right-continuous modification of $\bar{\Psi}^n$. Moreover, by Lemma III.8.4, we have that $\bar{U}_u^n = \bar{V}_u^n$, $\mathbb{P}$ -a.s. and so we obtain

$$\begin{aligned}
 \bar{S}_u^n &= \bar{\Psi}_u^n \\
 &= \int_t^{u \wedge \bar{\rho}^n} f(s, \bar{X}_s^n) ds - \int_{[t, u \wedge \bar{\rho}^n)} \langle c_+(s), d\bar{\xi}_s^{n,+} \rangle - \int_{[t, u \wedge \bar{\rho}^n)} \langle c_-(s), d\bar{\xi}_s^{n,-} \rangle \\
 &\quad + g(\bar{\rho}^n, \bar{X}_{\bar{\rho}^n}^n) \mathbf{1}_{\{\bar{\rho}^n \leq u\}} + \bar{V}_u^n \mathbf{1}_{\{\bar{\rho}^n > u\}} \\
 &= \int_t^{u \wedge \bar{\rho}^n} f(s, \bar{X}_s^n) ds - \int_{[t, u \wedge \bar{\rho}^n)} \langle c_+(s), d\bar{\xi}_s^{n,+} \rangle - \int_{[t, u \wedge \bar{\rho}^n)} \langle c_-(s), d\bar{\xi}_s^{n,-} \rangle \\
 &\quad + g(\bar{\rho}^n, \bar{X}_{\bar{\rho}^n}^n) \mathbf{1}_{\{\bar{\rho}^n \leq u\}} + v_n(u, \bar{X}_u^n) \mathbf{1}_{\{\bar{\rho}^n > u\}}, \quad \mathbb{P}\text{-a.s.}, \tag{III.8.39}
 \end{aligned}$$

where the last equality follows from the version of Proposition III.6.1 for n^2 -Lipschitz controls. Since both the left-hand side and the right-hand side of (III.8.39) are right-continuous as a function of $u \in [t, T]$, they are indistinguishable. Hence, from (III.8.37) and (III.8.39), the stopping time

$\bar{\sigma}^n$ in (III.8.15) reads as

$$\begin{aligned}
 \bar{\sigma}^n &= \inf \{s \geq t : v_n(s, \bar{X}_s^n) \mathbf{1}_{\{\bar{\rho}^n > s\}} = Y(s, s, \bar{\xi}^n) \mathbf{1}_{\{\bar{\rho}^n > s\}}\} \\
 &= \inf \{s \geq t : v_n(s, \bar{X}_s^n) \mathbf{1}_{\{\bar{\rho}^n > s\}} = g(s, \bar{X}_s^n) \mathbf{1}_{\{\bar{\rho}^n > s\}}\} \\
 &= \inf \{s \geq t : v_n(s, \bar{X}_s^n) = g(s, \bar{X}_s^n) \text{ or } s \geq \bar{\rho}^n\} \\
 &= \bar{\tau}^n \wedge \bar{\rho}^n, \quad \mathbb{P}\text{-a.s.},
 \end{aligned} \tag{III.8.40}$$

where we have denoted $\tau^{\bar{\xi}^n, n}$ (defined as in (III.8.34)) simply by $\bar{\tau}^n$. Then, from (III.8.24), we have

$$\begin{aligned}
 v_n(t, x) &= \mathbb{E}[Y(t, \bar{\sigma}^n, \bar{\xi}^n)] \\
 &= \mathbb{E}[Y(t, \bar{\tau}^n \wedge \bar{\rho}^n, \bar{\xi}^n)] \\
 &= \mathbb{E}[Y(t, \bar{\tau}^n, \bar{\xi}^n)] \\
 &\leq \sup_{\xi \in \mathcal{A}_t^n} \mathbb{E}[Y(t, \tau^{\xi, n}, \xi)] \\
 &= \sup_{\xi \in \mathcal{A}_t^n} J_{t,x}(\tau^{\xi, n}, \xi),
 \end{aligned} \tag{III.8.41}$$

where the third equality follows from the definition of $\bar{Y}^n$ (see (III.8.11)) and the inequality follows from the fact that $\bar{\xi}^n \in \mathcal{A}_t^n$. Thus, (III.8.33) is proved when $\mathbb{P}(\bar{\sigma}^n \geq u) > 0$ for every $u \in [t, T]$.

If there exists $\hat{u} \in [t, T]$ such that $\mathbb{P}(\bar{\sigma}^n \geq \hat{u}) = 0$, i.e., $\mathbb{P}(\bar{\sigma}^n < \hat{u}) = 1$ then let

$$\hat{T} := \inf\{s \geq t : \mathbb{P}(\bar{\sigma}^n < s) = 1\} \in [t, \hat{u}]. \tag{III.8.42}$$

Therefore $\bar{\sigma}^n \leq \hat{T}$, $\mathbb{P}$ -a.s. and either

$$\mathbb{P}(\bar{\sigma}^n < \hat{T}) = 1 \quad \Rightarrow \quad \mathbb{P}(\bar{\sigma}^n \geq s) > 0, \quad \forall s \in [t, \hat{T}), \tag{III.8.43}$$

or

$$\mathbb{P}(\bar{\sigma}^n = \hat{T}) > 0 \quad \Rightarrow \quad \mathbb{P}(\bar{\sigma}^n \geq s) > 0, \quad \forall s \in [t, \hat{T}], \tag{III.8.44}$$

In the former case, we apply the same procedure as above restricted to the interval $[t, \hat{T})$ and we obtain that

$$\begin{aligned}
 \bar{\sigma}^n &= \inf\{s \in [t, \hat{T}) : v_n(s, \bar{X}_s^n) = g(s, \bar{X}_s^n)\} \wedge \bar{\rho}^n \\
 &= \bar{\tau}^n \wedge \bar{\rho}^n, \quad \mathbb{P}\text{-a.s.}
 \end{aligned} \tag{III.8.45}$$

In the latter case, we apply the same procedure as above restricted to the interval $[t, \hat{T}]$ and we obtain that

$$\begin{aligned}
 \bar{\sigma}^n &= \inf\{s \in [t, \hat{T}] : v_n(s, \bar{X}_s^n) = g(s, \bar{X}_s^n)\} \wedge \bar{\rho}^n \\
 &= \bar{\tau}^n \wedge \bar{\rho}^n, \quad \mathbb{P}\text{-a.s.}
 \end{aligned} \tag{III.8.46}$$

In both cases, we reach the thesis (III.8.33) by proceeding as in (III.8.41). $\square$

In order to obtain (III.8.1), we prove one last lemma. We show that if the value functions for the n^2 -Lipschitz controls converge to the value function for continuous controls, then the optimal stopping times for n^2 -Lipschitz controls converge to the analogous stopping time for continuous

controls.

Lemma III.8.6. *Assume that the functions $(s, y) \mapsto v^c(s, y)$ and $(s, y) \mapsto v_n(s, y)$ are continuous on $[t, T] \times \mathbb{R}^d$ for every $n \in \mathbb{N}$ and that*

$$\lim_{n \rightarrow \infty} v_n(t, x) = v^c(t, x), \quad \forall (t, x) \in [0, T] \times \mathbb{R}^d. \quad (\text{III.8.47})$$

Moreover, assume that Assumption III.8.1 (A) holds. Let $\xi \in \mathcal{A}_t^c$, then

$$\lim_{n \rightarrow \infty} \tau^{\xi, n} = \tau^{\xi, c}, \quad \mathbb{P}\text{-a.s.}, \quad (\text{III.8.48})$$

where $\tau^{\xi, n}$ is defined as in (III.8.34) and

$$\tau^{\xi, c} := \inf\{s \geq t : v^c(s, X_s^\xi) = g(s, X_s^\xi)\}. \quad (\text{III.8.49})$$

Proof. Notice that since $v_n \leq v_{n+1}$ for every $n \in \mathbb{N}$, then $(\tau^{\xi, n})_{n \in \mathbb{N}}$ is non-decreasing $\mathbb{P}$ -a.s. and thus it admits a limit $\mathbb{P}$ -a.s. We denote this limit by $\tau^{\xi, \infty}$. Since $v_n \leq v^c$ for every $n \in \mathbb{N}$, then $\tau^{\xi, n} \leq \tau^{\xi, c}$, $\mathbb{P}$ -a.s. for every $n \in \mathbb{N}$ and so also $\tau^{\xi, \infty} \leq \tau^{\xi, c}$, $\mathbb{P}$ -a.s. We now want to show the opposite inequality. Fix $\omega \in \Omega$. If $\tau^{\xi, c}(\omega) = t$, then we have nothing to prove. Thus, assume that $\tau^{\xi, c}(\omega) > t$ and consider any $u \in [t, \tau^{\xi, c}(\omega))$. By continuity of v^c , g and X^ξ , we have that

$$\inf_{s \in [t, u]} (v^c(s, X_s^\xi) - g(s, X_s^\xi))(\omega) > 0. \quad (\text{III.8.50})$$

Since $(v_n)_{n \in \mathbb{N}}$ is non-decreasing, $v_n \rightarrow v^c$ pointwise (by hypothesis (III.8.47)) and both $(v_n)_{n \in \mathbb{N}}$ and v^c are continuous then, by Dini's theorem on uniform convergence, we have that $v_n \rightarrow v^c$ uniformly on compact sets. Therefore, for n sufficiently large, we obtain

$$\inf_{s \in [t, u]} (v_n(s, X_s^\xi) - g(s, X_s^\xi))(\omega) > 0. \quad (\text{III.8.51})$$

Thus, $\tau^{\xi, n}(\omega) > u$. Since $u \in [t, \tau^{\xi, c}(\omega))$ was arbitrary we obtain that for every $\varepsilon > 0$ there exists $n_\varepsilon \in \mathbb{N}$ such that $\tau^{\xi, n}(\omega) > \tau^{\xi, c}(\omega) - \varepsilon$ for every $n \geq n_\varepsilon$. Letting $n \rightarrow \infty$ and then $\varepsilon \rightarrow 0$, we obtain the desired result. $\square$

We can now prove the optimal stopping rule (III.8.1).

Theorem III.8.7 (Optimal stopping rule). *Assume that the functions $(s, y) \mapsto v^c(s, y)$ and $(s, y) \mapsto v_n(s, y)$ are continuous on $[t, T] \times \mathbb{R}^d$ for every $n \in \mathbb{N}$ and that*

$$\lim_{n \rightarrow \infty} v_n(t, x) = v(t, x), \quad \forall (t, x) \in [0, T] \times \mathbb{R}^d. \quad (\text{III.8.52})$$

Moreover, assume that

$$\lim_{n \rightarrow \infty} \sup_{\xi \in \mathcal{A}_t^c} J_{t, x}(\tau^{\xi, n}, \xi) \leq \sup_{\xi \in \mathcal{A}_t^c} \lim_{n \rightarrow \infty} J_{t, x}(\tau^{\xi, n}, \xi), \quad (\text{III.8.53})$$

where $\tau^{\xi, n}$ is defined as in (III.8.34). Under the hypotheses of Proposition III.6.1 and of Theorem III.7.2 and under Assumption III.8.1, we have that

$$v(t, x) = \sup_{\xi \in \mathcal{A}_t} J_{t, x}(\tau^\xi, \xi), \quad (\text{III.8.54})$$

where

$$\tau^\xi := \inf\{s \geq t : v(s, X_s^\xi) = g(s, X_s^\xi)\}. \quad (\text{III.8.55})$$

Remark III.8.8. Assumption (III.8.53) is, in general, hard to prove. This is a delicate technical issue that have already been discussed in various papers (see, e.g., [10] and [39]). It can be interpreted as a sort of monotone/dominated convergence theorem for non-linear expectations. Indeed, the supremum over the controls can be viewed as a non-linear expectation $\mathcal{E}$. To justify (III.8.53) in [10], the authors assume that the process Y is “ $\mathcal{E}$ -uniformly left-continuous” (see, Definition 2.3 and Theorem 2.1 in [10]). However, this hypothesis cannot be easily verified in our case because the optimisation domain $\mathcal{A}_t^c$ is not uniformly bounded.

Proof. It is immediate to see that $v(t, x) \geq \sup_{\xi \in \mathcal{A}_t} J_{t,x}(\tau^\xi, \xi)$ as $\tau^\xi \in \mathcal{T}_t$ for every $\xi \in \mathcal{A}_t$. Thus, we only need to prove the opposite inequality.

We first show that

$$v^c(t, x) \leq \sup_{\xi \in \mathcal{A}_t^c} J_{t,x}(\tau^{\xi,c}, \xi), \quad (\text{III.8.56})$$

where $\tau^{\xi,c}$ is defined as in (III.8.49). Notice that since $v_n \leq v^c \leq v$ for every $n \in \mathbb{N}$ then, by hypothesis (III.8.52), we obtain $v_n \rightarrow v^c = v$. Thus, for every $\varepsilon > 0$ there exists $n_\varepsilon > 0$ such that $v^c(t, x) - \varepsilon \leq v_n(t, x)$ for every $n \geq n_\varepsilon$. Then, by Lemma III.8.5, we obtain

$$v^c(t, x) - \varepsilon \leq \sup_{\xi \in \mathcal{A}_t^n} J_{t,x}(\tau^{\xi,n}, \xi) \leq \sup_{\xi \in \mathcal{A}_t^c} J_{t,x}(\tau^{\xi,n}, \xi), \quad \forall n \geq n_\varepsilon. \quad (\text{III.8.57})$$

Letting $n \rightarrow \infty$, using assumption (III.8.53), the dominated convergence theorem and Lemma III.8.6, we have that

$$v^c(t, x) - \varepsilon \leq \sup_{\xi \in \mathcal{A}_t^c} J_{t,x}(\tau^{\xi,c}, \xi). \quad (\text{III.8.58})$$

Finally, letting $\varepsilon \rightarrow 0$, we obtain the first desired result (III.8.56).

To conclude the proof, we need to show that (III.8.54) holds. As $v^c = v$, we have

$$v(t, x) = v^c(t, x) \leq \sup_{\xi \in \mathcal{A}_t^c} J_{t,x}(\tau^{\xi,c}, \xi). \quad (\text{III.8.59})$$

Now define

$$A := \{(t, x) \in [0, T] \times \mathbb{R}^d : v(t, x) - g(t, x) = 0\}, \quad (\text{III.8.60})$$

and

$$A^c := \{(t, x) \in [t, T] \times \mathbb{R}^d : v^c(t, x) - g(t, x) = 0\}. \quad (\text{III.8.61})$$

Again, since $v^c = v$, then $A = A^c$ and so, for every $\xi \in \mathcal{A}_t^c$,

$$\tau^{\xi,c} = \inf\{s \geq t : (s, X_s^\xi) \in A^c\} = \inf\{s \geq t : (s, X_s^\xi) \in A\} = \tau^\xi, \quad \mathbb{P}\text{-a.s.} \quad (\text{III.8.62})$$

Hence, from (III.8.59), we obtain

$$v(t, x) \leq \sup_{\xi \in \mathcal{A}_t^c} J_{t,x}(\tau^\xi, \xi) \leq \sup_{\xi \in \mathcal{A}_t} J_{t,x}(\tau^\xi, \xi), \quad (\text{III.8.63})$$

and the proof is complete. $\square$

III.9 Regularity of the value function

In this section we show some regularity properties of the value function. In particular, we prove that, under some assumptions, the value function satisfies the hypotheses required for the results of the previous sections.

Recall that, by definition, in our optimisation problem (III.3.3)-(III.3.4) the underlying reference probability system $\nu \in \mathcal{V}_t$ depends on the initial time $t \in [0, T]$. The first result of this section shows that we can fix a common reference probability system ν_0 that is suitable for every initial time $t \in [0, T]$. To prove this, let $\nu_0 := (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t \in [0, T]}, W) \in \mathcal{V}_0$ be a standard reference probability system starting at time 0, e.g., the canonical reference probability system. Thus, W is a Brownian motion starting at time 0 and $\mathcal{F}_t$ is the $\mathbb{P}$ -augmentation of $\mathcal{F}_t^0 := \sigma(W_s : s \in [0, t])$. Then, for every $(t, x) \in [0, T] \times \mathbb{R}^d$, we define

$$v_0(t, x) := \sup_{(\tau, \xi) \in \mathcal{T}_t^{\nu_0} \times \mathcal{A}_t^{\nu_0}} J_{t,x}^{\nu_0}(\tau, \xi). \quad (\text{III.9.1})$$

Proposition III.9.1. *For every $(t, x) \in [0, T] \times \mathbb{R}^d$, we have $v(t, x) = v_0(t, x)$.*

Proof. Let $(t, x) \in [0, T] \times \mathbb{R}^d$ and define $\nu_t := (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_s^t\}_{s \in [t, T]}, W^t)$, where $W_s^t := W_s - W_t$ for every $s \in [t, T]$ and $\mathcal{F}_s^t$ is the $\mathbb{P}$ -augmentation of $\mathcal{F}_s^{t,0} := \sigma(W_r^t : r \in [t, s])$. Then, $\nu_t \in \mathcal{V}_t$ and so, by Proposition III.4.5, we have

$$v(t, x) = \sup_{(\tau, \xi) \in \mathcal{T}_t^{\nu_t} \times \mathcal{A}_t^{\nu_t}} J_{t,x}^{\nu_t}(\tau, \xi). \quad (\text{III.9.2})$$

It is immediate to see that $v(t, x) \leq v_0(t, x)$ as $\mathcal{F}_s^t \subseteq \mathcal{F}_s$ for every $s \in [t, T]$. To show the opposite inequality, we proceed as in the proof of Proposition III.6.1 by using the regular conditional probabilities. Let $(\tau, \xi) \in \mathcal{T}_t^{\nu_0} \times \mathcal{A}_t^{\nu_0}$ and, by the same argument as in (III.4.2), let $\eta \in \mathcal{E}_t^{\nu_0}$ such that $J_{t,x}^{\nu_0}(\tau, \xi) = I_{t,x}^{\nu_0}(\eta, \xi)$. By Corollary III.4.2, let η^0 and ξ^0 be the $\{\mathcal{F}_t^0\}$ -adapted controls indistinguishable from η and ξ , respectively. Moreover, let $X := X^{t,x;\xi,\nu^0}$ and, by Lemma III.4.1, let X^0 be the $\{\mathcal{F}_t^0\}$ -adapted process which is indistinguishable from $X^{t,x;\xi^0,\nu^0}$ and let ρ^0 be its corresponding exit time from the domain $\mathcal{O}$. By Proposition III.5.4, we have that

$$\nu_\omega := (\Omega, \mathcal{F}_\omega, \mathbb{P}_\omega, \{\mathcal{F}_s^{t,\omega}\}_{s \in [t, T]}, W^t) \in \mathcal{V}_t, \quad \eta^0 \in \mathcal{E}_t^{\nu_\omega}, \quad \xi^0 \in \mathcal{A}_t^{\nu_\omega}, \quad \mathbb{P}\text{-a.e. } \omega, \quad (\text{III.9.3})$$

where $\mathbb{P}_\omega = \mathbb{P}(\cdot | \mathcal{F}_t^0)(\omega)$ is the regular conditional probability given $\mathcal{F}_t^0$ and $\mathcal{F}_s^{t,\omega}$ is the $\mathbb{P}_\omega$ -augmentation of $\mathcal{F}_s^{t,0}$. Finally, for $\mathbb{P}$ -a.e. ω , let $X^\omega := X^{t,x;\xi^0,\nu_\omega}$ and ρ^ω its corresponding exit time from the domain $\mathcal{O}$. Then, by an argument as in (III.6.15), we obtain that

$$\begin{aligned} J_{t,x}^{\nu_0}(\tau, \xi) &= I_{t,x}^{\nu_0}(\eta, \xi) \\ &= \mathbb{E}[Z(t, \eta, \xi, X, \rho)] \\ &= \mathbb{E}[\mathbb{E}[Z(t, \eta^0, \xi^0, X^0, \rho^0) | \mathcal{F}_t^0](\omega)] \\ &= \mathbb{E}[\mathbb{E}_\omega[Z(t, \eta^0, \xi^0, X^\omega, \rho^\omega)]] \\ &= \mathbb{E}[I_{t,x}^{\nu_\omega}(\eta^0, \xi^0)] \\ &\leq \mathbb{E}[v(t, x)] \\ &= v(t, x), \end{aligned} \quad (\text{III.9.4})$$

where the inequality follows from (III.9.3) and Proposition III.4.5. Taking the supremum over $(\tau, \xi) \in \mathcal{T}_t^{\nu_0} \times \mathcal{A}_t^{\nu_0}$ in (III.9.4), we obtain that $v_0(t, x) \leq v(t, x)$ and the theorem is proved. $\square$

While the formulation of the value function as in $v(t, x)$ is tailored to obtain the dynamic programming principle, the formulation as in $v_0(t, x)$ is more suitable to prove regularity properties, e.g., the continuity of the value function. Indeed, by introducing $v_0(t, x)$ we remove the dependence of the underlying reference probability system on the initial time t . Thus, in what follows we refer to the formulation $v_0(t, x)$ fixing the common probability system ν_0 and avoiding any explicit reference to it. For instance, we denote $\mathcal{A}_t^{\nu_0}$ by $\mathcal{A}_t$ also in accordance with the previous section.

In this section, we assume that the functions $(s, y) \mapsto \mu(s, y)$, $(s, y) \mapsto \sigma(s, y)$ are Borel-measurable and that there exist $L > 0$, $M > 0$ such that for every $y, z \in \mathbb{R}^d$ and $s \in [t, T]$,

$$|\mu(s, y)| \leq M(1 + |y|), \quad |\sigma(s, y)| \leq M(1 + |y|), \quad (\text{III.9.5})$$

$$|\mu(s, y) - \mu(s, z)| \leq L|y - z|, \quad |\sigma(s, y) - \sigma(s, z)| \leq L|y - z|. \quad (\text{III.9.6})$$

Under these conditions, the SDE (III.3.1) admits a unique (up to indistinguishability) $\{\mathcal{F}_s^t\}$ -adapted solution which is $\mathbb{P}$ -a.s. left continuous with finite right limits. This follows by defining, for $s \in [t, T]$, the Picard iterations

$$\begin{aligned} X_s^0 &:= x + \xi_s, \\ X_s^n &:= x + \int_t^s \mu(u, X_u^{n-1}) du + \int_t^s \sigma(u, X_u^{n-1}) dW_u + \xi_s, \quad n \geq 1, \end{aligned} \quad (\text{III.9.7})$$

and applying the usual procedure (see, e.g., [6, Theorem 9.2]).

We now provide some standard estimates for the SDE (III.3.1) that will be useful to prove the subsequent results.

Proposition III.9.2. *There exists $C_1 = C_1(M, T) > 0$ such that, for every $u \in [t, T]$ and $\xi \in \mathcal{A}_t$,*

$$\mathbb{E} \left[\sup_{s \in [t, u]} |X_s^{t, x; \xi}|^2 \right] \leq C_1 \left(1 + |x|^2 + \mathbb{E} \left[\sup_{s \in [t, u]} |\xi_s|^2 \right] \right), \quad (\text{III.9.8})$$

$$\mathbb{E} \left[\sup_{s \in [t, u]} |X_s^{t, x; \xi} - x|^2 \right] \leq C_1 \left((u - t)(1 + |x|^2) + \mathbb{E} \left[\sup_{s \in [t, u]} |\xi_s|^2 \right] \right). \quad (\text{III.9.9})$$

Moreover, there exists $C_2 = C_2(L, M, T) > 0$ such that, for every $(t_1, x_1) \in [0, T] \times \mathbb{R}^d$, $(t_2, x_2) \in [t_1, T] \times \mathbb{R}^d$, $\xi^1 \in \mathcal{A}_{t_1}$ and $\xi^2 \in \mathcal{A}_{t_2}$,

$$\begin{aligned} &\mathbb{E} \left[\sup_{s \in [t_2, T]} |X_s^{t_1, x_1; \xi^1} - X_s^{t_2, x_2; \xi^2}|^2 \right] \\ &\leq C_2 (1 + |x_1|^2) \left(t_2 - t_1 + |x_1 - x_2|^2 + \mathbb{E} \left[\sup_{s \in [t_2, T]} |\xi_s^1 - \xi_s^2|^2 \right] \right). \end{aligned} \quad (\text{III.9.10})$$

Proof. The proof follows from assumptions (III.9.5)-(III.9.6) by standard techniques for the estimates of SDEs (see, e.g., Theorem 9.1 and Proposition 9.1 in [6]). $\square$

Notice that if $\xi \in \mathcal{A}_t$, then

$$\mathbb{E} \left[\sup_{s \in [t, T]} |\xi_s|^2 \right] \leq 2 \left(\mathbb{E} [|\xi_T^+|^2] + \mathbb{E} [|\xi_T^-|^2] \right) < \infty. \quad (\text{III.9.11})$$

Remark III.9.3. In this section, we consider no exit from the domain, i.e., we assume $\mathcal{O} = \mathbb{R}^d$ which implies $\rho = T$. Thus, for $(\tau, \xi) \in \mathcal{T}_t \times \mathcal{A}_t$, the gain functional reads as

$$J_{t,x}(\tau, \xi) = \mathbb{E} \left[\int_t^\tau f(s, X_s^\xi) ds - \int_{[t,\tau)} \langle c_+(s), d\xi_s^+ \rangle - \int_{[t,\tau)} \langle c_-(s), d\xi_s^- \rangle + g(\tau, X_\tau^\xi) \right], \quad (\text{III.9.12})$$

where, for the sake of simplicity, we have also assumed that $g_1(T, \cdot) = g_2(T, \cdot)$ and we have simply denoted g_2 by g . We make this choice because, in the general setting, it is challenging to prove the continuity of the map $(t, x) \mapsto \rho^{t,x}$, which would be needed to obtain the regularity properties we provide. However, this continuity can be obtained at least in some specific cases. For instance, if $\mathcal{O} \subseteq \mathbb{R}^d$ is an open set, $\bar{\mathcal{O}}$ denotes the closure of $\mathcal{O}$ and

$$\rho = \inf \{s \geq t : X_s \notin \bar{\mathcal{O}}\}, \quad (\text{III.9.13})$$

one can obtain this continuity by an argument as in, e.g., [29, Proposition 5.2]. In the dividend problem [29], equality (III.9.13) holds as in their setting $\mathcal{O} = (0, \infty)$, the controlled diffusion X is a controlled Brownian motion with drift and the controls (the paid dividends) can only push X down. Thus, when X hits 0 (and thus exits from $\mathcal{O}$) it immediately exits from $\bar{\mathcal{O}} = [0, \infty)$ by the law of iterated logarithm, as the controller cannot push it up.

In this section we will make use of the following assumptions.

Assumption III.9.4.

(A) There exist $\alpha \in (0, 1]$ and $L_f, L_g > 0$ such that

$$|f(t, x) - f(t, y)| \leq L_f |x - y|^\alpha, \quad t \in [0, T], \quad x, y \in \mathbb{R}^d, \quad (\text{III.9.14})$$

$$|g(t, x) - g(t, y)| \leq L_g |x - y|^\alpha, \quad t \in [0, T], \quad x, y \in \mathbb{R}^d. \quad (\text{III.9.15})$$

(B) There exist $M_f, M_g > 0$ such that

$$|f(t, x)| \leq M_f(1 + |x|), \quad t \in [0, T], \quad x \in \mathbb{R}^d, \quad (\text{III.9.16})$$

$$|g(t, x)| \leq M_g(1 + |x|), \quad t \in [0, T], \quad x \in \mathbb{R}^d. \quad (\text{III.9.17})$$

(C) There exists $M_c > 0$ such that

$$\sup_{t \in [0, T]} |c_\pm(t)| \leq M_c. \quad (\text{III.9.18})$$

We now show that, under Assumption III.9.4 (A), the gain functional satisfies the hypothesis of Proposition III.6.1.

Proposition III.9.5. Under Assumption III.9.4 (A), for every compact set $K \subset \mathbb{R}^d$ there exists $C_K > 0$ such that, for every $t \in [0, T]$ and $(\tau, \xi) \in \mathcal{T}_t \times \mathcal{A}_t$, we have

$$|J_{t,x}(\tau, \xi) - J_{t,y}(\tau, \xi)| \leq C_K |x - y|^\alpha, \quad \forall x, y \in K. \quad (\text{III.9.19})$$

Proof. Let $t \in [0, T]$, $(\tau, \xi) \in \mathcal{T}_t \times \mathcal{A}_t$ and $x, y \in K$. We have

$$\begin{aligned} |J_{t,x}(\tau, \xi) - J_{t,y}(\tau, \xi)| &\leq \mathbb{E} \left[\int_t^\tau |f(s, X_s^{t,x}) - f(s, X_s^{t,y})| ds + |g(\tau, X_\tau^{t,x}) - g(\tau, X_\tau^{t,y})| \right] \\ &\leq L_f \mathbb{E} \left[\int_t^\tau |X_s^{t,x} - X_s^{t,y}|^\alpha ds \right] + L_g \mathbb{E} \left[|X_\tau^{t,x} - X_\tau^{t,y}|^\alpha \right] \\ &\leq 2(L_f T \vee L_g) \mathbb{E} \left[\sup_{s \in [t, T]} |X_s^{t,x} - X_s^{t,y}|^\alpha \right]. \end{aligned} \quad (\text{III.9.20})$$

Then, by Jensen's inequality and the estimate (III.9.10), we have that

$$\begin{aligned} \mathbb{E} \left[\sup_{s \in [t, T]} |X_s^{t,x} - X_s^{t,y}|^\alpha \right] &\leq \mathbb{E} \left[\sup_{s \in [t, T]} |X_s^{t,x} - X_s^{t,y}|^2 \right]^{\frac{\alpha}{2}} \\ &\leq C_2^{\frac{\alpha}{2}} (1 + |x|^2)^{\frac{\alpha}{2}} |x - y|^\alpha \\ &\leq C_2^{\frac{\alpha}{2}} \left(1 + \max_{x \in K} |x|^2 \right)^{\frac{\alpha}{2}} |x - y|^\alpha, \end{aligned} \quad (\text{III.9.21})$$

which completes the proof. $\square$

As a corollary of Proposition III.9.5, we also obtain that the function $x \mapsto v(t, x)$ is locally α -Hölder continuous uniformly in $t \in [0, T]$.

Corollary III.9.6. *Under Assumption III.9.4 (A), for every compact set $K \subset \mathbb{R}^d$ there exists $C_K > 0$ such that, for every $t \in [0, T]$, we have*

$$|v(t, x) - v(t, y)| \leq C_K |x - y|^\alpha, \quad \forall x, y \in K. \quad (\text{III.9.22})$$

Proof. From (III.9.19) there exists $C_K > 0$ such that, for every $t \in [0, T]$, $(\tau, \xi) \in \mathcal{T}_t \times \mathcal{A}_t$ and $x, y \in K$, we have that

$$J_{t,x}(\tau, \xi) \leq J_{t,y}(\tau, \xi) + C_K |x - y|^\alpha, \quad (\text{III.9.23})$$

$$J_{t,y}(\tau, \xi) \leq J_{t,x}(\tau, \xi) + C_K |x - y|^\alpha. \quad (\text{III.9.24})$$

Thus, taking the supremum over $(\tau, \xi) \in \mathcal{T}_t \times \mathcal{A}_t$, we obtain

$$v(t, x) \leq v(t, y) + C_K |x - y|^\alpha, \quad (\text{III.9.25})$$

$$v(t, y) \leq v(t, x) + C_K |x - y|^\alpha, \quad (\text{III.9.26})$$

which leads to the desired result. $\square$

Notice that the regularity proved in (III.9.22) for the value function v holds, a fortiori, for the value function for continuous controls v^c . Thus, to obtain the continuity of $(t, x) \mapsto v^c(t, x)$ (required in Theorem III.8.7), we only need to prove that $t \mapsto v^c(t, x)$ is continuous for every $x \in \mathbb{R}^d$.

Proposition III.9.7. *If Assumption III.9.4 hold and the function $t \mapsto g(t, x)$ is either continuous or non-decreasing on $[0, T]$ for every $x \in \mathbb{R}^d$, then the function $t \mapsto v^c(t, x)$ is continuous on $[0, T]$ for every $x \in \mathbb{R}^d$.*

Proof. Let $0 \leq t_1 < t_2 \leq T$, $x \in \mathbb{R}^d$ and $\varepsilon > 0$. Let $(\tau^\varepsilon, \xi^\varepsilon) \in \mathcal{T}_{t_2} \times \mathcal{A}_{t_2}^c$ be ε -optimal for $v(t_2, x)$. Then, $\tau^\varepsilon \geq t_2 > t_1$ and $\xi_s^\varepsilon = 0$ for every $s \in [0, t_1]$. Hence, $(\tau^\varepsilon, \xi^\varepsilon) \in \mathcal{T}_{t_1} \times \mathcal{A}_{t_1}^c$ and thus we have

$$v^c(t_2, x) - v^c(t_1, x) \leq \varepsilon + J_{t_2, x}(\tau^\varepsilon, \xi^\varepsilon) - J_{t_1, x}(\tau^\varepsilon, \xi^\varepsilon). \quad (\text{III.9.27})$$

Denoting $X^{t_i, x; \xi^\varepsilon}$ by X^i for $i = 1, 2$, we obtain

$$J_{t_2, x}(\tau^\varepsilon, \xi^\varepsilon) - J_{t_1, x}(\tau^\varepsilon, \xi^\varepsilon) = \mathbb{E} \left[\int_{t_2}^{\tau^\varepsilon} f(s, X_s^2) ds - \int_{t_1}^{\tau^\varepsilon} f(s, X_s^1) ds \right] \quad (\text{III.9.28})$$

$$- \mathbb{E} \left[\int_{[t_2, \tau^\varepsilon)} \langle c_+(s), d\xi_s^{\varepsilon, +} \rangle - \int_{[t_1, \tau^\varepsilon)} \langle c_+(s), d\xi_s^{\varepsilon, +} \rangle \right] \quad (\text{III.9.29})$$

$$- \mathbb{E} \left[\int_{[t_2, \tau^\varepsilon)} \langle c_-(s), d\xi_s^{\varepsilon, -} \rangle - \int_{[t_1, \tau^\varepsilon)} \langle c_-(s), d\xi_s^{\varepsilon, -} \rangle \right] \quad (\text{III.9.30})$$

$$+ \mathbb{E} \left[g(\tau^\varepsilon, X_{\tau^\varepsilon}^2) - g(\tau^\varepsilon, X_{\tau^\varepsilon}^1) \right]. \quad (\text{III.9.31})$$

For the term on the right-hand side of (III.9.28), we have

$$\begin{aligned} & \mathbb{E} \left[\int_{t_2}^{\tau^\varepsilon} f(s, X_s^2) ds - \int_{t_1}^{\tau^\varepsilon} f(s, X_s^1) ds \right] \\ & \leq \mathbb{E} \left[\int_{t_2}^T |f(s, X_s^2) - f(s, X_s^1)| ds \right] + \mathbb{E} \left[\int_{t_1}^{t_2} |f(s, X_s^1)| ds \right] \\ & \leq L_f T \mathbb{E} \left[\sup_{s \in [t_2, T]} |X_s^2 - X_s^1|^2 \right]^{\frac{\alpha}{2}} + M_f(t_2 - t_1) \left(1 + \mathbb{E} \left[\sup_{s \in [t_1, t_2]} |X_s^1|^2 \right]^{\frac{1}{2}} \right) \\ & \leq L_f T C_2^{\frac{\alpha}{2}} (1 + |x|^2)^{\frac{\alpha}{2}} (t_2 - t_1)^{\frac{\alpha}{2}} + M_f(t_2 - t_1) \\ & \quad + M_f C_1^{\frac{1}{2}} \left(1 + |x|^2 + \mathbb{E} \left[\sup_{s \in [t_1, t_2]} |\xi_s^\varepsilon|^2 \right] \right)^{\frac{1}{2}} (t_2 - t_1), \end{aligned} \quad (\text{III.9.32})$$

where we have used Assumption III.9.4, Jensen's inequality and the estimates (III.9.8)-(III.9.10). Notice that the terms in (III.9.29)-(III.9.30) are null as $\xi_s^{\varepsilon, \pm} = 0$ for every $s \in [t_1, t_2]$. For the term in (III.9.31), we have

$$\begin{aligned} \mathbb{E} \left[g(\tau^\varepsilon, X_{\tau^\varepsilon}^2) - g(\tau^\varepsilon, X_{\tau^\varepsilon}^1) \right] & \leq L_g \mathbb{E} \left[\sup_{s \in [t_2, T]} |X_s^2 - X_s^1|^2 \right]^{\frac{\alpha}{2}} \\ & \leq L_g C_2^{\frac{\alpha}{2}} (1 + |x|^2)^{\frac{\alpha}{2}} (t_2 - t_1)^{\frac{\alpha}{2}}, \end{aligned} \quad (\text{III.9.33})$$

where we have used Assumption III.9.4, Jensen's inequality and the estimate (III.9.10). Hence,

$$\begin{aligned} v^c(t_2, x) - v^c(t_1, x) & \leq \varepsilon + \tilde{C}_1 (1 + |x|^2)^{\frac{\alpha}{2}} (t_2 - t_1)^{\frac{\alpha}{2}} + M_f(t_2 - t_2) \\ & \quad + \tilde{C}_2 \left(1 + |x|^2 + \mathbb{E} \left[\sup_{s \in [t_1, t_2]} |\xi_s^\varepsilon|^2 \right] \right)^{\frac{1}{2}} (t_2 - t_1), \end{aligned} \quad (\text{III.9.34})$$

where $\tilde{C}_1, \tilde{C}_2 > 0$.

Now let $(\tau^\varepsilon, \xi^\varepsilon) \in \mathcal{T}_{t_1} \times \mathcal{A}_{t_1}^c$ be ε -optimal for $v(t_1, x)$. Define

$$\hat{\tau}^\varepsilon := \tau^\varepsilon \vee t_2, \quad \hat{\xi}_s^{\varepsilon, \pm} := (\xi_s^{\varepsilon, \pm} - \xi_{t_2}^{\varepsilon, \pm}) \mathbf{1}_{[t_2, T]}(s), \quad s \in [0, T], \quad (\text{III.9.35})$$

so that $(\hat{\tau}^\varepsilon, \hat{\xi}^\varepsilon) \in \mathcal{T}_{t_2} \times \mathcal{A}_{t_2}^c$ where $\hat{\xi}^\varepsilon := \hat{\xi}^{\varepsilon,+} - \hat{\xi}^{\varepsilon,-}$. Hence,

$$v^c(t_2, x) - v^c(t_1, x) \geq J_{t_2, x}(\hat{\tau}^\varepsilon, \hat{\xi}^\varepsilon) - J_{t_1, x}(\tau^\varepsilon, \xi^\varepsilon) - \varepsilon. \quad (\text{III.9.36})$$

Notice that $d\xi_s^{\varepsilon, \pm} = d\xi_s^{\varepsilon, \pm}$ on $[t_2, T]$ and so also $X^{t_2, x; \hat{\xi}^\varepsilon} = X^{t_2, x; \xi^\varepsilon}$. As above, we denote $X^{t_i, x; \xi^\varepsilon}$ by X^i for $i = 1, 2$ and thus we have

$$J_{t_2, x}(\hat{\tau}^\varepsilon, \hat{\xi}^\varepsilon) - J_{t_1, x}(\tau^\varepsilon, \xi^\varepsilon) = \mathbb{E} \left[\int_{t_2}^{\hat{\tau}^\varepsilon} f(s, X_s^2) ds - \int_{t_1}^{\tau^\varepsilon} f(s, X_s^1) ds \right] \quad (\text{III.9.37})$$

$$- \mathbb{E} \left[\int_{[t_2, \hat{\tau}^\varepsilon)} \langle c_+(s), d\xi_s^{\varepsilon, +} \rangle - \int_{[t_1, \tau^\varepsilon)} \langle c_+(s), d\xi_s^{\varepsilon, +} \rangle \right] \quad (\text{III.9.38})$$

$$- \mathbb{E} \left[\int_{[t_2, \hat{\tau}^\varepsilon)} \langle c_-(s), d\xi_s^{\varepsilon, -} \rangle - \int_{[t_1, \tau^\varepsilon)} \langle c_-(s), d\xi_s^{\varepsilon, -} \rangle \right] \quad (\text{III.9.39})$$

$$+ \mathbb{E} \left[g(\hat{\tau}^\varepsilon, X_{\hat{\tau}^\varepsilon}^2) - g(\tau^\varepsilon, X_{\tau^\varepsilon}^1) \right]. \quad (\text{III.9.40})$$

For the term on the right-hand side of (III.9.37), we have

$$\begin{aligned} & -\mathbb{E} \left[\int_{t_2}^{\hat{\tau}^\varepsilon} f(s, X_s^2) ds - \int_{t_1}^{\tau^\varepsilon} f(s, X_s^1) ds \right] \\ & \leq \mathbb{E} \left[\left| \int_{t_2}^{\tau^\varepsilon} f(s, X_s^2) ds - \int_{t_1}^{\tau^\varepsilon} f(s, X_s^1) ds \right| \mathbf{1}_{\{\tau^\varepsilon \geq t_2\}} + \left| \int_{t_1}^{\tau^\varepsilon} f(s, X_s^1) ds \right| \mathbf{1}_{\{\tau^\varepsilon < t_2\}} \right] \\ & \leq \mathbb{E} \left[\int_{t_2}^T |f(s, X_s^2) - f(s, X_s^1)| ds \right] + \mathbb{E} \left[\int_{t_1}^{t_2} |f(s, X_s^1)| ds \right] \\ & \leq L_f T \mathbb{E} \left[\sup_{s \in [t_2, T]} |X_s^2 - X_s^1|^2 \right]^{\frac{\alpha}{2}} + M_f(t_2 - t_1) \left(1 + \mathbb{E} \left[\sup_{s \in [t_1, t_2]} |X_s^1|^2 \right]^{\frac{1}{2}} \right) \\ & \leq L_f T C_2^{\frac{\alpha}{2}} (1 + |x|^2)^{\frac{\alpha}{2}} (t_2 - t_1)^{\frac{\alpha}{2}} + M_f(t_2 - t_1) \\ & \quad + M_f C_1^{\frac{1}{2}} \left(1 + |x|^2 + \mathbb{E} \left[\sup_{s \in [t_1, t_2]} |\xi_s^\varepsilon|^2 \right] \right)^{\frac{1}{2}} (t_2 - t_1), \end{aligned} \quad (\text{III.9.41})$$

where we have used Assumption III.9.4, Jensen's inequality and the estimates (III.9.8)-(III.9.10).

For the terms in (III.9.38)-(III.9.39), we obtain

$$\begin{aligned} & \mathbb{E} \left[\int_{[t_2, \hat{\tau}^\varepsilon)} \langle c_\pm(s), d\xi_s^{\varepsilon, \pm} \rangle - \int_{[t_1, \tau^\varepsilon)} \langle c_\pm(s), d\xi_s^{\varepsilon, \pm} \rangle \right] \\ & \leq \mathbb{E} \left[\left| \int_{[t_2, \tau^\varepsilon)} \langle c_\pm(s), d\xi_s^{\varepsilon, \pm} \rangle - \int_{[t_1, \tau^\varepsilon)} \langle c_\pm(s), d\xi_s^{\varepsilon, \pm} \rangle \right| \mathbf{1}_{\{\tau^\varepsilon \geq t_2\}} \right. \\ & \quad \left. + \left| \int_{[t_1, \tau^\varepsilon)} \langle c_\pm(s), d\xi_s^{\varepsilon, \pm} \rangle \right| \mathbf{1}_{\{\tau^\varepsilon < t_2\}} \right] \\ & \leq \mathbb{E} \left[\int_{[t_1, t_2)} |c_\pm(s)| |d\xi_s^{\varepsilon, \pm}| \right] \\ & \leq M_c \mathbb{E} \left[|\xi_{t_2}^{\varepsilon, \pm} - \xi_{t_1}^{\varepsilon, \pm}|^2 \right]^{\frac{1}{2}}, \end{aligned} \quad (\text{III.9.42})$$

where we have used Assumption III.9.4 and Jensen's inequality. For the term in (III.9.40), we have

$$\begin{aligned}
 & -\mathbb{E}\left[g(\hat{\tau}^\varepsilon, X_{\hat{\tau}^\varepsilon}^2) - g(\tau^\varepsilon, X_{\tau^\varepsilon}^1)\right] \\
 &= \mathbb{E}\left[\left(g(\tau^\varepsilon, X_{\tau^\varepsilon}^1) - g(\tau^\varepsilon, X_{\tau^\varepsilon}^2)\right)\mathbf{1}_{\{\tau^\varepsilon \geq t_2\}} + \left(g(\tau^\varepsilon, X_{\tau^\varepsilon}^1) - g(t_2, x)\right)\mathbf{1}_{\{\tau^\varepsilon < t_2\}}\right] \\
 &\leq \mathbb{E}\left[\left|g(\tau^\varepsilon, X_{\tau^\varepsilon}^1) - g(\tau^\varepsilon, X_{\tau^\varepsilon}^2)\right|\mathbf{1}_{\{\tau^\varepsilon \geq t_2\}}\right] + \mathbb{E}\left[\left|g(\tau^\varepsilon, X_{\tau^\varepsilon}^1) - g(t_2, x)\right|\mathbf{1}_{\{\tau^\varepsilon < t_2\}}\right] \\
 &\quad + \mathbb{E}\left[\left|g(\tau^\varepsilon, x) - g(t_2, x)\right|\mathbf{1}_{\{\tau^\varepsilon < t_2\}}\right]. \tag{III.9.43}
 \end{aligned}$$

If $t \mapsto g(t, x)$ is non-decreasing, then $\mathbb{E}[(g(\tau^\varepsilon, x) - g(t_2, x))\mathbf{1}_{\{\tau^\varepsilon < t_2\}}] \leq 0$, otherwise,

$$\begin{aligned}
 & -\mathbb{E}\left[g(\hat{\tau}^\varepsilon, X_{\hat{\tau}^\varepsilon}^2) - g(\tau^\varepsilon, X_{\tau^\varepsilon}^1)\right] \\
 &\leq L_g \mathbb{E}\left[\sup_{s \in [t_2, T]} |X_s^1 - X_s^2|^2\right]^{\frac{\alpha}{2}} + L_g \mathbb{E}\left[\sup_{s \in [t_1, t_2]} |X_s^1 - x|^2\right]^{\frac{\alpha}{2}} \\
 &\quad + \mathbb{E}\left[\sup_{s \in [t_1, t_2]} |g(s, x) - g(t_2, x)|\right] \\
 &\leq L_g C_2^{\frac{\alpha}{2}} (1 + |x|^2)^{\frac{\alpha}{2}} (t_2 - t_1)^{\frac{\alpha}{2}} + \sup_{s \in [t_1, t_2]} |g(s, x) - g(t_2, x)| \\
 &\quad + L_g C_1^{\frac{\alpha}{2}} \left((t_2 - t_1)(1 + |x|^2) + \mathbb{E}\left[\sup_{s \in [t_1, t_2]} |\xi_s^\varepsilon|^2\right]\right)^{\frac{\alpha}{2}}, \tag{III.9.44}
 \end{aligned}$$

where we have used Assumption III.9.4, Jensen's inequality and estimates (III.9.9)-(III.9.10). Hence,

$$\begin{aligned}
 v^c(t_2, x) - v^c(t_1, x) &\geq -\varepsilon + \hat{C}_1 (1 + |x|^2)^{\frac{\alpha}{2}} (t_2 - t_1)^{\frac{\alpha}{2}} - M_f (t_2 - t_1) \\
 &\quad + \hat{C}_2 \left(1 + |x|^2 + \mathbb{E}\left[\sup_{s \in [t_1, t_2]} |\xi_s^\varepsilon|^2\right]\right)^{\frac{1}{2}} (t_2 - t_1) \\
 &\quad - M_c \left(\mathbb{E}\left[|\xi_{t_2}^{\varepsilon,+} - \xi_{t_1}^{\varepsilon,+}|^2\right]^{\frac{1}{2}} + \mathbb{E}\left[|\xi_{t_2}^{\varepsilon,-} - \xi_{t_1}^{\varepsilon,-}|^2\right]^{\frac{1}{2}}\right) \\
 &\quad + \hat{C}_3 \left((t_2 - t_1)(1 + |x|^2) + \mathbb{E}\left[\sup_{s \in [t_1, t_2]} |\xi_s^\varepsilon|^2\right]\right)^{\frac{\alpha}{2}} \\
 &\quad - \sup_{s \in [t_1, t_2]} |g(s, x) - g(t_2, x)|, \tag{III.9.45}
 \end{aligned}$$

where $\hat{C}_1, \hat{C}_2, \hat{C}_3 < 0$. Therefore, by letting $t_1 \rightarrow t_2$ (or equivalently $t_2 \rightarrow t_1$) in (III.9.34) and (III.9.45), using the dominated convergence theorem and the continuity of $t \mapsto \xi_t^{\varepsilon, \pm}$ and $t \mapsto g(t, x)$, we obtain that

$$\lim_{t_1 \rightarrow t_2} |v^c(t_2, x) - v^c(t_1, x)| \leq \varepsilon \quad \left(\text{equiv. } \lim_{t_2 \rightarrow t_1} |v^c(t_2, x) - v^c(t_1, x)| \leq \varepsilon\right). \tag{III.9.46}$$

Since $\varepsilon > 0$ is arbitrary we reach our desired result. $\square$

As a consequence of Corollary III.9.6 and Proposition III.9.7, we obtain that the value function for continuous controls $(t, x) \mapsto v^c(t, x)$ is continuous. Notice that, a fortiori, we also obtain that the value functions for Lipschitz controls $(t, x) \mapsto v_n(t, x)$ are continuous for every $n \in \mathbb{N}$. These results are required in Theorem III.8.7.

Corollary III.9.8. *The value function for continuous controls $(t, x) \mapsto v^c(t, x)$ and the value*

functions for n^2 -Lipschitz controls $(t, x) \mapsto v_n(t, x)$ are continuous on $[0, T] \times \mathbb{R}^d$ for every $n \in \mathbb{N}$.

In the last part of this section we show that also hypothesis (III.8.52) holds, i.e., that the value function for singular controls v can be approximated by the value functions for n^2 -Lipschitz controls v_n . To obtain this result, we first introduce a lemma which shows how to build, for a fixed control $\xi \in \mathcal{A}_t$, a sequence of n^2 -Lipschitz controls $(\xi^n)_{n \in \mathbb{N}}$ such that $\xi^n \in \mathcal{A}_t^n$ and ξ^n converges to ξ pointwise. The construction of $(\xi^n)_{n \in \mathbb{N}}$ is inspired by [32, Proposition 5.5].

Lemma III.9.9. *For every $\xi \in \mathcal{A}_t$, there exists a sequence $(\xi^n)_{n \in \mathbb{N}}$ such that $\xi^n \in \mathcal{A}_t^n$ for every $n \in \mathbb{N}$ and, for every $s \in [t, T]$, we have*

$$\lim_{n \rightarrow \infty} \xi_s^n = \xi_s, \quad \mathbb{P}\text{-a.s.} \quad (\text{III.9.47})$$

Moreover, for every $\tau \in \mathcal{T}_t$, we obtain

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\int_t^T |X_s^{\xi^n} - X_s^\xi|^2 ds \right] = 0 \quad (\text{III.9.48})$$

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[|X_\tau^{\xi^n} - X_\tau^\xi|^2 \right] = 0. \quad (\text{III.9.49})$$

Proof. Let $\xi \in \mathcal{A}_t$. To simplify the notation, let us assume that ξ is $\mathbb{P}$ -a.s. non-decreasing. In the general case, one can simply apply the following argument to ξ^+ and ξ^- separately. Let $n \in \mathbb{N}$ and define

$$\xi_s^n := \int_{[t, T]} \varphi^n(s-r) d\tilde{\xi}_r, \quad s \in [t, T], \quad (\text{III.9.50})$$

where $\varphi^n(s) := (ns) \vee 0 \wedge 1$ and $\tilde{\xi} := \xi \wedge n$. For every $r, s, u \in [t, T]$, we have that

$$|\varphi^n(s-r) - \varphi^n(u-r)| \leq n|s-u|, \quad (\text{III.9.51})$$

and so we immediately obtain that ξ^n is n^2 -Lipschitz. Moreover, for every $s \in [t, T]$, we have that

$$\xi_s^n = \tilde{\xi}_{s-\frac{1}{n}} \mathbf{1}_{[t+\frac{1}{n}, T]}(s) + n \int_{[s-\frac{1}{n}, s)} (s-r) d\tilde{\xi}_r. \quad (\text{III.9.52})$$

Thus, $\xi^n \in \mathcal{A}_t^n$ and $|\xi_s^n| \leq |\xi_s|$, $\mathbb{P}$ -a.s. Also, for every $s \in [t, T]$, we obtain

$$|\xi_s^n - \xi_s| \leq |\tilde{\xi}_{s-\frac{1}{n}} \mathbf{1}_{[t+\frac{1}{n}, T]}(s) - \xi_s| + |\tilde{\xi}_s - \tilde{\xi}_{s-\frac{1}{n}}|. \quad (\text{III.9.53})$$

Letting $n \rightarrow \infty$ and using the $\mathbb{P}$ -a.s. left-continuity of ξ , we obtain the first desired result (III.9.47).

By using the assumptions (III.9.5)-(III.9.6) and standard techniques for the estimates of SDEs (e.g., Hölder's inequality and Doob's inequality), for every $s \in [t, T]$, we obtain

$$\begin{aligned} \mathbb{E} \left[|X_s^{\xi^n} - X_s^\xi|^2 \right] &\leq C \mathbb{E} \left[\int_t^s |X_u^{\xi^n} - X_u^\xi|^2 du + |\xi_s^n - \xi_s|^2 \right] \\ &= \int_t^s C \mathbb{E} \left[|X_u^{\xi^n} - X_u^\xi|^2 \right] du + C \mathbb{E} \left[|\xi_s^n - \xi_s|^2 \right], \end{aligned} \quad (\text{III.9.54})$$

for some constant $C > 0$ and where we have used Fubini's lemma in the equality. Since $\alpha^n(s) := \mathbb{E} \left[|X_s^{\xi^n} - X_s^\xi|^2 \right]$ and $\beta^n(s) := C \mathbb{E} \left[|\xi_s^n - \xi_s|^2 \right]$ are non-negative and bounded, we can apply Gronwall's lemma and have that

$$\alpha^n(s) \leq \beta^n(s) e^{CT}. \quad (\text{III.9.55})$$

By letting $n \rightarrow \infty$, using dominated convergence theorem and (III.9.47), we obtain that

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[|X_s^{\xi^n} - X_s^\xi|^2 \right] = 0, \quad s \in [t, T]. \quad (\text{III.9.56})$$

Consequently, by applying Fubini's theorem and dominated convergence theorem, we also have that

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\int_t^T |X_s^{\xi^n} - X_s^\xi|^2 ds \right] = \int_t^T \lim_{n \rightarrow \infty} \mathbb{E} \left[|X_s^{\xi^n} - X_s^\xi|^2 \right] ds = 0, \quad (\text{III.9.57})$$

and (III.9.48) is proved. Now let $\tau \in \mathcal{T}_t$, then in a similar way as in (III.9.54) we obtain

$$\mathbb{E} \left[|X_\tau^{\xi^n} - X_\tau^\xi|^2 \right] \leq C \mathbb{E} \left[\int_t^T |X_s^{\xi^n} - X_s^\xi|^2 ds \right] + C \mathbb{E} \left[|\xi_\tau^n - \xi_\tau|^2 \right]. \quad (\text{III.9.58})$$

By letting $n \rightarrow \infty$, using (III.9.57), (III.9.47) and dominated convergence theorem, we obtain the desired result (III.9.49) and the theorem is proved. $\square$

Proposition III.9.10. *Under Assumption III.9.4 (A) and if $s \mapsto c_\pm(s)$ is continuous on $[0, T]$, we have that*

$$\lim_{n \rightarrow \infty} v_n(t, x) = v(t, x), \quad \forall (t, x) \in [0, T] \times \mathbb{R}^d. \quad (\text{III.9.59})$$

Proof. Let $(t, x) \in [0, T] \times \mathbb{R}^d$. Clearly, we have $v_n(t, x) \leq v_{n+1}(t, x) \leq v(t, x)$ for every $n \in \mathbb{N}$ and so $\lim_{n \rightarrow \infty} v_n(t, x) \leq v(t, x)$. Thus, we need to prove the opposite inequality. To simplify the notation, let us consider the case where the controls $\xi \in \mathcal{A}_t$ are non-decreasing, i.e., $\xi = \xi^+$. Otherwise, simply apply the following argument to ξ^+ and ξ^- separately. Let $\varepsilon > 0$ and $(\tau^\varepsilon, \xi^\varepsilon) \in \mathcal{T}_t \times \mathcal{A}_t$ be ε -optimal for $v(t, x)$, i.e.,

$$J_{t,x}(\tau^\varepsilon, \xi^\varepsilon) \geq v(t, x) - \varepsilon. \quad (\text{III.9.60})$$

Since $\varepsilon > 0$ is fixed, we simply denote $(\tau^\varepsilon, \xi^\varepsilon)$ by (τ, ξ) . For every $n \in \mathbb{N}$, let $\xi^n \in \mathcal{A}_t^n$ be the n^2 -Lipschitz approximation of ξ as in (III.9.50). We now want to show that

$$\lim_{n \rightarrow \infty} J_{t,x}(\tau, \xi^n) = J_{t,x}(\tau, \xi). \quad (\text{III.9.61})$$

For every $n \in \mathbb{N}$, we have

$$\begin{aligned} |J_{t,x}(\tau, \xi^n) - J_{t,x}(\tau, \xi)| &\leq \mathbb{E} \left[\int_t^\tau |f(s, X_s^{\xi^n}) - f(s, X_s^\xi)| ds \right] \\ &\quad + \mathbb{E} \left[\left| \int_{[t,\tau)} \langle c(s), d\xi_s^n \rangle - \int_{[t,\tau)} \langle c(s), d\xi_s \rangle \right| \right] \\ &\quad + \mathbb{E} \left[|g(\tau, X_\tau^{\xi^n}) - g(\tau, X_\tau^\xi)| \right] \end{aligned} \quad (\text{III.9.62})$$

Then, by using Assumption III.9.4 (A) and Jensen's inequality, we obtain

$$\begin{aligned} |J_{t,x}(\tau, \xi^n) - J_{t,x}(\tau, \xi)| &\leq \tilde{C}_f \mathbb{E} \left[\int_t^\tau |X_s^{\xi^n} - X_s^\xi|^2 ds \right]^{\frac{\alpha}{2}} \\ &\quad + \mathbb{E} \left[\left| \int_{[t,\tau)} \langle c(s), d\xi_s^n \rangle - \int_{[t,\tau)} \langle c(s), d\xi_s \rangle \right| \right] \\ &\quad + \tilde{C}_g \mathbb{E} \left[|X_\tau^{\xi^n} - X_\tau^\xi|^2 \right]^{\frac{\alpha}{2}}, \end{aligned} \quad (\text{III.9.63})$$

where $\tilde{C}_f, \tilde{C}_g > 0$. Now, let $\omega \in \Omega$. If $\xi_T(\omega) = 0$ then, by definition (III.9.50) (and mono-

tonicity), we obtain $\xi^n(\omega) \equiv \xi(\omega) \equiv 0$. Also notice that, by definition (III.9.50), if $\xi_T(\omega) > 0$ then $\xi_T^n(\omega) > 0$ for every $n \in \mathbb{N}$. Thus, if $\xi_T(\omega) > 0$ then $s \mapsto \zeta_s^n(\omega) := \xi_s^n(\omega)/\xi_T^n(\omega)$ and $s \mapsto \zeta_s(\omega) := \xi_s(\omega)/\xi_T(\omega)$ are cumulative distribution functions on $[t, T]$ for every $n \in \mathbb{N}$. Moreover, by (III.9.47), we obtain that $\zeta_s^n \rightarrow \zeta_s$, $\mathbb{P}$ -a.s. for every $s \in [t, T]$ which is a continuity point for ζ . Therefore, since $s \mapsto c(s)$ is continuous (and bounded), we have on $\{\xi_T > 0\}$

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_{[t, \tau)} \langle c(s), d\xi_s^n \rangle &= \lim_{n \rightarrow \infty} \xi_T^n \int_{[t, \tau)} \langle c(s), d\zeta_s^n \rangle \\ &= \xi_T \int_{[t, \tau)} \langle c(s), d\zeta_s \rangle \\ &= \int_{[t, \tau)} \langle c(s), d\xi_s \rangle, \quad \mathbb{P}\text{-a.s.} \end{aligned} \quad (\text{III.9.64})$$

Hence, letting $n \rightarrow \infty$ in (III.9.63) and using (III.9.48), (III.9.49), (III.9.64) and the dominated convergence theorem, we obtain

$$\lim_{n \rightarrow \infty} |J_{t,x}(\tau, \xi^n) - J_{t,x}(\tau, \xi)| = 0. \quad (\text{III.9.65})$$

Finally, we have that

$$\lim_{n \rightarrow \infty} v_n(t, x) \geq \lim_{n \rightarrow \infty} J_{t,x}(\tau, \xi^n) = J_{t,x}(\tau, \xi) \geq v(t, x) - \varepsilon, \quad (\text{III.9.66})$$

and letting $\varepsilon \rightarrow 0$ we obtain our desired result. $\square$

III.10 Miscellaneous results

In this last section, we show some miscellaneous results. First, we show that under conditions (III.9.5)-(III.9.6), the assumption (III.4.24) holds.

Proposition III.10.1. *Let $\nu = (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_s^t\}, W)$, $\tilde{\nu} = (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}, \{\tilde{\mathcal{F}}_s^t\}, \tilde{W}) \in \mathcal{V}_t$. Let $\xi \in \mathcal{A}_t^\nu$, $\eta \in \mathcal{E}_t^\nu$, $\tilde{\xi} \in \mathcal{A}_t^{\tilde{\nu}}$ and $\tilde{\eta} \in \mathcal{E}_t^{\tilde{\nu}}$ be such that $\mathcal{L}_\mathbb{P}(\xi^\pm, \eta, W) = \mathcal{L}_{\tilde{\mathbb{P}}}(\tilde{\xi}^\pm, \tilde{\eta}, \tilde{W})$. Then,*

$$\mathcal{L}_\mathbb{P}(\xi^\pm, \eta, X^{\xi, \nu}, \rho^{\xi, \nu}) = \mathcal{L}_{\tilde{\mathbb{P}}}(\tilde{\xi}^\pm, \tilde{\eta}, X^{\tilde{\xi}, \tilde{\nu}}, \rho^{\tilde{\xi}, \tilde{\nu}}). \quad (\text{III.10.1})$$

Proof. Let us denote $X^{\tilde{\xi}, \tilde{\nu}}$ and $\rho^{\tilde{\xi}, \tilde{\nu}}$ by $\tilde{X}$ and $\tilde{\rho}$, respectively. We start by proving that

$$\mathcal{L}_\mathbb{P}(\xi^\pm, \eta, X) = \mathcal{L}_{\tilde{\mathbb{P}}}(\tilde{\xi}^\pm, \tilde{\eta}, \tilde{X}). \quad (\text{III.10.2})$$

For $s \in [t, T]$, we define the Picard iterations

$$\begin{aligned} X_s^0 &:= x + \xi_s, \\ X_s^n &:= x + \int_t^s \mu(u, X_u^{n-1}) du + \int_t^s \sigma(u, X_u^{n-1}) dW_u + \xi_s, \quad n \geq 1, \end{aligned} \quad (\text{III.10.3})$$

and similarly

$$\begin{aligned} \tilde{X}_s^0 &:= x + \tilde{\xi}_s, \\ \tilde{X}_s^n &:= x + \int_t^s \mu(u, \tilde{X}_u^{n-1}) du + \int_t^s \sigma(u, \tilde{X}_u^{n-1}) d\tilde{W}_u + \tilde{\xi}_s, \quad n \geq 1. \end{aligned} \quad (\text{III.10.4})$$

By standard arguments, exploiting the sub-linear growth (III.9.5) and the Lipschitz continuity (III.9.6) (see, e.g., [6, Theorem 9.2]), we obtain that

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} X_s^n = X_s, \forall s \in [t, T]\right) = \tilde{\mathbb{P}}\left(\lim_{n \rightarrow \infty} \tilde{X}_s^n = \tilde{X}_s, \forall s \in [t, T]\right) = 1. \quad (\text{III.10.5})$$

In order to obtain (III.10.2) we now prove, by induction, that

$$\mathcal{L}_{\mathbb{P}}(\xi^{\pm}, \eta, W, X^n) = \mathcal{L}_{\tilde{\mathbb{P}}}(\tilde{\xi}^{\pm}, \tilde{\eta}, \tilde{W}, \tilde{X}^n), \quad \forall n \geq 0. \quad (\text{III.10.6})$$

The case $n = 0$ follows immediately from the hypothesis $\mathcal{L}_{\mathbb{P}}(\xi^{\pm}, \eta, W) = \mathcal{L}_{\tilde{\mathbb{P}}}(\tilde{\xi}^{\pm}, \tilde{\eta}, \tilde{W})$. Thus, let us assume that for all $k \leq n$ we have that $\mathcal{L}_{\mathbb{P}}(\xi^{\pm}, \eta, W, X^k) = \mathcal{L}_{\tilde{\mathbb{P}}}(\tilde{\xi}^{\pm}, \tilde{\eta}, \tilde{W}, \tilde{X}^k)$. This also implies that $\mathcal{L}_{\mathbb{P}}(\xi^{\pm}, \eta, W, X^k, \mu(\cdot, X^k)) = \mathcal{L}_{\tilde{\mathbb{P}}}(\tilde{\xi}^{\pm}, \tilde{\eta}, \tilde{W}, \tilde{X}^k, \mu(\cdot, \tilde{X}^k))$ and so (see, e.g., [75, Theorem 8.3])

$$\mathcal{L}_{\mathbb{P}}\left(\xi^{\pm}, \eta, W, X^k, \int_t^{\cdot} \mu(u, X_u^k) du\right) = \mathcal{L}_{\tilde{\mathbb{P}}}\left(\tilde{\xi}^{\pm}, \tilde{\eta}, \tilde{W}, \tilde{X}^k, \int_t^{\cdot} \mu(u, \tilde{X}_u^k) du\right). \quad (\text{III.10.7})$$

Then, we also obtain that

$$\begin{aligned} \mathcal{L}_{\mathbb{P}}\left(\xi^{\pm}, \eta, W, X^k, \int_t^{\cdot} \mu(u, X_u^k) du, \sigma(\cdot, X^k)\right) \\ = \mathcal{L}_{\tilde{\mathbb{P}}}\left(\tilde{\xi}^{\pm}, \tilde{\eta}, \tilde{W}, \tilde{X}^k, \int_t^{\cdot} \mu(u, \tilde{X}_u^k) du, \sigma(\cdot, \tilde{X}^k)\right), \end{aligned} \quad (\text{III.10.8})$$

and so (see, e.g., [75, Theorem 8.6])

$$\begin{aligned} \mathcal{L}_{\mathbb{P}}\left(\xi^{\pm}, \eta, W, \int_t^{\cdot} \mu(u, X_u^k) du, \int_t^{\cdot} \sigma(u, X_u^k) dW_u\right) \\ = \mathcal{L}_{\tilde{\mathbb{P}}}\left(\tilde{\xi}^{\pm}, \tilde{\eta}, \tilde{W}, \int_t^{\cdot} \mu(u, \tilde{X}_u^k) du, \int_t^{\cdot} \sigma(u, \tilde{X}_u^k) d\tilde{W}_u\right). \end{aligned} \quad (\text{III.10.9})$$

Hence, we have

$$\mathcal{L}_{\mathbb{P}}(\xi^{\pm}, \eta, W, X^{k+1}) = \mathcal{L}_{\tilde{\mathbb{P}}}(\tilde{\xi}^{\pm}, \tilde{\eta}, \tilde{W}, \tilde{X}^{k+1}), \quad (\text{III.10.10})$$

which proves (III.10.6). Thus, from (III.10.5), (III.10.6) and the same limit argument as in (III.4.22), it follows that

$$\mathcal{L}_{\mathbb{P}}(\xi^{\pm}, \eta, X) = \mathcal{L}_{\tilde{\mathbb{P}}}(\tilde{\xi}^{\pm}, \tilde{\eta}, \tilde{X}), \quad (\text{III.10.11})$$

i.e., our first objective (III.10.2).

Now let $t \leq s_1 < s_2 < \dots < s_n \leq T$, $B \in \mathcal{B}(\mathbb{R}^{n(3d+1)})$ and $\hat{B} \in \mathcal{B}([t, T])$. To show the final thesis (III.4.24), we want to prove that

$$\mathbb{P}((\xi_{s_i}^{\pm}, \eta_{s_i}, X_{s_i})_{i=1}^n \in B, \rho \in \hat{B}) = \tilde{\mathbb{P}}((\tilde{\xi}_{s_i}^{\pm}, \tilde{\eta}_{s_i}, \tilde{X}_{s_i})_{i=1}^n \in B, \tilde{\rho} \in \hat{B}). \quad (\text{III.10.12})$$

To reach this result it is sufficient to show that, for every $u \in [t, T]$, we have

$$\mathbb{P}((\xi_{s_i}^{\pm}, \eta_{s_i}, X_{s_i})_{i=1}^n \in B, \rho \geq u) = \tilde{\mathbb{P}}((\tilde{\xi}_{s_i}^{\pm}, \tilde{\eta}_{s_i}, \tilde{X}_{s_i})_{i=1}^n \in B, \tilde{\rho} \geq u), \quad (\text{III.10.13})$$

which is equivalent to

$$\begin{aligned} \mathbb{P}((\xi_{s_i}^{\pm}, \eta_{s_i}, X_{s_i})_{i=1}^n \in B, \{X_{s+} \in \mathcal{O}, \forall s \in [t, u]\}) \\ = \tilde{\mathbb{P}}((\tilde{\xi}_{s_i}^{\pm}, \tilde{\eta}_{s_i}, \tilde{X}_{s_i})_{i=1}^n \in B, \{\tilde{X}_{s+} \in \mathcal{O}, \forall s \in [t, u]\}). \end{aligned} \quad (\text{III.10.14})$$

This last equation holds by (III.10.2). Thus, also (III.4.24) holds and the proof is completed. $\square$

We now show that (III.8.7), which is essential to obtain the optimal stopping rule, holds under Assumption III.9.4 (B)-(C). Notice that the integrability condition in (III.8.6) is also implied by (III.8.7). In the following proposition we set the same framework of Section III.9.

Proposition III.10.2. *Let $\xi \in \mathcal{A}_t$. Under Assumption III.9.4 (B)-(C), we obtain that*

$$\mathbb{E} \left[\sup_{s \in [t, T]} |Y(t, s, \xi)|^2 \right] < \infty. \quad (\text{III.10.15})$$

Proof. By Hölder's inequality, we have that

$$\begin{aligned} \mathbb{E} \left[\sup_{s \in [t, T]} |Y(t, s, \xi)|^2 \right] &\leq C \mathbb{E} \left[\int_t^T |f(s, X_s^\xi)|^2 ds + \int_{[t, T)} \langle |c_+(s)|^2, d\xi_s^+ \rangle \right. \\ &\quad \left. + \int_{[t, T)} \langle |c_-(s)|^2, d\xi_s^- \rangle + \sup_{s \in [t, T]} |g(s, X_s^\xi)|^2 \right]. \end{aligned} \quad (\text{III.10.16})$$

Now notice that, by Assumption III.9.4 (C), we obtain

$$\begin{aligned} \int_{[t, T)} \langle |c_+(s)|^2, d\xi_s^+ \rangle + \int_{[t, T)} \langle |c_-(s)|^2, d\xi_s^- \rangle &\leq M_c^2 \left(\sum_{i=1}^d \int_{[t, T)} d\xi_s^{+, i} + \sum_{i=1}^d \int_{[t, T)} d\xi_s^{-, i} \right) \\ &\leq M_c^2 \left(\sum_{i=1}^d \xi_T^{+, i} + \sum_{i=1}^d \xi_T^{-, i} \right) \\ &= M_c^2 \sqrt{d} [|\xi_T^+| + |\xi_T^-|]. \end{aligned} \quad (\text{III.10.17})$$

Thus, from Assumption III.9.4 (B)-(C) and (III.9.8), we obtain

$$\begin{aligned} \mathbb{E} \left[\sup_{s \in [t, T]} |Y(t, s, \xi)|^2 \right] &\leq \tilde{C}_1 \left(1 + \mathbb{E} \left[\sup_{s \in [t, T]} |X_s^\xi|^2 \right]^{\frac{1}{2}} \right) \\ &\quad + \tilde{C}_2 \left(\mathbb{E} [|\xi_T^+|^2]^{\frac{1}{2}} + \mathbb{E} [|\xi_T^-|^2]^{\frac{1}{2}} \right) \\ &< \infty, \end{aligned} \quad (\text{III.10.18})$$

where $\tilde{C}_1, \tilde{C}_2 > 0$ and we have also used Jensen's inequality. $\square$

III.11 Conclusions

In this chapter we have developed a detailed probabilistic study of singular control with discretionary stopping problems. After setting the framework and showing that our optimisation is independent on the reference probability system, we have rigorously proved the dynamic programming principle for this class of problems, which is the main result of our work. As the control problem we study is in general non-Markovian, we need to rely upon the technical concept of regular conditional probabilities to provide a pseudo-Markov structure essential to obtain the dynamic programming principle. In addition, we have provided an optimal stopping rule for this class of problems under some strong technical assumptions.

We have extended the work by Hausmann and Suo [52] by additionally including in the framework: (i) discretionary stopping, (ii) terminal gain functions and (iii) an exit time from a fixed domain. Our work is designed to directly provide the dynamic programming principle to a reader who is working on a specific problem of singular control with discretionary stopping. As we follow closely the analysis by Fabbri, Gozzi & Swiech [43], who obtain the dynamic programming principle for classical control problems, our work seems extendable to include also classical controls. We also aim to obtain the dynamic programming principle for the finite-fuel case and when considering stopping times and not only deterministic times. Finally, we would like to extend the optimal stopping rule by obtaining this result under milder assumptions.

Chapter IV

On a remarkable decomposition of measures by Dellacherie

In this chapter I present my doctoral research on some decompositions of measures and processes. The core decomposition of this work was defined in [34, Ch. II, Sect. 2, Thm. T14]. We provide an original proof for this decomposition, we study some of its properties and we show how several well-known decompositions can be derived from it.

IV.1 Introduction

In mathematics, the concept of decomposition is often associated with a notion of orthogonality. Namely, the different parts constituting a decomposition are, in some mathematical sense, orthogonal to each other. In measure theory and stochastic analysis, several decompositions of measures and processes have been defined and studied. Since most of these decompositions have been built in distinct frameworks, they reveal different concepts of orthogonality.

Whilst studying the connections between some of these decompositions, we discovered a decomposition of measure that was unknown to us (and we believe to many others). This decomposition was introduced by Dellacherie in [34, Ch. II, Sec. 2, Th. T14] and can be stated as follows: given a σ -finite measure $\mu : \Omega \rightarrow [0, \infty]$ on a measurable space $(\Omega, \mathcal{F})$ and a sub-family $\mathcal{G} \subseteq \mathcal{F}$ closed under countable unions, there exists a set $G_\mu \in \mathcal{G}$ and a unique decomposition

$$\mu = \mu_{\mathcal{G}} + \mu_{\mathcal{G}}^\perp, \tag{IV.1.1}$$

where $\mu(G_\mu^c) = 0$ and $\mu_{\mathcal{G}}^\perp(G) = 0$ for every $G \in \mathcal{G}$.

Although very simple in its statement, this decomposition turns out to be remarkably powerful and, as stated by Dellacherie, deserves to be classical (“mériterait d’être classique”). Indeed, as we show in this chapter, several well-known decompositions of measures and processes can be obtained as simple corollaries of Dellacherie’s decomposition (IV.1.1). This work aims to advertise this decomposition and to improve it in several directions which we now outline. In Section IV.2 we propose a simple original proof of decomposition (IV.1.1), based on the concept of essential supremum, and we outline some initial properties. In fact, Dellacherie’s proof in [34, Ch. II,

Sect. 2, Thm. T14] was initially not clear to us and relies on several non-obvious properties of the order between measures with no specific mention; we also show how well-known decompositions (and some new potentially interesting decompositions) for (positive) measures can be obtained as a simple corollary of (IV.1.1). In Section IV.3 we vastly extend the range of applicability of Dellacherie's decomposition by establishing a version of the theorem for vector measures. Thus, we show how this extension can be applied to signed measures to derive the Jordan decomposition and to the stochastic integral to obtain a decomposition of semimartingales due to Bichteler [15, Theorem 4.4.5]. Finally, in Section IV.4 we study some regularity properties of the decomposition with respect to the order and with respect to the topology.

IV.2 Dellacherie's decomposition for positive measures

In this section we present the central theorem of this work (see Theorem IV.2.1), introduced by Dellacherie in [34, Ch. II, Sect. 2, Thm. T14], for which we propose a simple original proof based on the notion of essential supremum. We also show how to easily derive well-known decompositions of measures from it and we propose some potentially interesting applications.

Throughout the whole chapter, we consider a fixed measurable space $(\Omega, \mathcal{F})$. We say that μ is a positive measure (or simply a measure) if it is a map $\mu : \mathcal{F} \rightarrow [0, \infty]$ such that $\mu(\emptyset) = 0$ and which is countably additive on $\mathcal{F}$.

Theorem IV.2.1. *Let μ be a σ -finite measure on $(\Omega, \mathcal{F})$ and $\mathcal{G} \subseteq \mathcal{F}$ be closed under countable unions. There exist a set $G_\mu \in \mathcal{G}$ and two measures $\mu_{\mathcal{G}}$ and $\mu_{\mathcal{G}}^\perp$ such that*

$$\mu = \mu_{\mathcal{G}} + \mu_{\mathcal{G}}^\perp, \quad (\text{IV.2.1})$$

where $\mu_{\mathcal{G}}(G_\mu^c) = 0$ and $\mu_{\mathcal{G}}^\perp(G) = 0$ for every $G \in \mathcal{G}$. Moreover, the decomposition (IV.2.1) is unique.

Proof. First, let us assume that μ is a finite measure and let

$$S := \mu\text{-ess sup}_{G \in \mathcal{G}} \mathbf{1}_G. \quad (\text{IV.2.2})$$

Then, there exists $\{G_n\}_{n \in \mathbb{N}}$ such that $\sup_n \mathbf{1}_{G_n} = S$, μ -a.e. (see, e.g., [79, Lemma 1.3]). Thus, setting $G_\mu := \cup_n G_n \in \mathcal{G}$, we obtain

$$\mathbf{1}_{G_\mu} = \sup_{n \in \mathbb{N}} \mathbf{1}_{G_n} = S, \quad \mu\text{-a.e.} \quad (\text{IV.2.3})$$

Hence, if $G \in \mathcal{G}$ then $\mathbf{1}_G \leq S = \mathbf{1}_{G_\mu}$, μ -a.e., and so $\mu(G \cap G_\mu^c) = 0$. Therefore, defining $\mu_{\mathcal{G}}(\cdot) := \mu(\cdot \cap G_\mu)$ and $\mu_{\mathcal{G}}^\perp := \mu(\cdot \cap G_\mu^c)$ we obtain the desired result (IV.2.1).

If μ is, more generally, σ -finite then there exists $f > 0$ such that $f \in L^1(\mu)$ (see, e.g., [81, Lemma 6.9]). Therefore, we can apply the same argument as above to the finite positive measure $d\nu = f d\mu$ and obtain the desired result.

To prove uniqueness, assume there exists $\tilde{G}_\mu \in \mathcal{G}$, $\tilde{\mu}_{\mathcal{G}}$, $\tilde{\mu}_{\mathcal{G}}^\perp$ such that $\mu = \tilde{\mu}_{\mathcal{G}} + \tilde{\mu}_{\mathcal{G}}^\perp$, $\tilde{\mu}_{\mathcal{G}}(\tilde{G}_\mu^c) = 0$ and

$\tilde{\mu}_{\mathcal{G}}^{\perp}(G) = 0$ for every $G \in \mathcal{G}$. Then, we must have

$$\mu_{\mathcal{G}}(F) + \mu_{\mathcal{G}}^{\perp}(F) = \tilde{\mu}_{\mathcal{G}}(F) + \tilde{\mu}_{\mathcal{G}}^{\perp}(F), \quad \forall F \in \mathcal{F}. \quad (\text{IV.2.4})$$

Let $F \in \mathcal{F}$, by setting $\hat{G} := G_{\mu} \cup \tilde{G}_{\mu} \in \mathcal{G}$, we obtain

$$\begin{aligned} \mu_{\mathcal{G}}(F) &= \mu_{\mathcal{G}}(F \cap \hat{G}) \\ &= \tilde{\mu}_{\mathcal{G}}(F \cap \hat{G}) - \mu_{\mathcal{G}}^{\perp}(F \cap \hat{G}) + \tilde{\mu}_{\mathcal{G}}^{\perp}(F \cap \hat{G}) \\ &= \tilde{\mu}_{\mathcal{G}}(F \cap \hat{G} \cap \tilde{G}_{\mu}) + \tilde{\mu}_{\mathcal{G}}(F \cap \hat{G} \cap \tilde{G}_{\mu}^c) \\ &= \tilde{\mu}_{\mathcal{G}}(F \cap \tilde{G}_{\mu}) \\ &= \tilde{\mu}_{\mathcal{G}}(F), \end{aligned} \quad (\text{IV.2.5})$$

where in the second equality we have used (IV.2.4) and in the third equality we have also used that $\mu_{\mathcal{G}}^{\perp}(F \cap \hat{G}) = \tilde{\mu}_{\mathcal{G}}^{\perp}(F \cap \hat{G}) = 0$ as $\hat{G} \in \mathcal{G}$. Consequently, we also obtain that $\mu_{\mathcal{G}}^{\perp}(F) = \tilde{\mu}_{\mathcal{G}}^{\perp}(F)$ for every $F \in \mathcal{F}$. Hence, the decomposition is unique. $\square$

Remark IV.2.2. The set $G_{\mu} \in \mathcal{G}$ is in general not unique. However, if $G_{\mu}^1, G_{\mu}^2 \in \mathcal{G}$ both satisfy decomposition (IV.2.1), then it immediately follows that $\mu(G_{\mu}^1 \setminus G_{\mu}^2) = \mu(G_{\mu}^2 \setminus G_{\mu}^1) = 0$. Namely, the set G_{μ} is unique up to μ -null sets. In fact, G_{μ} is a representative of the equivalence class

$$\mathcal{G}_{\mu} := \{G \in \mathcal{G} : \mu(G \Delta G_{\mu}) = 0\}, \quad (\text{IV.2.6})$$

where $\mu(G \Delta G_{\mu}) := \mu(G \setminus G_{\mu}) + \mu(G_{\mu} \setminus G)$. Hence, $\mu(F \cap G_{\mu}^1) = \mu(F \cap G_{\mu}^2)$ for every $G_{\mu}^1, G_{\mu}^2 \in \mathcal{G}_{\mu}$ and $F \in \mathcal{F}$, and so the measures $\mu_{\mathcal{G}}$ and $\mu_{\mathcal{G}}^{\perp}$ are well-defined. We will then call $\mu_{\mathcal{G}}$ and $\mu_{\mathcal{G}}^{\perp}$, respectively, the $\mathcal{G}$ -diffuse part of μ and the $\mathcal{G}$ -orthogonal part of μ .

Remark IV.2.3. As we expect by its construction using the concept of essential supremum, the set G_{μ} is a maximum of $\mathcal{G}$ with respect to the μ -a.e. order. Indeed, we have that $\mu(G \setminus G_{\mu}) = \mu(G \cap G_{\mu}^c) = 0$ for every $G \in \mathcal{G}$, i.e., $\mathbf{1}_G \leq \mathbf{1}_{G_{\mu}}$, μ -a.e. By identifying the sets with their indicator functions, this means that $G \subseteq G_{\mu}$, μ -a.e.

Remark IV.2.4. Notice that if $\mathcal{G}$ is not closed under countable unions but admits a maximum $G_{\mu} \in \mathcal{G}$, either with respect to the μ -a.e. order or more importantly with respect to the pointwise order (the order of inclusion of sets), then Dellacherie's decomposition (IV.2.1) trivially holds by setting $\mu_{\mathcal{G}}(\cdot) := \mu(\cdot \cap G_{\mu})$ and $\mu_{\mathcal{G}}^{\perp}(\cdot) := \mu(\cdot \cap G_{\mu}^c)$. Notice also that for this variant of Theorem IV.2.1, $\mu : \mathcal{F} \rightarrow [0, \infty]$ does not need to be countably additive (as we do not need the concept of essential supremum) but it suffices to be finitely additive.

We now illustrate some applications of Theorem IV.2.1. First, we show how to obtain two of the most relevant decompositions of measures, the Lebesgue decomposition and the atomic/non-atomic decomposition, as a simple and direct corollary of the theorem.

Recall that a measure μ and a measure ν are (mutually) singular if there exists $A \in \mathcal{F}$ such that $\mu(A) = \nu(A^c) = 0$. A measure μ is absolutely continuous with respect to a measure ν if $\nu(A) = 0$ implies $\mu(A) = 0$ for $A \in \mathcal{F}$.

Corollary IV.2.5 (Lebesgue decomposition). *Let μ and ν be two σ -finite measures on $(\Omega, \mathcal{F})$. There exist two measures μ^{ac} , μ^s such that*

$$\mu = \mu^{ac} + \mu^s, \quad (\text{IV.2.7})$$

where μ^{ac} is absolutely continuous with respect to ν , and μ^s and ν are singular. Moreover, the decomposition (IV.2.7) is unique.

Proof. Let $\mathcal{G} := \{A \in \mathcal{F} : \nu(A) = 0\}$, which is closed under countable unions. Thus, Theorem IV.2.1 applied to μ and $\mathcal{G}$ provides the unique decomposition $\mu = \mu_{\mathcal{G}} + \mu_{\mathcal{G}}^{\perp}$ where $\mu_{\mathcal{G}}$ is concentrated on a ν -null set $G_{\mu} \in \mathcal{G}$ and every ν -null set $G \in \mathcal{G}$ is a $\mu_{\mathcal{G}}^{\perp}$ -null set. Hence, $\mu^{ac} := \mu_{\mathcal{G}}^{\perp}$ and $\mu^s := \mu_{\mathcal{G}}$ satisfy the required decomposition (IV.2.7). $\square$

Recall that a set $A \in \mathcal{F}$ is an atom of a measure μ if $\mu(A) > 0$ and for every $B \in \mathcal{F}$ with $B \subseteq A$ either $\mu(B) = 0$ or $\mu(B) = \mu(A)$. We say that a measure μ is (purely) atomic if every measurable set of positive measure contains an atom and non-atomic if it gives null measure to every atom.

Corollary IV.2.6 (Atomic/non-atomic decomposition). *Let μ be a σ -finite measure on $(\Omega, \mathcal{F})$. There exist two measures μ^{pa} , μ^{na} such that*

$$\mu = \mu^{pa} + \mu^{na}, \quad (\text{IV.2.8})$$

where μ^{pa} is purely atomic and μ^{na} is non-atomic. Moreover, the decomposition (IV.2.8) is unique.

Proof. Let $\mathcal{G}$ be the family of all the countable unions of atoms, which is trivially closed under countable unions. Theorem IV.2.1 applied to μ and $\mathcal{G}$ provides the unique decomposition $\mu = \mu_{\mathcal{G}} + \mu_{\mathcal{G}}^{\perp}$ where $\mu_{\mathcal{G}}$ is concentrated on a countable union of atoms $G_{\mu} \in \mathcal{G}$ and $\mu_{\mathcal{G}}^{\perp}$ gives null measure to every atom. Hence, $\mu^{pa} := \mu_{\mathcal{G}}$ and $\mu^{na} := \mu_{\mathcal{G}}^{\perp}$ satisfy the required decomposition (IV.2.8). $\square$

One of the two key components of Theorem IV.2.1 is the family $\mathcal{G} \subseteq \mathcal{F}$ which only needs to be closed under countable unions. We now show some possible choices of $\mathcal{G}$ that lead to the corresponding Dellacherie decomposition of a measure μ into its $\mathcal{G}$ -diffuse part $\mu_{\mathcal{G}}$ and its $\mathcal{G}$ -orthogonal part $\mu_{\mathcal{G}}^{\perp}$. In particular, there are two main operations that provide suitable choices of $\mathcal{G}$.

1. Given a family $\mathcal{H} \subseteq \mathcal{F}$, then the family $\mathcal{G} := \mathcal{H}^{\sigma}$ containing all the countable unions of sets in $\mathcal{H}$ is trivially closed under countable unions.
2. Given a collection $\{\mathcal{G}_i\}_{i \in I}$ of $\mathcal{G}_i \subseteq \mathcal{F}$ all closed under countable unions, then also $\mathcal{G} := \cap_i \mathcal{G}_i$ is closed under countable unions.

Among the potentially interesting applications of these operations, we mention the following ones.

- (i) Let $\mathcal{F}$ be the Borel σ -algebra of a topological space Ω and $\mathcal{H} \subseteq \mathcal{F}$ be the family of closed sets, then $\mathcal{G} := \mathcal{H}^{\sigma}$ is the family of the so called F_{σ} sets, i.e., countable unions of closed sets.
- (ii) Let $\mathcal{F}$ be the Borel σ -algebra of a topological space Ω and $\mathcal{H} \subseteq \mathcal{F}$ be the family of nowhere dense sets, then $\mathcal{G} := \mathcal{H}^{\sigma}$ is the family of meagre sets (also called sets of first category; see, e.g., [1, Section 3.11]).
- (iii) Let $\{\mu_i\}_{i \in I}$ be measures on $(\Omega, \mathcal{F})$ and $\mathcal{G}_i$ be the family of μ_i -null sets, then $\mathcal{G} := \cap_i \mathcal{G}_i$ is the family of so called polar sets for the measures $\{\mu_i\}_{i \in I}$ (see, e.g., the definition in [20, Section 1.1]).

IV.3 Dellacherie's decomposition for vector measures

In this section we show how Theorem IV.2.1 can be extended to vector measures and how this extension allows to derive further well-known decompositions. We first introduce some definitions and results about vector measures.

Let $(V, |\cdot|)$ be a Banach space. A vector measure is a set-function $\Theta : \mathcal{F} \rightarrow V$ such that $\Theta(\emptyset) = 0$ and which is countably additive. Namely, for any sequence $(A_n)_{n \in \mathbb{N}} \subseteq \mathcal{F}$ of pairwise disjoint sets it holds

$$\lim_{k \rightarrow \infty} |\Theta(\bigcup_{n=1}^{\infty} A_n) - \sum_{n=1}^k \Theta(A_n)| = 0. \quad (\text{IV.3.1})$$

For a vector measure $\Theta : \mathcal{F} \rightarrow V$, we define two associated set functions. The (total) variation $|\Theta|$ of Θ is defined as

$$|\Theta|(A) := \sup \sum_{i=1}^n |\Theta(A_i)|, \quad A \in \mathcal{F}, \quad (\text{IV.3.2})$$

where the supremum is taken over all the finite partitions $(A_i)_{i=1}^n$ of A .

The semi-variation $\|\Theta\|$ of Θ is defined as

$$\|\Theta\|(A) := \sup \left| \sum_{i=1}^n \alpha_i \Theta(A_i) \right|, \quad A \in \mathcal{F}, \quad (\text{IV.3.3})$$

where the supremum is taken over all finite collections of real numbers α_i such that $|\alpha_i| \leq 1$ and all finite partitions $(A_i)_{i=1}^n$ of A . In particular, we remark that for the semi-variation the following properties hold (see, e.g., [37, Ch. 4, Sec. 10, Lem. 4])

- (i) $0 \leq |\Theta(A)| \leq \|\Theta\|(A) < \infty$ for every $A \in \mathcal{F}$.
- (ii) $\|\Theta\|(B) \leq \|\Theta\|(A)$ for every $B \subseteq A$ with $A, B \in \mathcal{F}$.

It is now easy to see that the definition of semi-variation IV.3.3 and the properties (i) and (ii) yield

$$\{A \in \mathcal{F} : \|\Theta\|(A) = 0\} = \{A \in \mathcal{F} : \Theta(B) = 0, \forall B \subseteq A, B \in \mathcal{F}\}. \quad (\text{IV.3.4})$$

This naturally provides the correct definition of Θ -null set for a vector measure $\Theta : \mathcal{F} \rightarrow V$. Namely, we say that $A \in \mathcal{F}$ is Θ -null set if $\Theta(B) = 0$ for every $B \subseteq A, B \in \mathcal{F}$. In this case, we also say that Θ is concentrated on A^c .

The semi-variation also allows us to introduce the concept of control measure for a vector measure, which generalises the concept of absolute continuity to vector measures and is essential in the extension of Dellacherie's decomposition theorem. We say that a (finite) positive measure $\theta : \mathcal{F} \rightarrow [0, \infty)$ is a control measure for a vector measure $\Theta : \mathcal{F} \rightarrow V$ if

$$\{A \in \mathcal{F} : \theta(A) = 0\} \subseteq \{A \in \mathcal{F} : \|\Theta\|(A) = 0\}. \quad (\text{IV.3.5})$$

By (IV.3.4), we thus have that if $\theta : \mathcal{F} \rightarrow [0, \infty)$ is a control measure for $\Theta : \mathcal{F} \rightarrow V$ then

$$\{A \in \mathcal{F} : \theta(A) = 0\} \subseteq \{A \in \mathcal{F} : \Theta(B) = 0, \forall B \subseteq A, B \in \mathcal{F}\}, \quad (\text{IV.3.6})$$

i.e., if $A \in \mathcal{F}$ is a θ -null set, then it is also a Θ -null set. Consequently, if θ is concentrated on

$A \in \mathcal{F}$, then also Θ is concentrated on $A \in \mathcal{F}$.

As proved in [37, Ch. 4, Sec. 10, Lem. 5], every vector measure admits a control measure. In fact, it is shown that every vector measure $\Theta : \mathcal{F} \rightarrow V$ admits an equivalent measure $\theta : \mathcal{F} \rightarrow [0, \infty)$, i.e.,

$$\{A \in \mathcal{F} : \theta(A) = 0\} = \{A \in \mathcal{F} : \|\Theta\|(A) = 0\}. \quad (\text{IV.3.7})$$

We are now ready to state and prove our extension of Theorem IV.2.1 to vector measures.

Theorem IV.3.1. *Let Θ be a vector measure on $(\Omega, \mathcal{F})$, $\mathcal{G} \subseteq \mathcal{F}$ be closed under countable unions. There exist a set $G_\Theta \in \mathcal{G}$ and two vector measures $\Theta_\mathcal{G}$ and $\Theta_\mathcal{G}^\perp$ such that*

$$\Theta = \Theta_\mathcal{G} + \Theta_\mathcal{G}^\perp, \quad (\text{IV.3.8})$$

where $\Theta_\mathcal{G}$ is concentrated on $G_\Theta \in \mathcal{G}$ and every $G \in \mathcal{G}$ is $\Theta_\mathcal{G}^\perp$ -null. Moreover, the decomposition (IV.3.8) is unique.

Proof. Let θ be a control measure for Θ . Since θ is a $(\sigma$ -finite) positive measure, we can apply Theorem IV.2.1 to θ and $\mathcal{G}$ and obtain $G_\theta \in \mathcal{G}$ such that $\theta_\mathcal{G}(\cdot) := \theta(\cdot \cap G_\theta)$ is concentrated on G_θ and $\theta_\mathcal{G} := \theta(G \cap G_\theta^c) = 0$ for every $G \in \mathcal{G}$. Let $G_\Theta := G_\theta \in \mathcal{G}$, $\Theta_\mathcal{G}(\cdot) := \Theta(\cdot \cap G_\Theta)$ and $\Theta_\mathcal{G}^\perp(\cdot) := \Theta(\cdot \cap G_\Theta^c)$. Since $\|\Theta_\mathcal{G}\|(\cdot) = \|\Theta\|(\cdot \cap G_\Theta)$ and θ is a control measure for Θ then $\theta_\mathcal{G}$ is a control measure for $\Theta_\mathcal{G}$. As $\theta_\mathcal{G}$ is concentrated on $G_\theta = G_\Theta$, then also $\Theta_\mathcal{G}$ is concentrated on G_Θ . Since θ is a control measure for Θ and $G \cap G_\Theta^c = G \cap G_\theta^c$ is a θ -null set for every $G \in \mathcal{G}$, then it is also a Θ -null set, i.e., every $G \in \mathcal{G}$ is a $\Theta_\mathcal{G}^\perp$ -null set. Hence, the desired properties for the decomposition (IV.3.8) are proved.

To prove uniqueness of the decomposition, consider another set $\tilde{G}_\Theta \in \mathcal{G}$ and two vector measures $\tilde{\Theta}_\mathcal{G}$ and $\tilde{\Theta}_\mathcal{G}^\perp$ such that $\Theta = \tilde{\Theta}_\mathcal{G} + \tilde{\Theta}_\mathcal{G}^\perp$, $\tilde{\Theta}_\mathcal{G}$ is concentrated on $\tilde{G}_\Theta$ and every $G \in \mathcal{G}$ is $\tilde{\Theta}_\mathcal{G}^\perp$ -null. Namely $\tilde{\Theta}_\mathcal{G}(F \cap \tilde{G}_\Theta^c) = 0$ and $\tilde{\Theta}_\mathcal{G}^\perp(F \cap G) = 0$ for every $F \in \mathcal{F}$ and $G \in \mathcal{G}$. Hence,

$$\begin{aligned} \tilde{\Theta}_\mathcal{G}(F) &= \tilde{\Theta}_\mathcal{G}(F \cap \tilde{G}_\Theta) + \tilde{\Theta}_\mathcal{G}(F \cap \tilde{G}_\Theta^c) \\ &= \Theta(F \cap \tilde{G}_\Theta) - \tilde{\Theta}_\mathcal{G}^\perp(F \cap \tilde{G}_\Theta) \\ &= \Theta(F \cap \tilde{G}_\Theta), \end{aligned} \quad (\text{IV.3.9})$$

and so also $\tilde{\Theta}_\mathcal{G}^\perp(F) = \Theta(F \cap \tilde{G}_\Theta^c)$. Finally, to show that $\Theta_\mathcal{G}(F) = \tilde{\Theta}_\mathcal{G}(F)$ (and so $\Theta_\mathcal{G}^\perp(F) = \tilde{\Theta}_\mathcal{G}^\perp(F)$) for every $F \in \mathcal{F}$, set $\hat{G} := G_\Theta \cup \tilde{G}_\Theta \in \mathcal{G}$ and use that

$$\Theta_\mathcal{G}(F) + \Theta_\mathcal{G}^\perp(F) = \tilde{\Theta}_\mathcal{G}(F) + \tilde{\Theta}_\mathcal{G}^\perp(F), \quad \forall F \in \mathcal{F}, \quad (\text{IV.3.10})$$

and an argument as in (IV.2.5). $\square$

Remark IV.3.2. Notice that, similarly to Theorem IV.2.1, the set G_Θ is unique up to unions with Θ -null sets.

A straightforward example of vector measures is given by signed measures, i.e., real-valued (vector) measures. It is possible to uniquely decompose a signed measure μ into a difference of two singular positive measures, i.e., $\mu = \mu^+ - \mu^-$ where μ^+ and μ^- are two singular positive measures. This decomposition is known as Jordan decomposition (see, e.g., [50, Ch. VI, Sec. 29]).

We now show how to obtain the Jordan decomposition of a signed measure μ as a simple corollary

of Theorem IV.3.1. To this extent, we first need to state and prove a general lemma about signed measures. We say that $A \in \mathcal{F}$ is a positive set (for a signed measure μ) if for every $B \in \mathcal{F}$ such that $B \subseteq A$ we have $\mu(B) \geq 0$. Analogously, we say that $A \in \mathcal{F}$ is a negative set if for every $B \in \mathcal{F}$ such that $B \subseteq A$ we have $\mu(B) \leq 0$.

Lemma IV.3.3. *Let μ be a signed measure on $(\Omega, \mathcal{F})$ and $A \in \mathcal{F}$ such that $\mu(A) > 0$. Then, there exists $\bar{A} \in \mathcal{F}$ such that $\bar{A} \subseteq A$, $\bar{A}$ is positive and $\mu(\bar{A}) > 0$.*

Proof. If A is positive then $\bar{A} = A$ satisfies the desired result. Otherwise there exists $B \in \mathcal{F}$, $B \subseteq A$ such that $\mu(B) < 0$. Let $n_1 \in \mathbb{N}$ be the smallest integer such that there exists $B_1 \in \mathcal{F}$, $B_1 \subseteq B$ and $\mu(B_1) < -1/n_1$. Then, let $A_1 := A \setminus B_1$ and notice that $\mu(A_1) = \mu(A) - \mu(B_1) > \mu(A) > 0$. If A_1 is positive then $\bar{A} = A_1$ satisfies the desired result. If not, continue inductively and at the k -th step either $A_{k-1} := A \setminus \bigcup_{i=1}^{k-1} B_i$ is positive (and then the proof is complete) or let $n_k \in \mathbb{N}$ be the smallest integer such that there exists $B_k \in \mathcal{F}$, $B_k \subseteq A_{k-1}$ and $\mu(B_k) < -1/n_k$. Notice that $n_k \geq n_{k-1}$ for every k . If we never stop, let $\bar{A} := A \setminus \bigcup_{i=1}^{\infty} B_i$. Notice that, since $(B_i)_{i \in \mathbb{N}}$ are disjoint, we have again $\mu(\bar{A}) = \mu(A) - \sum_{i=1}^{\infty} \mu(B_i) > \mu(A) > 0$. We just need to prove that $\bar{A}$ is positive. Suppose that, by contradiction, there exists $N \subseteq \bar{A}$ such that $\mu(N) < 0$. Since $0 < \mu(A) < \infty$, from $\mu(A) = \mu(\bar{A}) + \sum_{i=1}^{\infty} \mu(B_i)$ we obtain that the series $\sum_{i=1}^{\infty} \mu(B_i)$ converges. Thus, $\lim_{i \rightarrow \infty} \mu(B_i) = 0$ and so $\lim_{i \rightarrow \infty} 1/n_i = 0$. As $\mu(N) < 0$, we can find $k \in \mathbb{N}$ such that $\mu(N) < -1/(n_k - 1)$. Since $\bar{A} \subseteq A_{k-1}$, then also $N \subseteq A_{k-1}$ and so $\mu(N) < -1/(n_k - 1)$ would imply that the definition of n_k is violated by $n_k - 1$, and thus would lead to a contradiction. $\square$

Corollary IV.3.4 (Jordan decomposition). *Let μ be a signed measure on $(\Omega, \mathcal{F})$. There exist two positive singular measures μ^+ and μ^- such that $\mu = \mu^+ - \mu^-$ and the decomposition is unique.*

Proof. Let $\mathcal{G}$ be the family of positive sets, which is closed under countable unions. Thus, Theorem IV.3.1 provides a positive set $G_\mu \in \mathcal{G}$ and the unique decomposition $\mu = \mu_{\mathcal{G}} + \mu_{\mathcal{G}}^\perp$ where $\mu_{\mathcal{G}}(\cdot) := \mu(\cdot \cap G_\mu)$ and $\mu_{\mathcal{G}}^\perp(G) := \mu(G \cap G_\mu^c) = 0$ for every $G \in \mathcal{G}$. Hence, $\mu_{\mathcal{G}}(A) \geq 0$ for every $A \in \mathcal{F}$. Moreover, $\mu_{\mathcal{G}}^\perp(A) \leq 0$ for every $A \in \mathcal{F}$. Indeed, assume by contradiction that there exists $A \in \mathcal{F}$ such that $\mu_{\mathcal{G}}^\perp(A) > 0$ then, by Lemma IV.3.3, there exists $\bar{A} \in \mathcal{F}$ such that $\bar{A} \subseteq A$, $\bar{A}$ is positive (i.e., $\bar{A} \in \mathcal{G}$) and $\mu_{\mathcal{G}}^\perp(\bar{A}) > 0$ which is a contradiction. Thus, $\mu^+ := \mu_{\mathcal{G}}$ and $\mu^- := -\mu_{\mathcal{G}}^\perp$ form the Jordan decomposition of μ . $\square$

Remark IV.3.5. *Notice that the variation $|\mu|$ of the signed measure μ satisfies $|\mu| = \mu^+ + \mu^-$ and that (G_μ, G_μ^c) form a Hahn decomposition of Ω (see, e.g., [50, Ch. 6, Sec. 29]).*

Another remarkable example of vector measure is the stochastic integral. We only report some definitions and results on this topic (which is rather technical), as our aim here is simply to show an application of Theorem IV.3.8 for the stochastic integral. We refer to, e.g., [86] and [13] for a deeper study of the link between vector measures and stochastic integral.

Let $T > 0$ and $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0, T]}, \mathbb{P})$ be a complete filtered probability space. We say that a process $X := (X_t)_{t \in [0, T]}$ is an H^p -semimartingale ($1 \leq p < \infty$) if it admits a unique decomposition $X = M + A$ where M is a local martingale and A is a predictable finite variation process such that

$$\mathbb{E} \left[\sup_{t \in [0, T]} |M_t|^p \right] < \infty, \quad \text{and} \quad \mathbb{E} \left[\left(\int_0^T |dA_t| \right)^p \right] < \infty. \quad (\text{IV.3.11})$$

Let $\mathcal{P}$ be the predictable σ -algebra on $[0, T] \times \Omega$, then X is an H^p -semimartingale if and only if the stochastic integral $I_X : \mathcal{P} \rightarrow L^p$ defined by $I_X(A) := \mathbf{1}_A \cdot X := \int_0^T \mathbf{1}_A dX_t$ is a vector measure

(see, e.g., [86, Theorem 3.1]). Moreover, the finite positive measure defined on $\mathcal{P}$ by

$$\theta_X(dt \times d\omega) := (d[M, M]_t^{\frac{1}{2}}(\omega) + |dA_t(\omega)|)d\mathbb{P}(\omega) \quad (\text{IV.3.12})$$

is a control measure for the vector measure I_X (see, e.g., [86, Theorem 3.4]). This argument can be extended to semimartingales. Indeed, let X be a semimartingale on a complete filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0, T]}, \mathbb{Q})$ then (see, e.g., [80, Chapter 4, Theorem 34]) there exists an equivalent measure $\mathbb{P} \sim \mathbb{Q}$ such that X is an H^2 -semimartingale under $\mathbb{P}$. Thus, θ_X as in (IV.3.12) is a control measure for $X = M + A$, where $X = M + A$ is the canonical decomposition of X under $\mathbb{P}$.

We now show how to obtain one decomposition of processes that appears in [15, Theorem 4.4.5] as an application of Theorem IV.3.1 to the stochastic integral. To do so, we first recall some notions. For a stopping time τ , we denote by $\llbracket \tau \rrbracket$ its graph, i.e.,

$$\llbracket \tau \rrbracket := \{(t, \omega) \in [0, T] \times \Omega : \tau(\omega) = t\}. \quad (\text{IV.3.13})$$

We will say that a set $J \subseteq \Omega \times [0, T]$ is thin if $J = \cup_n \llbracket \tau_n \rrbracket$ for some stopping times $(\tau_n)_{n \in \mathbb{N}}$; we call such J a thin predictable set if $J = \cup_n \llbracket \tau_n \rrbracket$ for some predictable stopping times $(\tau_n)_{n \in \mathbb{N}}$, and a thin totally inaccessible set if $J = \cup_n \llbracket \tau_n \rrbracket$ for some totally inaccessible stopping times $(\tau_n)_{n \in \mathbb{N}}$. We will use that the set of jumps $J := \{\Delta X \neq 0\}$ of a cadlag adapted process X is a thin set (see, e.g., [56, Proposition 1.32]). We will say that a cadlag process X has predictable jump times if $\{\Delta X \neq 0\}$ is thin predictable, and that it has totally inaccessible jump times if $\{\Delta X \neq 0\}$ is thin totally inaccessible.

Corollary IV.3.6 (Bichteler decomposition). *Let X be a (cadlag) semimartingale. There exists a thin predictable set $P \in \mathcal{P}$ such that*

$$X = Y + Z, \quad (\text{IV.3.14})$$

where $Y = \mathbf{1}_P \cdot X$ and Z has totally inaccessible jump times. Moreover, the decomposition (IV.3.14) is unique.

Proof. Let $\mathcal{G}$ be the family of predictable thin sets, which is trivially closed under countable unions. By applying Theorem IV.3.1 to $\mathcal{G}$ and the vector measure $\Theta := I_X : \mathcal{P} \rightarrow L^p$, defined by $I_X(A) := \mathbf{1}_A \cdot X$, we obtain a predictable thin set $P := G_\Theta \in \mathcal{G}$. Let us set $Y := \mathbf{1}_P \cdot X$, $Z := \mathbf{1}_{P^c} \cdot X$. Since $\mathbf{1}_P \mathbf{1}_{P^c} = \mathbf{1}_P$ and the stochastic integral satisfies $H \cdot (K \cdot X) = (HK) \cdot X$, we have

$$Y = \mathbf{1}_P \cdot X = (\mathbf{1}_P \mathbf{1}_{P^c}) \cdot X = \mathbf{1}_P \cdot (\mathbf{1}_{P^c} \cdot X) = \mathbf{1}_P \cdot Y. \quad (\text{IV.3.15})$$

Moreover, Theorem IV.3.1 also gives that

$$0 = \Theta_G^\perp(G) = \Theta(G \cap G_\Theta^c) = I_X(G \cap P^c) = \mathbf{1}_{G \cap P^c} \cdot X, \quad \forall G \in \mathcal{G}. \quad (\text{IV.3.16})$$

Let τ be a predictable stopping time then $\llbracket \tau \rrbracket \in \mathcal{G}$ and so

$$0 = \mathbf{1}_{\llbracket \tau \rrbracket \cap P^c} \cdot X = (\mathbf{1}_{\llbracket \tau \rrbracket} \mathbf{1}_{P^c}) \cdot X = \mathbf{1}_{\llbracket \tau \rrbracket} \cdot (\mathbf{1}_{P^c} \cdot X) = \mathbf{1}_{\llbracket \tau \rrbracket} \cdot Z = \Delta Z_\tau. \quad (\text{IV.3.17})$$

Hence, if J is the set of jumps of Z , then $J \cap \llbracket \tau \rrbracket = \emptyset$. Let $J = \cup_n \llbracket \tau_n \rrbracket$ for some stopping times $(\tau_n)_{n \in \mathbb{N}}$, then

$$\llbracket \tau_n \rrbracket \cap \llbracket \tau \rrbracket \subseteq J \cap \llbracket \tau \rrbracket = \emptyset, \quad \forall n \in \mathbb{N}, \quad (\text{IV.3.18})$$

so every τ_n is totally inaccessible, i.e., J is thin totally inaccessible and the properties of decomposition (IV.3.14) are complete.

Let $X = \tilde{Y} + \tilde{Z}$ be another such decomposition where $\tilde{Y} = \mathbf{1}_{\tilde{P}} \cdot \tilde{Y}$ and $\tilde{P}$ is a thin predictable set. Then, $\hat{P} := P \cup \tilde{P}$ is also a thin predictable set and the semimartingale $W := Y - \tilde{Y} = Z - \tilde{Z}$ satisfies (i) $W = \mathbf{1}_{\hat{P}} \cdot W$ and (ii) $\mathbf{1}_{\hat{P}} \cdot W = 0$. Hence, $W = 0$ and the decomposition is unique. $\square$

Remark IV.3.7. Corollary IV.3.6 can then be applied to obtain further known decompositions. Applying Corollary IV.3.6 to a purely-discontinuous process of finite variation $V = X$, we obtain two purely-discontinuous processes of finite variation $V^{jpr} := Y$ and $V^{jti} := Z$, where V^{jpr} has predictable jump times and V^{jti} has totally inaccessible jump times, i.e., the decomposition in [59, Proposition 22.17]. Applying it to a continuous process of finite variation $V = X$, gives $Y = 0$, $Z = V$.

Analogously, applying Corollary IV.3.6 to a purely-discontinuous local martingale $M = X$ gives two purely-discontinuous local martingales $M^{jpr} := Y$ and $M^{jti} := Z$, where M^{jpr} has predictable jump times and M^{jti} has totally-inaccessible jump times. Applying it to a continuous local martingale $M = X$ gives $Y = 0$, $Z = M$. Thus, we have also obtained the decompositions in [14, Theorem 7.9].

IV.4 Regularity properties for Dellacherie's decomposition

In this section we show some regularity properties of Dellacherie's decomposition for positive measures. In particular, we study the dependence of the $\mathcal{G}$ -diffuse part $\mu_{\mathcal{G}}$ and of the equivalence class $\mathcal{G}_{\mu}$ on μ and on $\mathcal{G}$. Notice that similar properties clearly follow for the $\mathcal{G}$ -orthogonal part $\mu_{\mathcal{G}}^{\perp}$.

Let $\mathcal{M}$ be the space of signed measures (i.e., real-valued measures) on $(\Omega, \mathcal{F})$ and $\mathcal{M}_+ \subseteq \mathcal{M}$ be the subset of positive finite measures. We endow $\mathcal{M}$ with the setwise (partial) order $\succeq$ and with the total variation norm $|\cdot|_{\mathcal{M}}$ defined, for $\mu, \nu \in \mathcal{M}$, as $\mu \succeq \nu$ if $\mu(A) \geq \nu(A)$ for every $A \in \mathcal{F}$ and $|\mu|_{\mathcal{M}} := |\mu|(\Omega)$. It is interesting to study both the order structure derived from $\succeq$ and the topological structure derived from $|\cdot|_{\mathcal{M}}$. It is shown in [1, Theorem 10.53] that $(\mathcal{M}, \succeq)$ is a Riesz space, i.e., an ordered vector space such that each pair of elements has a supremum and an infimum (with respect to the order). It is also proved that any increasing (and bounded from above) net $\{\mu_i\}_{i \in I} \subseteq \mathcal{M}$ admits a supremum $\vee_{i \in I} \mu^i$ which satisfies

$$(\vee_{i \in I} \mu^i)(A) = \sup_{i \in I} \mu^i(A), \quad A \in \mathcal{F}. \quad (\text{IV.4.1})$$

Let $\mathcal{G} \subseteq \mathcal{F}$ be closed under countable unions. For the map $\mu \in \mathcal{M}_+ \mapsto \mu_{\mathcal{G}} \in \mathcal{M}_+$ we have the following results.

Proposition IV.4.1. *The map $\mu \mapsto \mu_{\mathcal{G}}$ is*

- (i) *increasing, i.e., if $\mu, \nu \in \mathcal{M}_+$ satisfy $\mu \succeq \nu$ then $\mu_{\mathcal{G}} \succeq \nu_{\mathcal{G}}$;*
- (ii) *additive, i.e., if $\mu, \nu \in \mathcal{M}_+$ then $(\mu + \nu)_{\mathcal{G}} = \mu_{\mathcal{G}} + \nu_{\mathcal{G}}$;*
- (iii) *positively-homogeneous, i.e., if $\mu \in \mathcal{M}_+$, $\alpha \geq 0$ then $(\alpha\mu)_{\mathcal{G}} = \alpha\mu_{\mathcal{G}}$;*

(iv) continuous from below with respect to the order $\succeq$, i.e., if $\{\mu^i\}_{i \in I} \subseteq \mathcal{M}_+$ is an increasing net bounded from above then

$$\left(\bigvee_{i \in I} \mu^i \right)_{\mathcal{G}} = \bigvee_{i \in I} \mu_{\mathcal{G}}^i. \quad (\text{IV.4.2})$$

(v) Lipschitz continuous with constant $K = 1$ with respect to the norm $|\cdot|$, i.e., if $\mu, \nu \in \mathcal{M}_+$ then $|\mu_{\mathcal{G}} - \nu_{\mathcal{G}}|_{\mathcal{M}} \leq |\mu - \nu|_{\mathcal{M}}$.

Proof. Let $\mu, \nu \in \mathcal{M}_+$, $\alpha \geq 0$ and $\{\mu^i\}_{i \in I} \subseteq \mathcal{M}_+$ be an increasing net bounded from above. For every measure $\rho \in \mathcal{M}_+$, by Theorem IV.2.1, let $G_{\rho} \in \mathcal{G}$ be such that $\rho_{\mathcal{G}}(\cdot) := \rho(\cdot \cap G_{\rho})$ and $\rho(G \cap G_{\rho}^c) = 0$ for every $G \in \mathcal{G}$.

We start by proving (i). Since $\mu(G \cap G_{\mu}^c) = 0$ for every $G \in \mathcal{G}$ then also $\mu(G_{\nu} \cap G_{\mu}^c) = 0$. Hence, for every $F \in \mathcal{F}$,

$$\begin{aligned} \mu(F \cap G_{\nu}) &= \mu(F \cap G_{\nu} \cap G_{\mu}) + \mu(F \cap G_{\nu} \cap G_{\mu}^c) \\ &\leq \mu(F \cap G_{\mu}) + \mu(G_{\nu} \cap G_{\mu}^c) \\ &= \mu(F \cap G_{\mu}). \end{aligned} \quad (\text{IV.4.3})$$

Therefore, for every $F \in \mathcal{F}$, we obtain

$$\mu_{\mathcal{G}}(F) = \mu(F \cap G_{\mu}) \geq \mu(F \cap G_{\nu}) \geq \nu(F \cap G_{\nu}) = \nu_{\mathcal{G}}(F), \quad (\text{IV.4.4})$$

and (i) is proved.

We now prove (ii). By defining $\rho := \mu + \nu \in \mathcal{M}_+$, we aim to show that $\rho_{\mathcal{G}} = \mu_{\mathcal{G}} + \nu_{\mathcal{G}}$. By the same argument as in (IV.4.3) applied to μ and ν , separately, notice that

$$\mu(F \cap G_{\rho}) \leq \mu(F \cap G_{\mu}) \quad \text{and} \quad \nu(F \cap G_{\rho}) \leq \nu(F \cap G_{\nu}), \quad \forall F \in \mathcal{F}. \quad (\text{IV.4.5})$$

Since $\rho \succeq \mu$, we also have that

$$0 = \rho(G_{\mu} \cap G_{\rho}^c) \geq \mu(G_{\mu} \cap G_{\rho}^c) \geq 0, \quad (\text{IV.4.6})$$

and so $\mu(G_{\mu} \cap G_{\rho}^c) = 0$. By a similar argument as in (IV.4.3), this yields $\mu(F \cap G_{\rho}) \geq \mu(F \cap G_{\mu})$ for every $F \in \mathcal{F}$. By the same argument applied to ν , we also obtain $\nu(F \cap G_{\rho}) \geq \nu(F \cap G_{\nu})$ for every $F \in \mathcal{F}$. These results combined with (IV.4.5) imply that

$$\mu(F \cap G_{\rho}) = \mu(F \cap G_{\mu}) \quad \text{and} \quad \nu(F \cap G_{\rho}) = \nu(F \cap G_{\nu}), \quad \forall F \in \mathcal{F}. \quad (\text{IV.4.7})$$

Hence, for every $F \in \mathcal{F}$,

$$\rho_{\mathcal{G}}(F) = \mu(F \cap G_{\rho}) + \nu(F \cap G_{\rho}) = \mu(F \cap G_{\mu}) + \nu(F \cap G_{\nu}) = \mu_{\mathcal{G}}(F) + \nu_{\mathcal{G}}(F), \quad (\text{IV.4.8})$$

and (ii) is proved.

We now prove (iii). By defining $\rho := \alpha\mu \in \mathcal{M}_+$ we aim to show that $\rho_{\mathcal{G}} = \alpha\mu_{\mathcal{G}}$. By the same argument as in (IV.4.3) applied to μ and ρ , separately, we obtain that $\mu(F \cap G_{\rho}) \leq \mu(F \cap G_{\mu})$ and $\rho(F \cap G_{\mu}) \leq \rho(F \cap G_{\rho})$ for every $F \in \mathcal{F}$. Hence, for every $F \in \mathcal{F}$,

$$\rho_{\mathcal{G}}(F) = \alpha\mu(F \cap G_{\rho}) \leq \alpha\mu(F \cap G_{\mu}) = \alpha\mu_{\mathcal{G}}(F), \quad (\text{IV.4.9})$$

and

$$\rho_{\mathcal{G}}(F) = \rho(F \cap G_{\rho}) \geq \rho(F \cap G_{\mu}) = \alpha\mu(F \cap G_{\mu}) = \alpha\mu_{\mathcal{G}}(F). \quad (\text{IV.4.10})$$

Thus, $\rho_{\mathcal{G}} = \alpha\mu_{\mathcal{G}}$ and (iii) is proved.

We now prove (iv). By defining $\rho := \vee_{i \in I} \mu^i$, we aim to show that $\rho_{\mathcal{G}} = \vee_{i \in I} \mu_{\mathcal{G}}^i$. Since $\rho \succeq \mu^i$ for every $i \in I$ then, by (i), $\rho_{\mathcal{G}} \succeq \mu_{\mathcal{G}}^i$ for every $i \in I$. Therefore, by definition of supremum, we obtain the first inequality

$$\rho_{\mathcal{G}} \succeq \vee_{i \in I} \mu_{\mathcal{G}}^i. \quad (\text{IV.4.11})$$

To prove the opposite inequality, apply for every $i \in I$ the same argument as in (IV.4.3) and obtain

$$\mu_{\mathcal{G}}^i(F) = \mu_{\mathcal{G}}^i(F \cap G_{\mu^i}) \geq \mu^i(F \cap G_{\rho}), \quad \forall F \in \mathcal{F}. \quad (\text{IV.4.12})$$

Hence, for every $i \in I$,

$$\vee_{i \in I} \mu_{\mathcal{G}}^i(F) \geq \mu^i(F \cap G_{\rho}), \quad \forall F \in \mathcal{F} \quad (\text{IV.4.13})$$

By taking the supremum, we obtain

$$\vee_{i \in I} \mu_{\mathcal{G}}^i(F) \geq \sup_{i \in I} \mu^i(F \cap G_{\rho}) = (\vee_{i \in I} \mu^i)(F \cap G_{\rho}) = \rho(F \cap G_{\rho}) = \rho_{\mathcal{G}}(F), \quad \forall F \in \mathcal{F}, \quad (\text{IV.4.14})$$

where in the first equality we have used (IV.4.1). Thus, (iv) is proved.

We now prove (v). Let $\mathcal{P}(A)$ denote the finite partitions of $A \in \mathcal{F}$. By defining $\hat{G} := G_{\mu} \cup G_{\nu} \in \mathcal{G}$, we obtain

$$\begin{aligned} |\mu_{\mathcal{G}} - \nu_{\mathcal{G}}|_{\mathcal{M}} &= |\mu_{\mathcal{G}} - \nu_{\mathcal{G}}|(\Omega) = \sup_{\{A_i\} \in \mathcal{P}(\Omega)} \sum_{i=1}^n |(\mu_{\mathcal{G}} - \nu_{\mathcal{G}})(A_i)| \\ &= \sup_{\{A_i\} \in \mathcal{P}(\Omega)} \sum_{i=1}^n |\mu(A_i \cap G_{\mu}) - \nu(A_i \cap G_{\nu})| \\ &= \sup_{\{A_i\} \in \mathcal{P}(\Omega)} \sum_{i=1}^n |\mu(A_i \cap \hat{G}) - \nu(A_i \cap \hat{G})| \\ &\leq \sup_{\{B_i\} \in \mathcal{P}(\hat{G})} \sum_{i=1}^n |(\mu - \nu)(B_i)| \\ &= |\mu - \nu|(\hat{G}) \leq |\mu - \nu|_{\mathcal{M}}, \end{aligned} \quad (\text{IV.4.15})$$

and the proof is complete. $\square$

Let $\mu \in \mathcal{M}_+$. We now set the framework to study the order properties for the maps $\mathcal{G} \mapsto \mu_{\mathcal{G}}$ and $\mathcal{G} \mapsto \mathcal{G}_{\mu}$, where $\mathcal{G}_{\mu}$ denotes the equivalence class (IV.2.6) outlined by the Dellacherie's decomposition of μ with respect to $\mathcal{G}$, i.e., the $(\mu$ -a.e.) maximum sets in $\mathcal{G}$. Notice that, for the sake of simplicity, in the following we do not distinguish between the equivalence class $\mathcal{G}_{\mu}$ and its representative G_{μ} , i.e., we identify $\mathcal{G}_{\mu}$ as one of its representatives.

Let $\mathbb{S} := \{\mathcal{G} : \mathcal{G} \subseteq \mathcal{F}\}$ be the class of subfamilies of $\mathcal{F}$ and define its subclass

$$\mathbb{G} := \{\mathcal{G} \in \mathbb{S} : (G_n)_{n \in \mathbb{N}} \subseteq \mathcal{G} \Rightarrow \cup_{n \in \mathbb{N}} G_n \in \mathcal{G}\} \quad (\text{IV.4.16})$$

of all subfamilies of $\mathcal{F}$ which are closed under countable unions. We endow $\mathbb{S}$ and $\mathbb{G}$ with the order of inclusion of sets.

Let $\mathcal{G} \in \mathbb{S}$. Since the intersection of families closed under countable unions is closed under countable unions, by taking the intersection of every $\mathcal{H} \in \mathbb{G}$ such that $\mathcal{G} \subseteq \mathcal{H}$ we obtain the smallest element of $\mathbb{G}$ that contains $\mathcal{G}$, which we denote by $\mathcal{G}^\sigma$. Notice that $\mathcal{G}^\sigma$ can be, equivalently, defined as the family of countable unions of sets in $\mathcal{G}$ and that

$$\mathbb{G} = \{\mathcal{G} \in \mathbb{S} : \mathcal{G} = \mathcal{G}^\sigma\}. \quad (\text{IV.4.17})$$

For an arbitrary family $\{\mathcal{G}^i\}_{i \in I} \subseteq \mathbb{G}$, its infimum $\bigwedge_{i \in I} \mathcal{G}^i$ and supremum $\bigvee_{i \in I} \mathcal{G}^i$ in $\mathbb{G}$ are given by

$$\bigwedge_{i \in I} \mathcal{G}^i = \bigcap_{i \in I} \mathcal{G}^i \quad (\text{IV.4.18})$$

$$\bigvee_{i \in I} \mathcal{G}^i = (\bigcup_{i \in I} \mathcal{G}^i)^\sigma. \quad (\text{IV.4.19})$$

Moreover, notice that

$$(\bigcup_{i \in I} \mathcal{G}^i)^\sigma = \{\bigcup_{i \in J} G_i : J \subseteq I \text{ is countable, } J \neq \emptyset \text{ and } G_i \in \mathcal{G}^i\}, \quad (\text{IV.4.20})$$

and so

$$\bigvee_{i \in I} \mathcal{G}^i = \{\bigcup_{i \in J} G_i : J \subseteq I \text{ is countable, } J \neq \emptyset \text{ and } G_i \in \mathcal{G}^i\}. \quad (\text{IV.4.21})$$

For instance, if $\mathcal{G}, \mathcal{H} \in \mathbb{G}$ then

$$\mathcal{G} \vee \mathcal{H} = \{G, H, G \cup H : G \in \mathcal{G}, H \in \mathcal{H}\}. \quad (\text{IV.4.22})$$

We now show some order properties for the map $\mathcal{G} \in \mathbb{G} \mapsto \mathcal{G}_\mu$.

Proposition IV.4.2. *The map $\mathcal{G} \mapsto \mathcal{G}_\mu$ is increasing, i.e., if $\mathcal{G} \subseteq \mathcal{H}$ then $\mathcal{G}_\mu \subseteq \mathcal{H}_\mu$, μ -a.e. and it satisfies the following continuity from below: if $\{\mathcal{G}^i\}_{i \in I} \subseteq \mathbb{G}$, then there exists $J \subseteq I$ countable such that*

$$(\bigvee_{i \in I} \mathcal{G}^i)_\mu = (\bigvee_{i \in J} \mathcal{G}^i)_\mu = \bigcup_{i \in J} \mathcal{G}_\mu^i. \quad (\text{IV.4.23})$$

Proof. Monotonicity is straightforward: as $\mathcal{H}_\mu$ is the maximum in $\mathcal{H}$ and $\mathcal{G}_\mu \in \mathcal{G} \subseteq \mathcal{H}$, then also $\mathcal{G}_\mu \subseteq \mathcal{H}_\mu$, μ -a.e.

To prove (IV.4.23) we first show $(\bigvee_{i \in J} \mathcal{G}^i)_\mu = \bigcup_{i \in J} \mathcal{G}_\mu^i$ for a countable set J . Namely, we want to prove that $\bigcup_{i \in J} \mathcal{G}_\mu^i \in \bigvee_{i \in J} \mathcal{G}^i$ and that $\bigcup_{i \in J} \mathcal{G}_\mu^i$ is a maximum in $\bigvee_{i \in J} \mathcal{G}^i$. By (IV.4.21), it is easy to see that $\bigcup_{i \in J} \mathcal{G}_\mu^i \in \bigvee_{i \in J} \mathcal{G}^i$. Now let $G \in \bigvee_{i \in J} \mathcal{G}^i$. By (IV.4.21), we have that $G = \bigcup_{i \in J_1} G_i$ with $J_1 \subseteq J$ and $G_i \in \mathcal{G}^i$. Since $\mathcal{G}_\mu^i$ is the maximum in $\mathcal{G}^i$, then $G \subseteq \bigcup_{i \in J} \mathcal{G}_\mu^i$, μ -a.e. and so

$$(\bigvee_{i \in J} \mathcal{G}^i)_\mu = \bigcup_{i \in J} \mathcal{G}_\mu^i. \quad (\text{IV.4.24})$$

To complete the proof, we now want to prove that there exists $J \subseteq I$ countable such that $(\bigvee_{i \in I} \mathcal{G}^i)_\mu = \bigcup_{i \in J} \mathcal{G}_\mu^i$. By (IV.4.21), let $J \subseteq I$ be countable such that $(\bigvee_{i \in I} \mathcal{G}^i)_\mu = \bigcup_{i \in J} G_i$ with $G_i \in \mathcal{G}^i$ and such that $\bigcup_{i \in J} G_i$ is a maximum in $\bigvee_{i \in I} \mathcal{G}^i$. Since $\bigcup_{i \in J} \mathcal{G}_\mu^i \in \bigvee_{i \in I} \mathcal{G}^i$, by maximality of $\bigcup_{i \in J} G_i$, we obtain $\bigcup_{i \in J} \mathcal{G}_\mu^i \subseteq \bigcup_{i \in J} G_i$, μ -a.e. Since by the previous step $\bigcup_{i \in J} \mathcal{G}_\mu^i$ is the maximum in $\bigvee_{i \in J} \mathcal{G}^i$ and since $\bigcup_{i \in J} G_i \in \bigvee_{i \in J} \mathcal{G}^i$, we then obtain that $\bigcup_{i \in J} G_i \subseteq \bigcup_{i \in J} \mathcal{G}_\mu^i$, μ -a.e., which completes the proof. $\square$

For the map $\mathcal{G} \in \mathbb{G} \mapsto \mu_{\mathcal{G}} \in \mathcal{M}_+$, we have the following results.

Proposition IV.4.3. *The map $\mathcal{G} \mapsto \mu_{\mathcal{G}}$ is*

(i) increasing, i.e., if $\mathcal{G} \subseteq \mathcal{H}$ then $\mu_{\mathcal{G}} \preceq \mu_{\mathcal{H}}$;

(ii) continuous from below with respect to the order, i.e., if $\{\mathcal{G}^i\}_{i \in I} \subseteq \mathbb{G}$ then

$$\mu_{\bigvee_{i \in I} \mathcal{G}^i} = \bigvee_{i \in I} \mu_{\mathcal{G}^i}. \quad (\text{IV.4.25})$$

Proof. To prove monotonicity, recall that $\mu_{\mathcal{G}}(\cdot) := \mu(\cdot \cap \mathcal{G}_{\mu})$ and $\mu_{\mathcal{H}}(\cdot) := \mu(\cdot \cap \mathcal{H}_{\mu})$ with $\mathcal{G}_{\mu} \in \mathcal{G}$ and $\mathcal{H}_{\mu} \in \mathcal{H}$. Moreover, $\mathcal{H}_{\mu}$ is the maximum in $\mathcal{H}$ and so $\mu(F \cap H) \leq \mu(F \cap \mathcal{H}_{\mu})$ for every $H \in \mathcal{H}$ (in particular, for $\mathcal{G}_{\mu} \in \mathcal{G} \subseteq \mathcal{H}$) and $F \in \mathcal{F}$. Thus, also $\mu_{\mathcal{G}} \preceq \mu_{\mathcal{H}}$.

To prove (IV.4.25), we first show it holds for $I = \{1, 2\}$. By the (just proved) monotonicity of $\mathcal{G} \mapsto \mu_{\mathcal{G}}$, we easily have that $\mu_{\mathcal{G}^1 \vee \mathcal{G}^2} \succeq \mu_{\mathcal{G}^1} \vee \mu_{\mathcal{G}^2}$. To show that the equality holds, we need to show that if $\nu \succeq \mu_{\mathcal{G}^1} \vee \mu_{\mathcal{G}^2}$ then $\nu \succeq \mu_{\mathcal{G}^1 \vee \mathcal{G}^2}$. Since (by Proposition IV.4.2) $\mu_{\mathcal{G}^1 \vee \mathcal{G}^2}(\cdot) := \mu(\cdot \cap (\mathcal{G}_{\mu}^1 \cup \mathcal{G}_{\mu}^2))$, then showing $\nu \succeq \mu_{\mathcal{G}^1 \vee \mathcal{G}^2}$ is equivalent to showing that $\nu(F) \geq \mu(F)$ for every $F \subseteq \mathcal{G}_{\mu}^1 \cup \mathcal{G}_{\mu}^2$, $F \in \mathcal{F}$. For the same reason, $\nu \succeq \mu_{\mathcal{G}^1} \vee \mu_{\mathcal{G}^2} \succeq \mu_{\mathcal{G}^i}$, $i = 1, 2$, implies that $\nu(F) \geq \mu(F)$ for every $F \subseteq \mathcal{G}_{\mu}^1$ and every $F \subseteq \mathcal{G}_{\mu}^2$, $F \in \mathcal{F}$. Now let $F \subseteq \mathcal{G}_{\mu}^1 \cup \mathcal{G}_{\mu}^2$, $F \in \mathcal{F}$ and notice that $F = (F \cap \mathcal{G}_{\mu}^1) \cup (F \setminus \mathcal{G}_{\mu}^1)$ where $F \cap \mathcal{G}_{\mu}^1 \subseteq \mathcal{G}_{\mu}^1$ and $F \setminus \mathcal{G}_{\mu}^1 \subseteq \mathcal{G}_{\mu}^2$ are disjoint. Hence,

$$\nu(F) = \nu(F \cap \mathcal{G}_{\mu}^1) + \nu(F \setminus \mathcal{G}_{\mu}^1) \geq \mu(F \cap \mathcal{G}_{\mu}^1) + \mu(F \setminus \mathcal{G}_{\mu}^1) = \mu(F), \quad (\text{IV.4.26})$$

which shows $\mu_{\mathcal{G}^1 \vee \mathcal{G}^2} = \mu_{\mathcal{G}^1} \vee \mu_{\mathcal{G}^2}$.

To prove the general statement (IV.4.25), notice that again one inequality is easy by monotonicity. Thus, we only need to show that if $\nu \succeq \bigvee_{i \in I} \mu_{\mathcal{G}^i}$ then $\nu(F) \geq \mu(F)$ for every $F \subseteq (\bigvee_{i \in I} \mathcal{G}^i)_{\mu}$, $F \in \mathcal{F}$. Notice that $\nu \succeq \bigvee_{i \in I} \mu_{\mathcal{G}^i}$ implies that $\nu(F) \geq \mu(F)$ for every $F \subseteq \mathcal{G}_{\mu}^i$, $i \in I$, $F \in \mathcal{F}$. By Proposition IV.4.2, let $J \subseteq I$ be countable and such that $(\bigvee_{i \in I} \mathcal{G}^i)_{\mu} = \bigcup_{i \in J} \mathcal{G}_{\mu}^i$ and let $(j_n)_{n \in \mathbb{N}}$ be an enumeration of J . Then, we can write an arbitrary $F \subseteq (\bigvee_{i \in I} \mathcal{G}^i)_{\mu} = \bigcup_{n \in \mathbb{N}} \mathcal{G}_{\mu}^{j_n}$, $F \in \mathcal{F}$ as the countable union of disjoint sets $F^n \subseteq \mathcal{G}_{\mu}^{j_n}$, $n \in \mathbb{N}$ defined as

$$F^1 := F \cap \mathcal{G}_{\mu}^{j_1}, \quad F^{n+1} := (F \cap \mathcal{G}_{\mu}^{j_{n+1}}) \setminus (\bigcup_{k=1}^n \mathcal{G}_{\mu}^{j_k}), \quad n \in \mathbb{N}. \quad (\text{IV.4.27})$$

Hence,

$$\nu(F) = \sum_{n=1}^{\infty} \nu(F^n) \geq \sum_{n=1}^{\infty} \mu(F^n) = \mu(F), \quad (\text{IV.4.28})$$

and the proof is complete. $\square$

IV.5 Conclusions

In this chapter we have studied a decomposition of measures introduced by Dellacherie [34, Ch. II, Sec. 2, Th. T14] for which we have first provided an original simpler proof. Although not well-known, this decomposition is straightforward in its statement and remarkably powerful: indeed, we have shown how to obtain several well-known (and some new potentially interesting) decompositions as simple corollaries of this decomposition. We have then proved an extension of this decomposition to vector measures. By an application of this extension to signed measures we have obtained the Jordan decomposition and by an application of this extension to the stochastic integral we have obtained a decomposition of semimartingales due to Bichteler. Finally, we have studied some regularity properties of this decomposition.

Our work has enhanced the strength of this decomposition which deserves to be better known and further studied. We also believe more (well-known or new) decompositions of measures and processes could be obtained as simple applications and we are currently working in this direction. For instance, recently we have been able to obtain the decomposition of a stopping time into its accessible part and totally inaccessible part as a simple corollary of Dellacherie's decomposition.

This page was intentionally left blank.

Bibliography

- [1] C. D. Aliprantis and K. C. Border. *Infinite dimensional analysis: a hitchhiker's guide*. Springer, 2006.
- [2] L. H. Alvarez. Singular stochastic control, linear diffusions, and optimal stopping: A class of solvable problems. *SIAM Journal on Control and Optimization*, 39(6):1697–1710, 2001.
- [3] B. Asri, S. Hamadène, and K. Oufdil. On the stochastic control-stopping problem. *arXiv preprint arXiv:2005.06789*, 2020.
- [4] R. Atar and A. Budhiraja. Singular control with state constraints on unbounded domain. *The Annals of Probability*, 34(5):1864–1909, 2006.
- [5] M. Avellaneda and M. D. Lipkin. A market-induced mechanism for stock pinning. *Quantitative Finance*, 3(6):417–425, 2003.
- [6] P. Baldi. *Stochastic Calculus*. Springer, 2017.
- [7] F. M. Baldursson. Singular stochastic control and optimal stopping. *Stochastics: An International Journal of Probability and Stochastic Processes*, 21(1):1–40, 1987.
- [8] J. Bather and H. Chernoff. Sequential decisions in the control of a spaceship. In *Fifth Berkeley Symposium on Mathematical Statistics and Probability*, volume 3, pages 181–207, 1967.
- [9] E. J. Baurdoux, N. Chen, B. A. Surya, and K. Yamazaki. Optimal double stopping of a Brownian bridge. *Advances in Applied Probability*, 47(4):1212–1234, 2015.
- [10] E. Bayraktar and S. Yao. Optimal stopping for non-linear expectations—Part II. *Stochastic processes and their applications*, 121(2):212–264, 2011.
- [11] R. Bellman. Dynamic programming. *Science*, 153(3731):34–37, 1966.
- [12] V. E. Beneš, L. A. Shepp, and H. S. Witsenhausen. Some solvable stochastic control problemst. *Stochastics: An International Journal of Probability and Stochastic Processes*, 4(1):39–83, 1980.
- [13] K. Bichteler. Stochastic integrators. *Bulletin of the American Mathematical Society*, 1(5):761–765, 1979.
- [14] K. Bichteler. Stochastic integration and L^p -theory of semimartingales. *The Annals of Probability*, 9(1):49–89, 1981.
- [15] K. Bichteler. *Stochastic integration with jumps*, volume 89. Cambridge University Press, 2002.

- [16] F. Boetius and M. Kohlmann. Connections between optimal stopping and singular stochastic control. *Stochastic Processes and their Applications*, 77(2):253–281, 1998.
- [17] V. S. Borkar. Optimal control of diffusion processes. *Pitman Research Notes in Math.*, 203, 1989.
- [18] B. Bouchard, R. Elie, and N. Touzi. Stochastic target problems with controlled loss. *SIAM Journal on Control and Optimization*, 48(5):3123–3150, 2010.
- [19] B. Bouchard and M. Nutz. Weak dynamic programming for generalized state constraints. *SIAM Journal on Control and Optimization*, 50(6):3344–3373, 2012.
- [20] B. Bouchard and M. Nutz. Arbitrage and duality in nondominated discrete-time models. *Annals of Applied Probability*, 25(2):823–859, 2015.
- [21] B. Bouchard and N. Touzi. Weak dynamic programming principle for viscosity solutions. *SIAM Journal on Control and Optimization*, 49(3):948–962, 2011.
- [22] W. M. Boyce. Stopping rules for selling bonds. *The Bell Journal of Economics and Management Science*, 1(1):27–53, 1970.
- [23] X. Chen and F. Yi. A problem of singular stochastic control with optimal stopping in finite horizon. *SIAM Journal on Control and Optimization*, 50(4):2151–2172, 2012.
- [24] M. B. Chiarolla. Singular stochastic control of a singular diffusion process. *Stochastics: An International Journal of Probability and Stochastic Processes*, 62(1-2):31–63, 1997.
- [25] J. Claisse, D. Talay, and X. Tan. A pseudo-markov property for controlled diffusion processes. *SIAM Journal on Control and Optimization*, 54(2):1017–1029, 2016.
- [26] M. H. Davis and A. R. Norman. Portfolio selection with transaction costs. *Mathematics of operations research*, 15(4):676–713, 1990.
- [27] M. H. Davis and M. Zervos. A problem of singular stochastic control with discretionary stopping. *The Annals of Applied Probability*, 4(1):226–240, 1994.
- [28] T. De Angelis. A note on the continuity of free-boundaries in finite-horizon optimal stopping problems for one-dimensional diffusions. *SIAM Journal on Control and Optimization*, 53(1):167–184, 2015.
- [29] T. De Angelis and E. Ekström. The dividend problem with a finite horizon. *The Annals of Applied Probability*, 27(6):3525–3546, 2017.
- [30] T. De Angelis, G. Ferrari, and J. Moriarty. A nonconvex singular stochastic control problem and its related optimal stopping boundaries. *SIAM Journal on Control and Optimization*, 53(3):1199–1223, 2015.
- [31] T. De Angelis and Y. Kitapbayev. Integral equations for rost’s reversed barriers: existence and uniqueness results. *Stochastic Processes and their Applications*, 127(10):3447–3464, 2017.
- [32] T. De Angelis, N. Merkulov, and J. Palczewski. On the value of non-markovian dynkin games with partial and asymmetric information. *arXiv preprint arXiv:2007.10643*, 2020.
- [33] T. De Angelis and G. Peskir. Global C^1 regularity of the value function in optimal stopping problems. *Annals of Applied Probability*, 30(3):1007–1031, 2020.

- [34] C. Dellacherie. *Capacités et processus stochastiques*, volume 357. Springer, 1972.
- [35] J. Detemple, Y. Kitapbayev, and L. Zhang. American option pricing under stochastic volatility models via Picard iterations. CMU conference slides, 2018.
- [36] C. Doléans-Dade. On the existence and unicity of solutions of stochastic integral equations. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 36(2):93–101, 1976.
- [37] N. Dunford and J. T. Schwartz. *Linear Operators. Part 1: General Theory*. New York Interscience, 1957.
- [38] A. Dvoretzky. Existence and properties of certain optimal stopping rules. In *Proc. Fifth Berkeley Symp. Math. Statist. Prob*, volume 1, pages 441–452, 1967.
- [39] I. Ekren, N. Touzi, and J. Zhang. Optimal stopping under nonlinear expectation. *Stochastic Processes and Their Applications*, 124(10):3277–3311, 2014.
- [40] E. Ekström and J. Vaicenavicius. Optimal stopping of a brownian bridge with an unknown pinning point. *Stochastic Processes and their Applications*, 130(2):806–823, 2020.
- [41] E. Ekström and H. Wanntorp. Optimal stopping of a Brownian bridge. *Journal of Applied Probability*, 46(1):170–180, 2009.
- [42] P. Ernst and L. Shepp. Revisiting a theorem of la shepp on optimal stopping. *Communications on Stochastic Analysis*, 9(3):419–423, 2015.
- [43] G. Fabbri, F. Gozzi, and A. Swiech. *Stochastic optimal control in infinite dimension*. Springer, 2017.
- [44] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*, volume 25. Springer Science & Business Media, 2006.
- [45] H. Föllmer. Optimal stopping of constrained Brownian motion. *Journal of Applied Probability*, 9(3):557–571, 1972.
- [46] A. Friedman. *Stochastic differential equations and applications I*. New York: Academic Press, 1975.
- [47] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*, volume 224. Springer, 2015.
- [48] K. Glover. Optimally stopping a brownian bridge with an unknown pinning time: A bayesian approach. *Stochastic Processes and their Applications*, 2020.
- [49] X. Guo and P. Tomecek. Connections between singular control and optimal switching. *SIAM Journal on Control and Optimization*, 47(1):421–443, 2008.
- [50] P. R. Halmos. *Measure theory*, volume 18. Springer, 2013.
- [51] U. G. Haussmann and W. Suo. Singular optimal stochastic controls I: Existence. *SIAM Journal on Control and Optimization*, 33(3):916–936, 1995.
- [52] U. G. Haussmann and W. Suo. Singular optimal stochastic controls II: Dynamic programming. *SIAM Journal on Control and Optimization*, 33(3):937–959, 1995.

- [53] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes*. North-Holland, 1989.
- [54] S. Jacka. Avoiding the origin: a finite-fuel stochastic control problem. *The Annals of Applied Probability*, 12(4):1378–1389, 2002.
- [55] S. D. Jacka. Optimal stopping and the american put. *Mathematical Finance*, 1(2):1–14, 1991.
- [56] J. Jacod and A. Shiryaev. *Limit theorems for stochastic processes*, volume 288. Springer Science & Business Media, 2013.
- [57] P. Jaillet, D. Lamberton, and B. Lapeyre. Variational inequalities and the pricing of american options. *Acta Applicandae Mathematica*, 21(3):263–289, 1990.
- [58] M. Jeannin, G. Iori, and D. Samuel. Modeling stock pinning. *Quantitative Finance*, 8(8):823–831, 2008.
- [59] O. Kallenberg. *Foundations of modern probability*. Springer Science & Business Media, 2006.
- [60] I. Karatzas. The monotone follower problem in stochastic decision theory. *Applied Mathematics and Optimization*, 7(1):175–189, 1981.
- [61] I. Karatzas. A class of singular stochastic control problems. *Advances in Applied Probability*, 15(2):225–254, 1983.
- [62] I. Karatzas, D. Ocone, H. Wang, and M. Zervos. Finite-fuel singular control with discretionary stopping. *Stochastics: An International Journal of Probability and Stochastic Processes*, 71(1-2):1–50, 2000.
- [63] I. Karatzas and S. E. Shreve. Connections between optimal stopping and singular stochastic control I. monotone follower problems. *SIAM Journal on Control and Optimization*, 22(6):856–877, 1984.
- [64] I. Karatzas and S. E. Shreve. Connections between optimal stopping and singular stochastic control II. reflected follower problems. *SIAM Journal on Control and Optimization*, 23(3):433–451, 1985.
- [65] I. Karatzas and S. E. Shreve. *Methods of mathematical finance*, volume 39. Springer, 1998.
- [66] I. Karatzas and I.-M. Zamfirescu. Martingale approach to stochastic control with discretionary stopping. *Applied Mathematics and Optimization*, 53(2):163–184, 2006.
- [67] I. J. Kim. The analytic valuation of american options. *The Review of Financial Studies*, 3(4):547–572, 1990.
- [68] N. V. Krylov. *Controlled diffusion processes*, volume 14. Springer Science & Business Media, 2008.
- [69] T. Leung, J. Li, and X. Li. Optimal timing to trade along a randomized Brownian bridge. *International Journal of Financial Studies*, 6(3):75, 2018.
- [70] R. C. Merton. Lifetime portfolio selection under uncertainty: The continuous-time case. *The review of Economics and Statistics*, pages 247–257, 1969.
- [71] V. S. Mikhalevich. Sequential bayes solutions and optimal methods of statistical acceptance control. *Teoriya Veroyatnostei i ee Primeneniya*, 1(4):437–465, 1956.

- [72] H. Morimoto. A singular control problem with discretionary stopping for geometric brownian motions. *SIAM journal on control and optimization*, 48(6):3781–3804, 2010.
- [73] H. Morimoto. *Stochastic control and mathematical modeling*. Cambridge University Press, 2010.
- [74] M. Nisio. Lectures on stochastic control theory, isi lect. *Notes*, 9, 1981.
- [75] M. Ondreját. *Uniqueness for stochastic evolution equations in Banach spaces*, volume 426. Dissertationes Math., 2004.
- [76] K. R. Parthasarathy. *Probability measures on metric spaces*, volume 352. American Mathematical Soc., 2005.
- [77] J. Pederson and G. Perkir. Solving non-linear optimal stopping problems by the method of time-change. *Stochastic analysis and applications*, 18(5):811–835, 2000.
- [78] G. Peskir. On the american option problem. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 15(1):169–181, 2005.
- [79] G. Peskir and A. Shiryaev. *Optimal stopping and free-boundary problems*. Springer, 2006.
- [80] P. E. Protter. *Stochastic integration and differential equations*. Springer, 2005.
- [81] W. Rudin. *Real and complex analysis*. Tata McGraw-Hill education, 2006.
- [82] L. A. Shepp. Explicit solutions to some problems of optimal stopping. *The Annals of Mathematical Statistics*, 40(3):993, 1969.
- [83] J. L. Snell. Applications of martingale system theorems. *Transactions of the American Mathematical Society*, 73(2):293–312, 1952.
- [84] A. Wald. *Sequential analysis*. Wiley, New York, 1947.
- [85] J. Yong and X. Y. Zhou. *Stochastic controls: Hamiltonian systems and HJB equations*, volume 43. Springer Science & Business Media, 1999.
- [86] M. Yor. Quelques interactions entre mesures vectorielles et intégrales stochastiques. In *Séminaire de Théorie du Potentiel Paris, No. 4*, pages 264–281. Springer, 1979.
- [87] H. Zhu. Generalized solution in singular stochastic control: the nondegenerate problem. *Applied Mathematics and Optimization*, 25(3):225–245, 1992.