Stopper vs. Singular-Controller Games

Andrea Bovo

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds
School of Mathematics

October 2022

The fact that you have a place where you can return home, will lead you to happiness.

Kaworu Nagisa

to my parents,

Understanding 100% of everything is impossible. That's why we spend all our lives trying to understand the thinking of others. That's what makes life so interesting.

Ryoji Kaji

Acknowledgements

Several things have happened during this journey, including a pandemic and a war. Thus, I couldn't be more grateful to the world for trying so hard to make the PhD more difficult than it already is. Luckily, my supervisors helped me more than I expected and deserved: in particular Dr Tiziano De Angelis, who always had a good idea, suggestion and word, about mathematics and not only...your patience and passion lead me through difficult hills; Dr Elena Issoglio, who has always been on my side and understood my difficulties, your calm was an important ingredient to overcome the dark times; Dr Jan Palczewski who has always been interested and enthusiastic about my research, even when I was far from my goals. I want to say 'Thank you' to all of you because I felt part of a family.

A great thanks goes to my examiners, Prof Amanda Turner and Prof Harry Zheng, who read and appreciated this dissertation. You gave me food for thought about this work and about future research.

I am also very grateful to all the Probability and Financial Mathematics Group who organised meetings and social events which helped to open my interest on new topics. I want to say thank you to Jason, Cheng, Nikita, Jacob S. and Luis Mario who shared similar topics and made mathematics a nice thing to talk about. A particular thanks goes to Jacob C.R. who explained to me a lot of things about PhD life and has always been available for a chat. I want to express the gratitude to all PhD students in Leeds who created a nice environment where I felt accepted. Thank you to: Celeste, Dario, Fiona, Lorenzo, Luca, Matteo Sp., Matteo St., Pietro, Rosario, Rukia. A final thanks goes to the School of Mathematics which has been an incredible environment in which to grow.

Outside these mathematical acknowledgements, there are definitely some human acknowledgements that I would like to give to those who helped in the writing of this thesis.

A great thanks goes to my parents, Nadia and Pietro, who always give me the possibility to pursue all my strange passions, such as mathematics. There are things that you cannot learn by studying and I really appreciate that you have passed me some of them. Thank you for how you raised me. A thanks goes to my siblings, Federica and Daniele, who always bring me back to the reality of life. I want to say thank you to my niece and nephew, Anastasia and Leonardo, for teaching me the importance of little things. A thanks goes also to my brother's wardrobe which still dresses me every day.

A thanks to my old friends Federico, Giorgia and Nicolò for their friendship despite everything. You were always there for a meeting, beer, or whatever, even during the lockdown. A thanks to Andrea and Valerian for their point of view and nice conversations, and for very nice Holidays together. Another thanks goes to all of my friends who have shared their time with me: 'Letterati', 'Hai paura del Bugo', 'Emotions' and many others.

A big thanks goes to my second family in Leeds: 'Casa Italia', a safe harbour in a seaside town. There are so many things that I would like to mention: food, cinema, beers, etc., but it would be never enough. I just want to express my gratitude just reminding one thing to each of you: to Francesco, for the great tennis matches; to Giovanni, for losing in Warzone games; to Gabriele, for the nice chats about comics, and to 'El Diablo' Leonardo, for the great music. I say without doubts you have completed me in all my favourite interests and you have made my journey in Leeds delicious.

I have always said I will never use this idiomatic form, but here we go. Last but not least, I want to say thank you to Rebecca who has always tollerated my being myself. I am not as easy as I could seem and you have always pampered me more than what I deserve. I really appreciate you because you listen to my deliriums and you make our time together always lovely. Finally, thank you because you push me up every time I am trying to destroy myself. I want to say thank you also to your family who always made me feel like a part of it. At this point, I should also say thank you to the pandemic because we wouldn't be here together without it...maybe.

Abstract

We study a class of zero-sum games between a singular-controller and a stopper over finite-time horizon. In the first part of the thesis, the underlying process is a multi-dimensional (locally non-degenerate) controlled stochastic differential equation (SDE) evolving in an unbounded domain. We prove that such games admit a value and provide an optimal strategy for the stopper. The value of the game is shown to be the maximal solution, in a suitable Sobolev class, of a variational inequality of 'min-max' type with obstacle constraint and gradient constraint. Although the variational inequality and the game are solved on an unbounded domain we do not require boundedness of either the coefficients of the controlled SDE or of the cost functions in the game. In the second part we extend the result to two classes of games that may be referred to as "degenerate" cases: (i) we study games with a constrained control direction and (ii) games with degenerate diffusion coefficient. Through approximation procedures, we obtain the existence of the value of the game and the optimal strategy for the stopper.

Contents

List of abbreviations

- **PDE**: Partial Differential Equation
- **SDE**: Stochastic Differential Equation
- **BSDE**: Backward Stochastic Differential Equation
- **ZSG**: Zero-Sum Game
- **HJB**: Hamilton-Jacobi-Bellmann
- $\mathbb{R}^{d+1}_{0,T}$: the general state space $[0,T] \times \mathbb{R}^d$
- B_m: the d-dimensional ball of radius m centred in 0
- $\mathcal{O}_m$: the parabolic cylinder $[0,T) \times B_m$
- $\partial_P \mathcal{O}_m$: the parabolic boundary of $\mathcal{O}_m$, i.e., the set defined as $([0,T) \times \partial B_m) \cup (\{T\} \times \overline{B}_m)$
- $\| \cdot \|_m$: the supremum norm on the set $\overline{\mathcal{O}}_m$
- ξ_m: a C^∞ cut-off function which is 1 inside B_{m-1} and it is 0 outside B_m
- **càglàd**: (continue à gauche, limite à droite)
- **càdlàg**: (continue à droite, limite à gauche)

Chapter 1

Introduction

The topic of this thesis is zero-sum (stochastic) games (ZSGs) between a stopper and a controller. Here zero-sum refers to a two-player game in which the players exchange a *payoff* when the game ends. The first player (the *controller*) uses a control that affects the dynamics of the underlying process and the second player (the *stopper*) decides when the game ends by choosing a stopping time. The controller pays the stopper, so the controller tries to minimise the payoff while the stopper tries to maximise it. This game is a combination of two well-known and widely studied problems: an optimal stopping problem and a singular control problem.

Optimal stopping was initially introduced by Wald in his theory of the sequential probability ratio test in 1945 (see [66]). Few years later, the problem was studied from a more probabilistic perspective by Snell (see [60]) who characterised the solution using martingale theory and showed that the value process related to the optimal stopping problem is the smallest supermartingale that dominates the payoff process (the so-called Snell's envelope). Snell's contribution led to what we know nowadays as the martingale approach. In the same years, problems with Markovian structure were studied by several authors, including, in particular, Dynkin [22], Shiryaev [58] and McKean [53]. This led to the so-called Markovian approach which has strong ties with the theory of partial differential equations. Optimal stopping problems are encountered in several research areas and they have received wide attention. For example, in Wald's sequential testing problem an optimiser wants to perform a hypothesis test on a stochastic system. After a period of observation of the system's dynamics, the observer should decide whether to keep or reject the null hypothesis. The time at which the decision is made should be chosen as an optimal stopping time in a suitable optimal stopping problem. Another famous application of optimal stopping theory is in mathematical finance, where it is used to compute the price of American options. An introduction to all these aspects can be found in the monographs by Shiryaev [59], El Karoui [23] and Peskir and Shiryaev [56]. In particular Chapters 1, 6, 7 and 8 in [56] contain results obtained using both the martingale and the Markovian approach, applications in mathematical statistics, in mathematical finance and in financial engineering. Other applications in mathematical finance from a more economic perspective can be found in the monograph by Dixit and Pindyck [21]. Finally, in [56] it is also shown an important connection between optimal stopping problems and free boundary problems. Free boundary problems concern the study of solutions of PDEs inside domains that must be determined as part of the solution itself. It turns out that the value function V of an optimal stopping problem is the solution U of a free boundary problem in which the free boundary is determined via an obstacle constraint on U.

Similarly, free boundary problems arise in the study of singular control problems in which an optimiser tries to minimise (or maximise) in expectation a functional dependent on a singularly controlled stochastic process. The peculiarity of singular controls is that they are stochastic processes which are not absolutely continuous with respect to the Lebesgue measure as functions of time. In contrast the so-called 'classical' controls admit a density (Radon-Nikodym derivative) with respect to the Lebesgue measure as functions of time. One of the first contribution on singular stochastic control (SSC) problems was made by Bather and Chernoff [3] in 1967. Later on, Beneš et al. [7] solved explicitly a problem of singular control in 1980. Singular controls have found applications in several research areas such as aerospace engineering [4], mathematical finance [18] and [9], mathematical biology [2], and others. The value function V of a singular control problem is the solution U (in a suitable sense) of a free boundary problem in which the free boundary is determined via a gradient constraint on U itself. The existence of an optimal control is related to the solvability of a so-called Skorohod reflection problem. The Skorohod problem is the problem of keeping an (optimally) controlled dynamics in the domain defined by the free boundary. This problem can be solved in some special cases (see, e.g., [51], [46], [31]) but it is a difficult task for general multi-dimensional setting (see Remark 5.2 in [10]). Indeed, the existence of an optimal control in the multi-dimensional framework relies on the regularity of the value function, the free boundary and the direction of reflection at this boundary; there is no general theory to address the problem and we do not consider such question in this dissertation. To conclude, it is worth mentioning a connection between singular stochastic control and optimal stopping (see [40] and [41]). In dimension one, such connection can be easily stated: it turns out that under suitable assumptions the derivative with respect to the controlled state variable of the value function of a singular control problem is the value function of a suitable optimal stopping problem; consequently, the free boundary for the singular control problem coincides with the optimal stopping boundary in the associated stopping problem.

The aim of this dissertation is to combine optimal stopping and singular control in a ZSG. Here, we introduce the framework used in Chapter 3 which leads to our main result (variations of this formulation will be considered in Chapter 4): we consider a class of ZSGs on a finite-time horizon $[0,T]$ between a controller and a stopper. The underlying stochastic dynamics $X^{[n,\nu]}$ is given by a d-dimensional, singularly controlled, stochastic differential equation of the form

$$\mathrm{d}X_t^{[n,\nu]} = b(X_t^{[n,\nu]})\mathrm{d}t + \sigma(X_t^{[n,\nu]})\mathrm{d}W_t + n_t\mathrm{d}\nu_t, \tag{1.1}$$

where W is a d'-dimensional Brownian motion (with $d \leq d'$) and the control pair $(n_t, \nu_t)_{t\in[0,T]}$ is given by a unitary vector $n_t(\omega) \in \mathbb{R}^d$ and a real valued, right-continuous, increasing process $\nu_t(\omega)$. The resulting dynamics is 'singular' because the mapping $t \mapsto \nu_t(\omega)$ need not be absolutely continuous with respect to the Lebesgue measure; it does not need to be even continuous. The stopper chooses a stopping time τ deciding when the game ends and she receives

$$\begin{aligned} e^{-r\tau}g(t+\tau, X_\tau^{[n,\nu]}) + \int_0^\tau e^{-rs}h(t+s, X_s^{[n,\nu]})\,\mathrm{d}s \\ + \int_{[0,\tau]} e^{-rs}f(t+s, X_s^{[n,\nu]}) \circ\mathrm{d}\nu_s \end{aligned} \tag{1.2}$$

from the controller. Here $g(t,x)$ is the terminal payoff, $h(t,x)$ is the running payoff and $f(t,x)$ is the cost per unit of control exerted. Using a blend of analytical and probabilistic techniques we prove that the game admits a value v which is the maximal solution in a suitable Sobolev space of the variational inequalities:

$$\begin{aligned}&\min\big\{\max\big\{\partial_t u+\mathcal{L}u-ru+h,g-u\big\},f-|\nabla u|_d\big\}=0,\\&\max\big\{\min\big\{\partial_t u+\mathcal{L}u-ru+h,f-|\nabla u|_d\big\},g-u\big\}=0,\end{aligned}\tag{1.3}$$

with terminal condition $u(T,x)=g(T,x)$; the above equations should be understood in the almost everywhere sense on $[0,T]\times\mathbb{R}^d$. Here, the operator $\mathcal{L}$ is the infinitesimal generator of the uncontrolled SDE, $r\geq 0$ is a constant discount rate and $|\cdot|_d$ is the Euclidean norm in $\mathbb{R}^d$. We are able to provide an optimal stopping rule for the stopper and we can derive an ϵ-optimal strategy for the controller for any $\epsilon>0$ (Remark 3.35).

The two variational problems in (1.3) have not received much attention in the literature and they pose a number of challenges. The first obvious one is that swapping the order of 'min' and 'max' is non-trivial and it relates in some sense to proving the equivalence between the so-called upper and lower value of the game. Secondly, a solution of the variational problem is subject to two hard constraints: an obstacle constraint $u\geq g$ and a gradient constraint $|\nabla u|_d\leq f$. Thirdly, we solve the problem on an *unbounded domain* but without imposing boundedness of the coefficients of the SDE or of the payoff functions, and without requiring uniform ellipticity of the matrix $\sigma\sigma^\top$ in the whole space. These seem important technical improvements even when compared to variational inequalities on unbounded domains for singular control problems (e.g., Chow et al. [17], Soner and Shreve [61, 62], Menaldi and Taksar [54] and Zhu [69]) or optimal stopping games (e.g., Friedman [28] and Stettner [63]). The two hard constraints characterise the free boundaries of the problem which can be defined as the boundaries of the sets $\{(t,x):u(t,x)>g(t,x)\}$ and $\{(t,x):|\nabla u(t,x)|_d<f(t,x)\}$. A priori, we do not know anything about these two sets, and their boundaries could be very irregular.

The study of controller-stopper ZSGs originates from work by Maitra and Sudderth [52] in 1996. A 'gambler' selects a conditional distribution $\varsigma=(\varsigma_0,\varsigma_1,\dots)$ from a suitable class for a discrete-time process $(X_n)_{n\in\mathbb{N}}\subset\mathcal{S}$ with $X_0=x$, and a stopper ends the game at a stopping time τ of her choosing. The stopper pays an amount $u(X_\tau)$ to the gambler, for some bounded function u. We say that the game *admits a value* if the following holds

$$v(x) = \sup_{\varsigma} \inf_{\tau} \mathsf{E}[u(X_\tau)|X_0 = x] = \inf_{\tau} \sup_{\varsigma} \mathsf{E}[u(X_\tau)|X_0 = x] \quad \text{for all } x \in \mathcal{S}.$$

It is proven in [52] that the game admits a value and an ϵ-optimal Markov family of strategies under general assumptions on the state space, the reward function, the class of stopping times and the class of admissible controls. We will give more details on this and next examples of problems in Chapter 2.

The above problem was later cast in a continuous-time infinite horizon framework by Karatzas and Sudderth [42] who consider a one-dimensional Itô diffusion in a interval whose drift and diffusion coefficients are chosen by the controller from a suitable class. They obtain (almost explicit) optimal strategies for both players using methods based on the general theory of one-dimensional linear diffusions. Weerasinghe [67] studied a similar problem, in which the underlying dynamics is a one-dimensional SDE whose diffusion coefficient is controlled and is allowed to vanish, and finds that the game admits a value that is not continuously differentiable as function of the initial state of the process.

Following those early contributions, the literature on controller-stopper ZSGs (and to some extent also nonzero-sum games) has grown steadily. A wide variety of methods has been deployed spanning, for example, martingale theory (Karatzas and Zamfirescu [44]), backward stochastic differential equations (e.g., Hamadène and Lepeltier [32], Hamadène [33], Choukroun et al. [16]) and solution of variational problems via viscosity theory (e.g., Bayraktar and Huang [5] and Bayraktar and Young [6]). A common denominator of those papers is that the controller uses so-called 'classical' controls, i.e., progressively measurable maps $(t,\omega) \mapsto \alpha_t(\omega)$ that enter the drift and diffusion coefficient, b and σ of the controlled SDE in the form

$$\mathrm{d}X_t^\alpha = b(X_t^\alpha, \alpha_t)\mathrm{d}t + \sigma(X_t^\alpha, \alpha_t)\mathrm{d}W_t.$$

From an analytical point of view, those games are connected to Hamilton-Jacobi-Bellman (HJB) equations with obstacle constraint but without gradient constraint, hence different from (1.3). A main difference across those contribution is that [5] does not require the diffusion coefficient to be uniformly non-degenerate in contrast to [33] and [44] where the condition is a crucial assumption to perform a Girsanov transformation of the probability measure.

Much less attention instead has been devoted to the study of games in which the controller can adopt singular controls as in (1.1). A notable contribution to this strand of the literature was given by Hernandez-Hernandez et al. [35] who consider a ZSG in which $X^{[n,\nu]}$ is real valued. They provide a general verification theorem and explicit optimal strategies for both players in specific examples. They also show that the value function of the game need not be smooth if the stopping payoff is not continuously differentiable. The methods in [35] rely crucially on the infinite horizon and one-dimensional set-up that allow to link the variational problem with ordinary differential equations and require an educated guess on the structure of the optimal strategies. In this thesis, instead, we develop a general theory for multi-dimensional state-dynamics and prove the existence and the variational characterisation of the value function. The finite time horizon and the dimensionality of the process lead us to consider PDEs and to use a mix of analytical and probabilistic methods. It appears that our study is the first one involving a multi-dimensional singularly controlled underlying process.

ZSGs of controller-stopper type have been motivated by several applications. In [43] authors show a connection with the prices of American put-options in the presence of an 'up-and-out' barrier with constraints on the short-selling of stock. In [6] the authors show another connection with a minimisation of lifetime ruin probability for an individual who can invest in a Black-Scholes financial market and the rate of consumption is stochastic. In [35] the authors explain numerous other important applications of ZSGs like the ones we consider here. Such applications include models for a central bank controlling exchange rates up to the time of a possible political veto and models for the control of inflation.

This thesis is structured as follows. In Section 1.1 we set out the notation. In Chapter 2, we give preliminary results and a brief review of the literature on ZSGs.

In Chapter 3, we solve the game (1.2) in the above framework and show the main result of this dissertation. Our method of proof builds upon penalisation techniques that address simultaneously the two hard constraints embedded in (1.3): $u \geq g$ and $|\nabla u|_d \leq f$. We find bounds on the Sobolev norm of the solution of the penalised PDE problem, uniformly with respect to the penalisation parameters, thanks to analytical techniques rooted in early work by Evans [24] and new probabilistic tricks developed *ad-hoc* in our framework. Indeed, it turns out that the co-existence of two hard constraints in (1.3), the 'min-max' structure of the problem, its parabolic nature and unboundedness of the domain make the use of purely analytical ideas as in Evans [24] (see also [68] and [39]) not sufficient to provide the necessary bounds (see also the references given in the previous paragraph and more recent work by Hynd [38] and Kelbert and Moreno-Franco [45], for comparison). In the process of obtaining our main result (Theorem 3.6) we also contribute a detailed proof of the existence and uniqueness of the solution for the penalised problem (Theorem 3.21 for bounded domain, and Theorem 3.24 and Proposition 3.25 for unbounded domain). Finally, the existence of an optimal stopping time τ_* is interesting in its own right as it may enable free-boundary techniques for the study of the optimal strategy of the controller.

In Chapter 4 we use the results from Chapter 3 to extend our analysis to two special classes of ZSGs that are not covered directly by Theorem (3.6): (i) the class in which the minimiser is allowed to use controls in selected directions of the state space and (ii) the class in which the diffusion of the underlying process is allowed to be degenerate. We approximate those problems with games that satisfy conditions of Chapter 3. Through the approximations we show that those two types of games admit a value and provide an optimal strategy for the stopper. In the case where the controller affects selected directions (Section 4.1) we allow the underlying process of the approximated game to be affected by a control in all the direction with a weight in the direction not originally affected. In the limit we let the weight go to zero and recover the original problem formulation. In the degenerate case (Section 4.2), we keep the same payoff as in the original game and we add a non-degenerate independent diffusion to the original underlying process in order to make it a non-degenerate process. We show that the value functions of the approximated games converge uniformly to the value of the initial game as the parameter goes to zero.

Finally, a technical appendix completes the thesis. We give a probabilistic proof of the well-known maximum principle for PDEs (Lemma B.3). We give a solution of [47, Exercise 10.1.14] which proves a type of Morrey's inequality for parabolic spaces (Lemma A.4). We present a possible choice of cut-off functions used in Chapter 3 with their properties (Section B.1) and we also present a possible choice of family of penalty functions based on this family of cut-off functions (Section B.2). We prove stability results for PDE (Lemma B.1 and Section B.5).

We prove the existence and the uniqueness of a non-explosive solution of the ϵ-optimal controlled SDE (Section B.6). Finally, we extend [20, Lemma 5.1] to semimartingales with jumps (Lemma C.1).

1.1 Notation

We conclude this introduction giving notation which is used in the whole dissertation related to $\mathbb{R}^d$, in particular we introduce parabolic Hölder spaces and their embedding in parabolic Sobolev spaces. Fix $d, d' \in \mathbb{N}$ and $T \in (0, \infty)$. Given $u \in \mathbb{R}^d$ we let $|u|_d$ be its euclidean norm. For vectors $u, v \in \mathbb{R}^d$ their scalar product is denoted by $\langle u, v \rangle$. Given a matrix $M \in \mathbb{R}^{d \times d'}$, with entries M_{ij}, $i = 1, \dots d$, $j = 1, \dots d'$, we denote its norm by

$$|M|_{d\times d'} \coloneqq \Big(\sum_{i=1}^{d} \sum_{j=1}^{d'} M_{ij}^2 \Big)^{1/2}$$

and, if $d = d'$, we let $\mathrm{tr}(M) \coloneqq \sum_{i=1}^{d} M_{ii}$.

The d-dimensional *open* ball centred in 0 with radius m is denoted by B_m and the general state space in this thesis is going to be

$$\mathbb{R}^{d+1}_{0,T} := [0, T] \times \mathbb{R}^d.$$

Finally, given a bounded set A we denote by $\overline{A}$ its closure.

For a smooth function $f : \mathbb{R}^{d+1}_{0,T} \to \mathbb{R}$ we denote its partial derivatives by $\partial_t f$, $\partial_{x_i} f$, $\partial_{tx_j} f$, $\partial_{x_i x_j} f$, for $i, j = 1, \dots d$. We will also use $f_t = \partial_t f$, $f_{x_i} = \partial_{x_i} f$, $f_{tx_i} = \partial_{tx_i} f$ and $f_{x_i x_j} = \partial_{x_i x_j} f$ to simplify long expressions. By ∇f we intend the spatial gradient, i.e., $\nabla f = (\partial_{x_1} f, \dots \partial_{x_d} f)$, and by $D^2 f$ the spatial Hessian matrix with entries $\partial_{x_i x_j} f$ for $i, j = 1, \dots d$.

As usual $C^\infty(\mathbb{R}^{d+1}_{0,T})$ is the space of functions with infinitely many continuous derivatives and $C^\infty_c(\mathbb{R}^{d+1}_{0,T})$ is the subset of $C^\infty(\mathbb{R}^{d+1}_{0,T})$ of functions with compact support. Continuous functions on a domain D are denoted by $C(D)$. For an open bounded set $\mathcal{O} \subset \mathbb{R}^{d+1}_{0,T}$ we let $C^0(\overline{\mathcal{O}})$ be the space of continuous functions

$f : \overline{\mathcal{O}} \to \mathbb{R}$ equipped with the supremum norm

$$\|f\|_{C^0(\overline{\mathcal{O}})} := \sup_{(t,x)\in\overline{\mathcal{O}}} |f(t,x)|. \tag{1.4}$$

Analogously, $C^0(\mathbb{R}^{d+1}_{0,T})$ is the space of bounded and continuous functions $f : \mathbb{R}^{d+1}_{0,T} \to \mathbb{R}$ equipped with the norm $\|f\|_\infty := \|f\|_{C^0(\mathbb{R}^{d+1}_{0,T})}$ as in (1.4) but with $\overline{\mathcal{O}}$ replaced by $\mathbb{R}^{d+1}_{0,T}$.

For bounded $\mathcal{O} \subset \mathbb{R}^{d+1}_{0,T}$, we consider the following function spaces:

- $C^{0,1}(\overline{\mathcal{O}})$ be the class of continuous functions with $\partial_{x_i} f \in C(\overline{\mathcal{O}})$ for $i = 1, \dots d$;
- $C^{1,2}(\overline{\mathcal{O}})$ be the class of continuous functions with $\partial_t f, \partial_{x_i} f, \partial_{x_i x_j} f \in C(\overline{\mathcal{O}})$ for $i, j = 1, \dots d$;
- $C^{1,3}(\overline{\mathcal{O}})$ be the class of continuous functions with

$$\partial_t f,\ \partial_{x_i} f,\ \partial_{x_i x_j} f,\ \partial_{x_i x_j x_k} f,\ \partial_{t x_i} f \in C(\overline{\mathcal{O}})$$

  for $i, j, k = 1, \dots d$ (notice the mixed derivatives $\partial_{t x_i} f$).

The above definitions extend obviously to continuously differentiable functions on $\mathbb{R}^{d+1}_{0,T}$.

Since we are studying the finite time horizon problem, we will deal with the so-called parabolic spaces. Let $\mathrm{d}(z_1, z_2) = (|t-s| + |x-y|^2_d)^{\frac{1}{2}}$ be the parabolic distance between points $z_1 = (t,x)$ and $z_2 = (s,y)$ in $\mathbb{R}^{d+1}_{0,T}$. For a fixed $\alpha \in (0,1)$ and a continuous function $f : \overline{\mathcal{O}} \to \mathbb{R}$ we set (see [29, p. 61])

$$\|f\|_{C^\alpha(\overline{\mathcal{O}})} := \|f\|_{C^0(\overline{\mathcal{O}})} + \sup_{\substack{z_1,z_2\in\overline{\mathcal{O}} \\ z_1\neq z_2}} \frac{|f(z_1) - f(z_2)|}{\mathrm{d}^\alpha(z_1,z_2)}.$$

We say that $f \in C^\alpha(\overline{\mathcal{O}})$ if $f \in C^0(\overline{\mathcal{O}})$ and $\|f\|_{C^\alpha(\overline{\mathcal{O}})} < \infty$. We work with the following norms, defined for sufficiently smooth functions f:

$$\|f\|_{C^{0,1,\alpha}(\overline{\mathcal{O}})} := \|f\|_{C^{\alpha}(\overline{\mathcal{O}})} + \sum_{i=1}^{d} \|\partial_{x_i} f\|_{C^{\alpha}(\overline{\mathcal{O}})};$$

$$\|f\|_{C^{1,2,\alpha}(\overline{\mathcal{O}})} := \|f\|_{C^{0,1,\alpha}(\overline{\mathcal{O}})} + \|\partial_t f\|_{C^{\alpha}(\overline{\mathcal{O}})} + \sum_{i,j=1}^{d} \|\partial_{x_i x_j} f\|_{C^{\alpha}(\overline{\mathcal{O}})};$$

$$\|f\|_{C^{1,3,\alpha}(\overline{\mathcal{O}})} := \|f\|_{C^{1,2,\alpha}(\overline{\mathcal{O}})} + \sum_{i=1}^{d} \|\partial_{t x_i} f\|_{C^{\alpha}(\overline{\mathcal{O}})} + \sum_{i,j,k=1}^{d} \|\partial_{x_i x_j x_k} f\|_{C^{\alpha}(\overline{\mathcal{O}})}.$$

For $(j,k) \in \{(0,0);(0,1);(1,2);(1,3)\}$ and bounded $\mathcal{O}$ let us define

$$C^{j,k,\alpha}(\overline{\mathcal{O}}) := \big\{ f \in C^{j,k}(\overline{\mathcal{O}}) \big|\, \|f\|_{C^{j,k,\alpha}(\overline{\mathcal{O}})} < \infty \big\},$$
$$C^{j,k,\alpha}_{\ell oc}(\mathbb{R}^{d+1}_{0,T}) := \{ f \in C^{j,k}(\mathbb{R}^{d+1}_{0,T}) \big|\, f \in C^{j,k,\alpha}(\overline{\mathcal{O}}) \text{ for all bounded } \mathcal{O} \subset \mathbb{R}^{d+1}_{0,T} \}.$$

For B and B' open balls in $\mathbb{R}^d$ and $S \in [0,T)$, let $\mathcal{O}_B := [0,T) \times B$ and $\mathcal{O}_{S,B'} := [0,S) \times B'$. We denote $C^{j,k,\alpha}_{Loc}(\mathcal{O}_B)$ the class of functions $f \in C(\overline{\mathcal{O}}_B)$ such that $f \in C^{j,k,\alpha}(\overline{\mathcal{O}}_{S,B'})$ for all $S < T$ and B' such that $\overline{B'} \subset B$. Finally, we let

$$C^{j,k,\alpha}_{Loc}(\mathbb{R}^{d+1}_{0,T}) := \big\{ f \in C(\mathbb{R}^{d+1}_{0,T}) \,\big|\, f \in C^{j,k,\alpha}_{Loc}(\mathcal{O}_B) \text{ for all open balls } B \subset \mathbb{R}^d \big\}.$$

Notice that, the derivatives of functions in $C^{j,k,\alpha}_{\ell oc}(\mathbb{R}^{d+1}_{0,T})$ are Hölder continuous on $\overline{\mathcal{O}}_B$ for any ball $B \subset \mathbb{R}^d$. Instead, the derivatives of functions in $C^{j,k,\alpha}_{Loc}(\mathcal{O}_B)$ need not be continuous along the parabolic boundary of $\mathcal{O}_B$ and derivatives of functions in $C^{j,k,\alpha}_{Loc}(\mathbb{R}^{d+1}_{0,T})$ may be discontinuous at T.

To simplify long formulae, sometimes we use the notations:

$$\|\nabla f\|_{C^0(\overline{\mathcal{O}})} := \Big(\sum_{i=1}^{d} \|f_{x_i}\|^2_{C^0(\overline{\mathcal{O}})} \Big)^{\frac{1}{2}} \text{ and } \|D^2 f\|_{C^0(\overline{\mathcal{O}})} := \Big(\sum_{i,j=1}^{d} \|f_{x_i x_j}\|^2_{C^0(\overline{\mathcal{O}})} \Big)^{\frac{1}{2}}. \quad (1.5)$$

For $p \in [1,\infty]$ we say that $f \in L^p_{\ell oc}(\mathbb{R}^{d+1}_{0,T})$ if $f \in L^p(\mathcal{O})$ for any bounded $\mathcal{O} \subset \mathbb{R}^{d+1}_{0,T}$. Denote by $\mathcal{D}^{1,2}_{\mathcal{O}}$ the class of functions $f \in L^p_{\ell oc}(\mathbb{R}^{d+1}_{0,T})$ whose partial derivatives $\partial_t f$, $\partial_{x_i} f$, $\partial_{x_i x_j} f$ exist in the weak sense on $\mathcal{O}$, for $i,j = 1, \dots d$, and let

$$\|f\|_{W^{1,2,p}(\mathcal{O})} \coloneqq \|f\|_{L^p(\mathcal{O})} + \|\partial_t f\|_{L^p(\mathcal{O})} + \sum_{i=1}^{d} \|\partial_{x_i} f\|_{L^p(\mathcal{O})} + \sum_{i,j=1}^{d} \|\partial_{x_i x_j} f\|_{L^p(\mathcal{O})}.$$

Then we define $W^{1,2,p}(\mathcal{O}) \coloneqq \{f \in \mathcal{D}^{1,2}_{\mathcal{O}} \mid \|f\|_{W^{1,2,p}(\mathcal{O})} < \infty\}$ and

$$W^{1,2,p}_{\ell oc}(\mathbb{R}^{d+1}_{0,T}) \coloneqq \left\{ f \in L^p_{\ell oc}(\mathbb{R}^{d+1}_{0,T}) \middle| f \in W^{1,2,p}(\mathcal{O}),\, \forall\, \mathcal{O} \subseteq \mathbb{R}^{d+1}_{0,T},\, \mathcal{O} \text{ bounded} \right\}.$$

The first non-standard but classical result we present is a compact Sobolev embedding for parabolic spaces. For $\alpha = 1 - \frac{d+2}{p}$ and $p > d+2$, and for any bounded $\mathcal{O} \subset \mathbb{R}^{d+1}_{0,T}$, we have

$$C^{1,2}(\overline{\mathcal{O}}) \subset W^{1,2,p}(\mathcal{O}) \hookrightarrow C^{0,1,\alpha}(\overline{\mathcal{O}}), \tag{1.6}$$

where the first inclusion is obvious and the second one is the compact Sobolev embedding. This result can be viewed without proof in [26, eq. (E.9)] or solving [47, Exercise 10.1.14]. In this dissertation we give the solution of [47, Exercise 10.1.14] in Appendix A.1 using hints from the book itself.

Chapter 2

Literature Review

This chapter provides a theoretical background and presents main results on ZSGs from the contributions provided in the introduction. We review those contributions, starting from the discrete time case and then moving on the continuous time case. In the case of 'classical' controls we consider both the one-dimensional and the multi-dimensional settings. Finally, we present a class of ZSGs with singular controls in the one dimensional case. Since this dissertation is based on the case where a controller pays a stopper, i.e., the controller is a minimiser of an objective function and the stopper is a maximiser, we introduce some concepts on ZSG using this set up. This framework is not the only possible: the controller can be a maximiser and the stopper can be a minimiser, both players can stop and control at the same time the process, the two players can only stop (in this particular framework the game is called Dynkin game).

2.1 Preliminary Concepts

In a zero-sum game two players exchange a payoff at the of the game. The player who receives the payoff aims at maximising her winnings. The player who pays the payoff aims at minimising her costs. Informally we say that each player selects a strategy in order to optimise their performance. In our framework, a player, called stopper, decides when the game ends choosing a stopping time τ from a suitable admissible class $\mathcal{T}$ and a player, called controller, chooses a control ν from an admissible class $\mathcal{A}$. In general, the admissible class $\mathcal{T}$ is composed by all stopping times, but sometimes the class that we consider is a subset of $\mathcal{T}$, in those cases we will specify the particular subset of $\mathcal{T}$. The class $\mathcal{A}$ and how the control influence the dynamics may vary depending on the specific setup. The game's payoff is a functional of the stopping time and the underlying process. For example, take two measurable functions g and h representing the terminal cost and the running cost, respectively, then the expected payoff with underlying process X^ν starting from x can be defined as

$$\mathcal{J}_x(\tau,\nu) \coloneqq \mathsf{E}_x\left[g(\tau, X^\nu_\tau) + \int_0^\tau h(s, X^\nu_s)\,\mathrm{d}s\right]. \tag{2.1}$$

We can associate to $\mathcal{J}_x(\tau,\nu)$ the *lower* and the *upper* value of the game, defined respectively by

$$\underline{v}(x) \coloneqq \sup_{\tau\in\mathcal{T}}\inf_{\nu\in\mathcal{A}} \mathcal{J}_x(\tau,\nu) \quad \text{and} \quad \overline{v}(x) \coloneqq \inf_{\nu\in\mathcal{A}}\sup_{\tau\in\mathcal{T}} \mathcal{J}_x(\tau,\nu), \tag{2.2}$$

so that $\underline{v}(x) \le \overline{v}(x)$. The lower and upper value functions are always well-defined and if equality holds for all x then we say that the game admits a *value*

$$v(x) \coloneqq \underline{v}(x) = \overline{v}(x).$$

Moreover, if we find a pair $(\tau_*,\nu_*) \in \mathcal{T}\times\mathcal{A}$ such that

$$\mathcal{J}_x(\tau,\nu_*) \le \mathcal{J}_x(\tau_*,\nu_*) \le \mathcal{J}_x(\tau_*,\nu), \tag{2.3}$$

for all $(\tau,\nu) \in \mathcal{T}\times\mathcal{A}$ then this pair is called *saddle point* of the game and it implies the existence of the value of the game. Indeed, considering the first inequality in (2.3) and taking the supremum on the class of stopping times, we get

$$\overline{v}(x) \le \sup_{\tau\in\mathcal{T}} \mathcal{J}_x(\tau,\nu_*) \le \mathcal{J}_x(\tau_*,\nu_*);$$

considering the second inequality in (2.3) and taking the infimum on the class of controls, we get

$$\mathcal{J}_x(\tau_*,\nu_*) \le \inf_{\nu\in\mathcal{A}} \mathcal{J}_x(\tau_*,\nu) \le \underline{v}(x).$$

It follows that $\underline{v}(x) = \overline{v}(x)$ and

$$\mathcal{J}_x(\tau_*, \nu_*) = v(x) = \underline{v}(x) = \overline{v}(x).$$

In general, the optimal strategies may not exist and it is only possible to provide a weaker concept. Assume that the value of the game exists, then η-optimal strategies $\tau_\eta \in \mathcal{T}$ and $\nu_\eta \in \mathcal{A}$, for $\eta > 0$ are strategies such that

$$\inf_{\nu \in \mathcal{A}} \mathcal{J}_x(\tau_\eta, \nu) > v(x) - \eta, \quad \text{and} \quad \sup_{\tau \in \mathcal{T}} \mathcal{J}_x(\tau, \nu_\eta) < v(x) + \eta.$$

2.2 Stopper vs. Controller: Discrete Time

One of the first contributions in this field is given by Maitra and Sudderth in [52] who studied a class of ZSGs in a discrete time framework. In this setting, the underlying process is a discrete stochastic process with values in a space S and a payoff of the following form

$$\mathcal{J}_x(\tau, \nu) := \mathsf{E}_x[g(X^\nu_\tau)],$$

where g is a bounded function. Differently from Section 2.1, the framework in [52] is presented with a controller who maximises and a stopper who minimises. We can recover the formulation from Section 2.1 replacing g with $-g$. Let S be a nonempty Borel subset of a Polish space and let $\mathcal{M}(S)$ be the collection of probability measures defined on the Borel subsets of S. The controller selects at each time a transition probability from a subset of $\mathcal{M}(S)$. Precisely, for each $y \in S$, the controller can choose a transition probability only from $\Gamma(y) \subset \mathcal{M}(S)$. The admissible class $\mathcal{A}$ is then composed by strategies ν described as follows: let $(X^\nu_n)_{n \in \mathbb{N}}$ be the underlying process starting from $X^\nu_0 := x_0$. The strategy ν is a sequence $(\nu_0, \nu_1, \ldots)$, where $\nu_0 \in \Gamma(x_0)$ and $\nu_n = \nu_n(x_0, \ldots, x_{n-1}) \in \Gamma(X^\nu_n)$ with $x_m = X^\nu_m$ for $1 \leq m < n$. In other words, at each time n, the controller uses the past of the process X^ν to select a transition probability from the set $\Gamma(X^\nu_n)$. The value function of the game (if it exists) can be written as

$$v(x) = \sup_{\nu \in \mathcal{A}} \inf_{\tau \in \mathcal{T}} \mathsf{E}_x[g(X^\nu)] = \inf_{\tau \in \mathcal{T}} \sup_{\nu \in \mathcal{A}} \mathsf{E}_x[g(X^\nu)].$$

It is proved that the game admits a value using a martingale approach. It turns out that the value of the game is the largest deficient function which is dominated by g. Deficient function is a concept related to sub-harmonic functions and sub-martingale processes and for further details the reader may refer to [37]. Moreover, it is proved that for all $\eta > 0$ the controller has an η-optimal Markov family of strategies for all $x \in S$, i.e., the η-optimal strategy ν_n at each time $n \in \mathbb{N}$ depends only on the state at the previous time-step $\nu_n(x_0, \dots, x_{n-1}) = \nu_n(x_{n-1})$.

2.3 Stopper vs. Classical Controller: Continuous Time

We start considering a so-called classical controller, i.e., a type of player who is allowed to use only controls that are absolutely continuous with respect to the Lebesgue measure as a function of time. We concentrate first on contributions about one-dimensional cases and later, we present the ones in a multi-dimensional cases. From now on, let T be a finite time horizon, $(\Omega, \mathcal{F}, \mathsf{P})$ be a filtered probability space with filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ and equipped with a d-dimensional Brownian motion $(W_t)_t$. For the ease of exposition, the dimension d of the Brownian motion agrees with the dimension of the underlying process specified in each framework.

2.3.1 One-dimensional case

Karatzas and Sudderth [42] studied a game where the underlying process X is a one-dimension Itô diffusion

$$\mathrm{d}X_t = b(t)\mathrm{d}t + \sigma(t)\mathrm{d}W_t, \quad t \in [0, \infty)$$

which evolves in a bounded interval $(\alpha, \beta) \subset \mathbb{R}$ with absorption at the boundary points α, β. The coefficients b and σ are real-valued, $\mathbb{F}$-progressively measurable processes which satisfy a suitable integrability condition, which we do not specify for the ease of exposition. The controller acts as follows: a family $\{\mathcal{A}(y) : y \in [\alpha, \beta]\}$ with $\mathcal{A}(y) \subset \mathbb{R} \times (0, \infty)$ is fixed at the beginning of the game and whenever the underlying process X is in a given state $X_t = x \in [\alpha, \beta]$, the player can choose a local-drift volatility pair $\nu = (\nu_1(t), \nu_2(t))$ from $\mathcal{A}(x)$, i.e, $X = X^\nu$ is solution of

$$\mathrm{d}X_t^\nu = \nu_1(t)\mathrm{d}t + \nu_2(t)\mathrm{d}W_t, \quad t \in [0, \infty).$$

The admissible class $\mathcal{A}$ in this case is composed by all processes X^ν which can be constructed as above. In this game, (2.1) is expressed with a discount factor $r \in [0, \infty)$, i.e., the expected payoff reads

$$\mathcal{J}_x(\tau, X^\nu) := \mathsf{E}_x[e^{-r\tau} g(X_\tau^\nu)].$$

The key point is that the pair (ν_1, ν_2) can be chosen in the form of a Markov pair. That is, for the minimiser it is sufficient to find two functions $b(x)$ and $\sigma(x)$, with suitable regularity, and set $\nu_1(t) = b(X_t)$ and $\nu_2(t) = \sigma(X_t)$. Then the controlled dynamics becomes a one-dimensional stochastic differential equation and it enables the use of concepts from diffusion theory like the scale function and the speed measure. Using these tools, for any given pair of functions (b, σ) the maximiser solves an optimal stopping problem for a one dimensional linear diffusion. That determines the candidate optimal stopping time τ_*. Finally, the controller selects the pair (b^*, σ^*) which provides the best response to τ_* and the proof is concluded by showing that indeed the treble $[(b^*, \sigma^*), \tau_*]$ forms a saddle point.

This approach is based on two main ingredients: (i) the scale function for one-dimensional processes and (ii) the fact that the controls affect the dynamics in an absolutely continuous way. In our framework, both these ingredients are not present and we allow the controller to affect the process with singular controls.

A problem more closely related to ours is given by Weerasinghe in [67]. In this work, a one-dimensional process is considered on infinite time horizon and the game is connected to an ordinary differential equation. A similar connection is found in this thesis, but since we consider a multi-dimensional process on a finite horizon the link is between the game and a partial differential equation of parabolic type with d spatial dimensions.

In the framework of [67], the underlying process is one-dimensional and the controller acts only on the diffusion coefficient:

$$\mathrm{d}X_t^\nu = b(X_t^\nu)\mathrm{d}t + \nu_t \mathrm{d}W_t,$$

where b satisfies standard assumptions and ν belongs to the class $\mathcal{A}$ of $\mathbb{F}$-progressively measurable processes uniformly bounded: $0 \leq \nu(t) \leq \sigma_0$, for a fixed $\sigma_0 \in (0, \infty)$. The payoff (2.1) reads

$$\mathcal{J}_x(\tau, \nu) := \mathsf{E}_x\left[\int_0^\tau e^{-rs} h(X_s^\nu)\,\mathrm{d}s\right],$$

where h is twice continuously differentiable defined on $\mathbb{R}$. It is proved that the value function of the game exists and a saddle point of the game is provided:

Theorem 2.1 ([67, Thm. 2.2]) *There exist an interval* $(\alpha_*, \beta_*) \supseteq (\alpha, \beta)$ *and a function* v *such that* $v(\alpha_*) = v(\beta_*) = 0$, $v(x) > 0$ *for all* $x \in (\alpha_*, \beta_*)$, $v_{xx}(x) \leq 0$ *for all* $x \in [\alpha_*, \beta_*]$ *and*

$$\tfrac{\sigma_0^2}{2} v_{xx}(x) + b(x) v_x(x) - r v(x) + h(x) = 0, \quad \textit{for all } x \in (\alpha_*, \beta_*). \tag{2.4}$$

Moreover, the pair (τ_*, ν_*) *is a saddle point, where* $\tau_* := \inf\{t \geq 0 | X_t^{\nu_*} \notin (\alpha_*, \beta_*)\}$ *and* $(\nu_*)_t := \sigma_0 \mathbb{1}_{\{X_t^{\nu_*} \in (\alpha_*, \beta_*)\}}$ *with* X^{ν_*} *a weak solution of*

$$\mathrm{d}X_t^{\nu_*} = b(X_t^{\nu_*})\,\mathrm{d}t + (\nu_*)_t \mathrm{d}W_s.$$

Furthermore, the value of the game is $v_*(x) := v(x)\mathbb{1}_{\{x \in (\alpha_*, \beta_*)\}}$.

It is proved first that there exist an interval (α_*, β_*) and a function v solution of the ODE in (2.4) with such properties. Then, it is proved that the pair (τ_*, ν_*) is a saddle point of the game. The approach of the proof is similar to the heuristic argument we use in Chapter 3 to find the system of equations for our variational inequality (see Problem A). Thus, we will not give here all the details. The regularity of h implies that v composed with the controlled process X^ν is sufficiently regular to apply Itô's formula to $e^{-rs} v(X_s^\nu)$ for a stopping time $\tau \in \mathcal{T}$. Taking expectations, we get

$$\mathsf{E}_x[e^{-r\tau} v(X_\tau^\nu)] = v(x) + \mathsf{E}_x\left[\int_0^\tau e^{-rs}\big(\tfrac{\nu_s^2}{2} v_{xx}(X_s^\nu) + b(x) v_x(X_s^\nu) - r v(X_s^\nu)\big)\,\mathrm{d}s\right].$$

First we show that $v(x) \leq \underline{v}(x)$. Letting $\tau = \tau_*$ as in Theorem 2.1, we have

$$\begin{aligned} 0 &= v(x) + \mathsf{E}_x\left[\int_0^{\tau_*} e^{-rs}\big(\tfrac{\sigma_0^2}{2} v_{xx}(X_s^\nu) + b(x)v_x(X_s^\nu) - rv(X_s^\nu) + \tfrac{\nu_s^2-\sigma_0^2}{2} v_{xx}(X_s^\nu)\big)\,\mathrm{d}s\right] \\ &= v(x) + \mathsf{E}_x\left[\int_0^{\tau_*} e^{-rs}\big(-h(X_s^\nu) + \tfrac{\nu_s^2-\sigma_0^2}{2} v_{xx}(X_s^\nu)\big)\,\mathrm{d}s\right], \end{aligned}$$

where the first inequality is by $v(X_{\tau_*}^\nu) = 0$ and the second equality uses that v is solution of the ODE. Arranging terms, using that v_{xx} is non-positive and $\nu_s \leq \sigma_0$ for all s, we have

$$v(x) \leq \mathsf{E}_x\left[\int_0^{\tau_*} e^{-rs} h(X_s^\nu)\,\mathrm{d}s\right] = \mathcal{J}_x(\tau_*, \nu).$$

Similarly, we prove that $v(x) \geq \overline{v}(x)$. Using the optimal control ν_* defined in Theorem 2.1, we obtain

$$\mathsf{E}_x[e^{-r\tau} v(X_\tau^{\nu_*})] = v(x) + \mathsf{E}_x\left[\int_0^{\tau} e^{-rs}\big(-h(X_s^{\nu_*})\big)\,\mathrm{d}s\right].$$

Using that $v(y) \geq 0$ for all y, we have

$$\mathcal{J}_x(\tau, \nu_*) = \mathsf{E}_x\left[\int_0^{\tau} e^{-rs} h(X_s^{\nu_*})\,\mathrm{d}s\right] \leq v(x).$$

This is sufficient to obtain that (τ_*, ν_*) is a saddle point.

Remark 2.2: Notice that the terminal payoff of this problem is $g \equiv 0$ on $\mathbb{R}$. Theorem 2.1 implies that $v_*(x) > 0$ for $x \in (\alpha_*, \beta_*)$ and $v_*(x) = 0$ otherwise. Thus, the optimal stopping time can be written as

$$\tau_* := \inf\{t \geq 0 | v_*(X_t^{\nu_*}) = g(X_t^{\nu_*})\}.$$

The stopping time τ_* is the same optimal stopping rule found in this dissertation as optimal strategy of the stopper. ■

2.3.2 Multi-dimensional case

After those early contributions for one-dimensional problems, the focus moved to the multi-dimensional case. In 2006, Hamadène [33] used a BSDE approach to solve a class of ZSGs. Let $\mathcal{T}_T$ be the subset of $\mathcal{T}$ composed by bounded stopping times with $\tau \le T$ P-a.s. and $\mathcal{A}$ be the class of $\mathbb{F}$-adapted processes with values in a suitable space A. In this class of games, the two players can both choose stopping times and controls: $(\tau, \nu) \in \mathcal{T}_T \times \mathcal{A}$ and $(\rho, \mu) \in \mathcal{T}_T \times \mathcal{A}$ for the first and second player, respectively. The underlying process is a weak solution of the following differential path-dependent equation

$$\mathrm{d}X_t^{[\nu,\mu]} = b(t, X_\cdot^{[\nu,\mu]}, \nu_t, \mu_t)\mathrm{d}t + \sigma(t, X_\cdot^{[\nu,\mu]})\mathrm{d}W_t,$$

where b and σ satisfy standard assumptions, σ is invertible and $|\sigma^{-1}(t,x)|_{d\times d}$ has a polynomial growth. The expected payoff is

$$\begin{aligned}\mathcal{J}_x(\tau, \nu, \rho, \mu) := \mathsf{E}_x\bigg[\int_0^{\tau\wedge\rho} h(s, X_\cdot^{[\nu,\mu]}, \nu_s, \mu_s)\,\mathrm{d}s + G^1_\rho \mathbb{1}_{\{\rho<\tau<T\}} \\ + G^2_\tau \mathbb{1}_{\{\tau=\rho<T\}} + G^3_\tau \mathbb{1}_{\{\tau<\rho\}} + G\mathbb{1}_{\{\tau=\rho=T\}}\bigg],\end{aligned}$$

where G^1_s, G^2_s, G^3_s are measurable processes and h is a continuous function with linear growth. In this work, it is shown that there exists a connection between the value of the game and the solution of a particular BSDE with two reflecting boundaries defined thanks to a saddle point of the game. Thus, a saddle point $(\tau_*, \nu_*, \rho_*, \mu_*)$ is guessed and then it is shown that the initial value of the solution of the correspondent BSDE system is the value of the game. The saddle point can be described as follows: τ_* and ρ_* can be described as the first time that the value process of the game hits two level functions, G^1 and G^3; ν_* and μ_* can be describe as feedback controls, i.e., there exists a function called Hamiltonian associated to the game and the controls are defined at each time as a saddle point of this Hamiltonian. Thanks to these controls the authors are able to obtain the value of the game. The Hamiltonian associated to the game is a key tool which we will find also in the next contributions. In our case, the use of singular controls does not allow the use of an Hamiltonian. Indeed, we are able to provide a similar approach 'Hamiltonian-feedback control' when we introduce a penalised problem in Section 3.2 but this construction does not apply to the original game.

A problem in a similar framework but solved with a completely different approach can be found in Karatzas and Zamfirescu [44], where a Martingale approach is developed. The admissible classes here are $\mathcal{T}_t^T$ and $\mathcal{A}$, the subset of $\mathcal{T}$ composed by τ bounded stopping times $t \leq \tau \leq T$ P-a.s. and the class of predictable processes with values in some space A. The underlying process is a solution of a path-dependent differential equation (as in [33])

$$\mathrm{d}X_t^\nu = b(t, X^\nu, \nu_t)\,\mathrm{d}t + \sigma(t, X^\nu)\,\mathrm{d}W_t,$$

where b, σ satisfy standard assumptions and σ is such that $|\sigma^{-1}(t,x)|_{d\times d} \leq D_1$ is uniformly bounded for some real constant D_1. The expected payoff of this game is

$$\mathcal{J}_{t,x}(\tau,\nu) := \mathsf{E}_x\bigg[g(X_\tau^\nu) + \int_t^\tau h(s, X^\nu, \nu_s)\,\mathrm{d}s\bigg]. \tag{2.5}$$

The main idea is similar to the one used in [42] and it is based on results from optimal stopping and stochastic control theory when these problems are studied separately. Fixing an admissible control in the game, an optimal stopping problem can be associated to the payoff in (2.5) and it can be solved obtaining an optimal strategy for the stopper. In particular, the value function of the optimal stopping problem composed with the process X^ν turns out to be a sub-martingale. Moreover, there exists a sequence of admissible controls $(\nu_k)_{k\in\mathbb{N}}$ such that

$$\bigg(\sup_{\tau\in\mathcal{T}_t^T} \mathcal{J}_{t,x}(\tau,\nu_k)\bigg)_{k\in\mathbb{N}}$$

is a decreasing sequence. The monotonicity of this sequence is reflected to the sequence of optimal stopping times $(\tau_k)_{k\in\mathbb{N}}$ associated to the optimal stopping problem associated to ν_k. Thanks to these properties, the integrability conditions on the functions and the fact that the control does not affect the diffusion coefficient σ, they are able to prove that the game played with (2.5) admits a value. Moreover, necessary and sufficient conditions for a pair to be a saddle point are provided. The optimal stopping time is the same as the one we obtain in our analysis and it can be found in other works presented before. The optimal control instead is similar to the one found in [33] and it is based on the Hamiltonian function: it is a feedback control which can be achieved only because the controller is allowed to use controls which are absolutely continuous.

Those previous contributions ([33] and [44]) treat the multi-dimensional case, where the diffusion coefficient is not affected from the control. A contribution where the controller affects both the coefficient of the underlying process is presented by Bayraktar and Huang [5]. In this framework, the admissible class $\mathcal{A}$ is defined as the class of $\mathbb{F}$-adapted processes with values in a suitable space A. The expected reward for $(\tau, \nu) \in \mathcal{T}_t^T \times \mathcal{A}$ is

$$\mathcal{J}_{t,x}(\tau,\nu) := \mathsf{E}_x\left[e^{-\int_t^\tau r(s,X_s^\nu)\,\mathrm{d}s} g(X_\tau^\nu) + \int_t^\tau e^{-\int_t^s r(\lambda,X_\lambda^\nu)\,\mathrm{d}\lambda} h(s,X_s^\nu,\nu_s)\mathrm{d}s\right],$$

where

$$\mathrm{d}X_t^\nu = b(t,X_t^\nu,\nu_t)\,\mathrm{d}t + \sigma(t,X_t^\nu,\nu_t)\,\mathrm{d}W_t.$$

Differently from previous contributions, in here the existence of the value function is proved directly and it is not a by product of the existence of saddle points. Thanks to the Hamiltonian function H introduced also in previous frameworks, the value of the game is the unique viscosity solution v of the variational inequality:

$$\max\left\{\partial_t v + H(t,x,\nabla v, D^2 v) - rv, v - g\right\} = 0, \quad \text{on } \mathbb{R}^{d+1}_{0,T},$$

with terminal condition $v(T,x) = g(x)$ for all $x \in \mathbb{R}^d$. In our problem (see (1.3)) there are two variational inequalities with two constraints. This is the difference between the use of 'classical' controls and singular controls. In the latter, the variational inequality requires a gradient constraint on the value function of the game. Instead, 'classical' controls allow the use of the Hamiltonian function inside the variational inequality and thus it requires only an obstacle constraint inside. Since there is only one constraint inside the variational inequality, a comparison principle for sub-solutions and super-solutions is proved which leads to the existence of the value of the game. In our framework, since there are two constraints in the variational inequality then we are not able to prove a comparison principle and indeed, we just obtain a maximal relationship between the value of the game and the solution of the variation inequality (1.3).

2.4 Stopper vs. Singular Controller

The common denominator of all those previous contributions is the use of 'classical' controls. Controls which are not absolutely continuous as functions of time are singular controls (e.g. jump processes, local time). In this last section, we present the main contributions to this type of ZSGs.

2.4.1 One-dimensional case

One of the first contribution in the topic of ZSG of stopper vs. controller type with singular controls is by Hernandez-Hernandez et al. [35] (see also [36]). The class of admissible controls $\mathcal{A}$ is composed by all $\mathbb{F}$-adapted finite-variation càglàd processes ν such that $\nu_0 = 0$. They consider a one-dimensional process X^ν with $\nu \in \mathcal{A}$ whose dynamics is

$$\mathrm{d}X_t^\nu = b(X_t^\nu)\mathrm{d}t + \sigma(X_t^\nu)\mathrm{d}W_t + \mathrm{d}\nu_t, \quad X_0 = x \in \mathbb{R},$$

where b, σ are Lipschitz continuous functions and σ is uniformly elliptic. The expected payoff is the following

$$\mathcal{J}_x(\nu,\tau) := \mathsf{E}_x\left[e^{-\Lambda_\tau} g(X_{\tau+}^\nu)\mathbb{1}_{\{\tau<\infty\}} + \int_0^\tau e^{-\Lambda_s} h(X_s^\nu)\,\mathrm{d}s + \int_{[0,\tau]} e^{-\Lambda_s}\,\mathrm{d}|\nu_s|\right], \quad (2.6)$$

where $\tau \in \mathcal{T}$, the functions g, h, r are continuous with $r > r_0 \in \mathbb{R}$, $|\nu_s|$ denotes the total variation process of ν and $\Lambda_t := \int_0^t r(X_s^\nu)\,\mathrm{d}s$. In this formulation, the game does not instantly finish at τ but the controller is allowed to move the process X up to time τ included, i.e., $X_{\tau+}^\nu$ may be different from X_τ^ν. Compared with the previous contributions, the payoff herein presents an extra term, i.e., the last one in (2.6). This term is called action cost and it gives a cost each time the control is exerted, continuously or singularly. In particular, it is proportional to the measure induced by the control and an action function f which is $f \equiv 1$ in this case. In our frameworks, we have an action function dependent on time and space, see Chapter 3 and only on time, see Chapter 4.

Remark 2.3: Every process $\nu \in \mathcal{A}$ admits the decomposition $\nu = \nu^c + \nu^j$, where ν^c, ν^j are $\mathbb{F}$-adapted finite-variation càglàd processes such that ν^c has continuous sample paths

$$\nu_0^c = \nu_0^j = 0 \quad \text{and} \quad \nu_t^j = \sum_{0 \le s < t} \Delta\nu_s \quad \text{for all } t > 0,$$

where $\Delta\nu_s = \nu_{s+} - \nu_s$ for $s \ge 0$. Given such decomposition, there exists $\mathbb{F}$-adapted continuous processes $(\nu^c)^+, (\nu^c)^-$ such that

$$(\nu^c)_0^+ = (\nu^c)_0^- = 0, \quad \nu^c = (\nu^c)^+ - (\nu^c)^- \quad \text{and} \quad |\nu^c| = (\nu^c)^+ + (\nu^c)^-,$$

where $|\nu^c|$ is the total variation process of ν^c. ■

The problem is solved by a guess and verifying approach, i.e., a set of properties for the value of the game and the structure of the optimal strategies are guessed and it is verified that if a function satisfies these conditions then it is the value of the game. These properties can be connected to the conditions in our problem: variational inequalities with both an obstacle and a gradient constraint (see (3.6)). Differently from this guess and verifying approach, we take a more 'constructive' approach. We prove the existence of the value function by a penalisation method and we connect it to a solution of the set of variational inequalities. In particular the value of the game is a maximal solution of the variational inequalities. The system of properties in [35] ensures the existence of an optimal control which can be described as follows: the controlled process is kept inside a region of the space, where the gradient constraint is satisfied with strict inequality and the process is reflected along the free boundary of this region. The optimal stopping strategy for the stopper is the same found in our analysis. The pair constructed according to this recipe forms a saddle point of the game which gives directly the existence of the value of the game.

The variational inequality found in the verification theorem is similar to the one obtained by us in Chapter 3. In our case, the variational inequality is also a characterisation for the value function. Indeed, we show independently that the value function exists and then, we show that it is also the maximal solution of the variational inequality. The framework of [35], i.e., the infinite horizon of the problem and the one-dimension of the underlying process, leads to the study of ODEs. In our case, we deal with a multi-dimensional underlying process in finite time horizon and these lead us to consider PDEs of parabolic type in place of ODEs. We do not require boundedness neither of the coefficients of the SDE nor the functions in the payoff. We prove directly the existence of the value of the game, we show which type of variational inequalities it satisfies and we provide an optimal strategy for the stopper. Unfortunately, we are not able to provide an optimal strategy for the controller. In singular control it is difficult to prove the existence of an optimal strategy in multi-dimensional problems and we are not able to provide it in the generality of our framework.

Chapter 3

Zero-sum game between Controller and Stopper

In this Chapter, we study a zero-sum game between a minimiser who controls all the directions of an underlying process and pays at the end of a game a stopper who has decided the terminal time. Under the Assumptions 3.4 and 3.5, we prove our main result (Theorem 3.6) which shows that the value of the game is the maximal solution of a variational inequality and provides an optimal strategy for the stopper.

3.1 Setting and Main Results

Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a complete probability space, $\mathbb{F} = (\mathcal{F}_s)_{s\in[0,\infty)}$ be a right-continuous filtration completed by the P-null sets and $(W_s)_{s\in[0,\infty)}$ be a $\mathbb{F}$-adapted, d'-dimensional Brownian motion. Fix $T \in (0, \infty)$ and $d \leq d'$. For $t \in [0, T]$, we denote

$$\mathcal{T}_t := \{\tau \,|\, \tau \text{ is } \mathbb{F}\text{-stopping time with } \tau \in [0, T-t], \ \mathsf{P}\text{-a.s.}\}$$

and we let $\mathcal{A}_t$ be the class of processes

$$\mathcal{A}_t := \left\{ (n,\nu) \;\middle|\; \begin{array}{l} (n_s)_{s\in[0,\infty)} \text{ is progressively measurable, } \mathbb{R}^d\text{-valued,} \\ \text{with } |n_s|_d = 1, \ \mathsf{P}\text{-a.s. for all } s \in [0,\infty); \\ (\nu_s)_{s\in[0,\infty)} \text{ is } \mathbb{F}\text{-adapted, real valued, non-decreasing and} \\ \text{right-continuous with } \nu_{0-} = 0, \ \mathsf{P}\text{-a.s., and } \mathsf{E}[|\nu_{T-t}|^2] < \infty \end{array} \right\}.$$

The notation $\nu_{0-} = 0$ accounts for a possible jump of ν at time zero. For a given pair $(n, \nu) \in \mathcal{A}_t$ we consider the following (controlled) stochastic differential equation:

$$X_s^{[n,\nu]} = x + \int_0^s b(X_u^{[n,\nu]})\,\mathrm{d}u + \int_0^s \sigma(X_u^{[n,\nu]})\,\mathrm{d}W_u + \int_{[0,s]} n_u\,\mathrm{d}\nu_u, \tag{3.1}$$

for $s \in [0, T-t]$, where $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d'}$ are continuous functions. For P-a.e. ω, the map $s \mapsto n_s(\omega)$ is Borel-measurable on $[0, T]$ and $s \mapsto \nu_s(\omega)$ defines a measure on $[0, T]$; thus the Lebesgue-Stieltjes integral $\int_{[0,s]} n_u(\omega)\mathrm{d}\nu_u(\omega)$ is well-defined for P-a.e. ω. Under our Assumption 3.4 on (b, σ) there is a unique $\mathbb{F}$-adapted solution of (3.1) by, e.g., [48, Thm. 2.5.7]. We denote

$$\mathsf{P}_x(\,\cdot\,) = \mathsf{P}(\,\cdot\,|X_{0-}^{[n,\nu]} = x) \quad \text{and} \quad \mathsf{E}_x[\,\cdot\,] = \mathsf{E}[\,\cdot\,|X_{0-}^{[n,\nu]} = x].$$

We study a class of 2-player zero-sum games (ZSGs) between a (singular) controller and a stopper. The stopper picks a stopping time $\tau \in \mathcal{T}_t$ and the controller chooses a pair $(n, \nu) \in \mathcal{A}_t$. At time τ the game ends and the controller pays to the stopper a random payoff depending on τ and on the path of $X^{[n,\nu]}$ up to time τ. Given continuous functions $f, g, h : \mathbb{R}^{d+1}_{0,T} \to [0, \infty)$, a fixed discount rate $r \geq 0$ and $(t, x) \in \mathbb{R}^{d+1}_{0,T}$, the game's *expected* payoff reads

$$\begin{aligned}\mathcal{J}_{t,x}(n, \nu, \tau) = \mathsf{E}_x\Big[e^{-r\tau} g(t+\tau, X_\tau^{[n,\nu]}) + \int_0^\tau e^{-rs} h(t+s, X_s^{[n,\nu]})\,\mathrm{d}s \\ + \int_{[0,\tau]} e^{-rs} f(t+s, X_s^{[n,\nu]}) \circ \mathrm{d}\nu_s \Big],\end{aligned} \tag{3.2}$$

where

$$\begin{aligned}\int_{[0,\tau]} e^{-rs} f(t+s, X_s^{[n,\nu]}) \circ \mathrm{d}\nu_s := \int_0^\tau e^{-rs} f(t+s, X_s^{[n,\nu]})\,\mathrm{d}\nu_s^c \\ + \sum_{0\leq s\leq \tau} e^{-rs} \int_0^{\Delta\nu_s} f(t+s, X_{s-}^{[n,\nu]} + \lambda n_s)\,\mathrm{d}\lambda.\end{aligned} \tag{3.3}$$

Here ν^c is the continuous part of the process ν in the decomposition $\nu_s = \nu_s^c + \sum_{u\leq s} \Delta\nu_u$, with $\Delta\nu_u = \nu_u - \nu_{u-}$.

Remark 3.1: The sum in (3.3) is well-defined. ■

Remark 3.2: If $f(t,x) = f(t)$ the integral (3.3) reduces to the standard Lebesgue-Stieltjes integral $\int_{[0,\tau]} f(s)\mathrm{d}\nu_s$. In general, the integral in (3.3), is different from the definition of a Lebesgue-Stieltjes integral. We follow the approach proposed in Zhu [69] and adopt the definition (3.3) instead of the Lebesgue-Stieltjes integral. We take this approach because (3.3) gives a cost of exerting control that is consistent with the gradient constraint appearing in Hamilton-Jacobi-Bellman (HJB) equations for singular stochastic control (see, e.g., [50]). In problems where the control is monotone (commonly called of the monotone follower type) the Lebesgue-Stieltjes integral and the one in (3.3) can be connected by the following argument: consider the control ν^n that at a given time t makes n instantaneous jumps of size h/n for a fixed h. Taking the limit as $n \to \infty$ the classical Lebesgue-Stieltjes integral becomes the integral in (3.3). It is shown in [1, Cor. 1] that the control obtained in the limit, i.e., $\nu^\infty := \lim_{n\to\infty} \nu^n$ is an optimal strategy in the sense that it optimises the functional, but it is not an admissible strategy in the class $\mathcal{A}_t$. We report here the example from [1, Cor. 1] which could give a better understanding. Let $n_s \equiv 1$ be a scalar and let ν_s be a non-decreasing, right-continuous process. We focus on jumps at time zero of the control and, in particular, we let the control ν^n be such that at time zero it makes n instantaneous jumps of size h/n. We have that the Riemann-Stieltjes integral with $\tau = 0$ can be written as

$$\int_0^{[0]} f(X_s^{\nu^n})\,\mathrm{d}\nu_s^n = \sum_{k=1}^{n} f(X_{k-1}^{\nu^n})(\nu_k^n - \nu_{k-1}^n) = \sum_{k=1}^{n} f(x_{k-1})(x_k - x_{k-1}),$$

where $X_k^{\nu^n}$ and ν_k^n are the values of the processes X^{ν^n} and ν^n after k jumps of size h/n, respectively (notice that $X_k^{\nu^n} = x + kh/n =: x_k$). From this integral, it is clear that if we pass to the limit as $n \to \infty$, we obtain that

$$\lim_{n\to\infty} \int_0^{[0]} f(X_s^{\nu^n})\,\mathrm{d}\nu_s^n = \int_0^h f(x+\lambda)\,\mathrm{d}\lambda.$$

On the other-hand, if we choose the definition of the integral as in (3.3) and we consider the control ν^n that makes n jumps of size h/n at time zero, we get for all $n \in \mathbb{N}$ that

$$\int_0^{[0]} f(X_s^{\nu^n}) \circ \mathrm{d}\nu_s^n = \sum_{k=1}^{n} \int_{(k-1)h/n}^{kh/n} f(X_{0-}^{\nu^n} + \lambda)\,\mathrm{d}\lambda = \int_0^h f(x+\lambda)\,\mathrm{d}\lambda.$$

This shows that the two integrals, in this particular setting, give the same cost. ■

The game admits *lower* and *upper* value, defined respectively by

$$\underline{v}(t,x) \coloneqq \sup_{\tau\in\mathcal{T}_t} \inf_{(n,\nu)\in\mathcal{A}_t} \mathcal{J}_{t,x}(n,\nu,\tau) \quad \text{and} \quad \overline{v}(t,x) \coloneqq \inf_{(n,\nu)\in\mathcal{A}_t} \sup_{\tau\in\mathcal{T}_t} \mathcal{J}_{t,x}(n,\nu,\tau), \tag{3.4}$$

so that $\underline{v}(t,x) \le \overline{v}(t,x)$. If equality holds then we say that our game admits a value

$$v(t,x) \coloneqq \underline{v}(t,x) = \overline{v}(t,x). \tag{3.5}$$

For $a(x) \coloneqq (\sigma\sigma^\top)(x) \in \mathbb{R}^{d\times d}$, the infinitesimal generator of $X^{[e_1,0]}$ (where e_1 is the unit vector with 1 in the first entry) reads

$$(\mathcal{L}\varphi)(x) = \tfrac{1}{2}\mathrm{tr}\big(a(x)D^2\varphi(x)\big) + \langle b(x), \nabla\varphi(x)\rangle, \text{ for any } \varphi \in C^2(\mathbb{R}^d).$$

By density arguments the linear operator $\mathcal{L}$ admits a unique extension $\bar{\mathcal{L}}$ to $W^{2,p}_{\ell oc}(\mathbb{R}^d)$ and, with a slight abuse of notation, we set $\bar{\mathcal{L}} = \mathcal{L}$.

A heuristic use of the dynamic programming principle, suggests that the value of the game v should be solution of a free boundary problem of the following form:

Problem A. Fix $p > d+2$. Find a function $u \in W^{1,2,p}_{\ell oc}(\mathbb{R}^{d+1}_{0,T})$ such that, letting

$$\mathcal{I} \coloneqq \big\{(t,x) \in \mathbb{R}^{d+1}_{0,T} \,\big|\, |\nabla u(t,x)|_d < f(t,x)\big\} \quad \text{and}$$
$$\mathcal{C} \coloneqq \big\{(t,x) \in \mathbb{R}^{d+1}_{0,T} \,\big|\, u(t,x) > g(t,x)\big\},$$

u satisfies:

$$\begin{cases} (\partial_t u + \mathcal{L}u - ru)(t,x) = -h(t,x), & \text{for all } (t,x) \in \mathcal{C} \cap \mathcal{I}; \\ (\partial_t u + \mathcal{L}u - ru)(t,x) \geq -h(t,x), & \text{for a.e. } (t,x) \in \mathcal{C}; \\ (\partial_t u + \mathcal{L}u - ru)(t,x) \leq -h(t,x), & \text{for a.e. } (t,x) \in \mathcal{I}; \\ u(t,x) \geq g(t,x), & \text{for all } (t,x) \in \mathbb{R}^{d+1}_{0,T}; \\ |\nabla u(t,x)|_d \leq f(t,x), & \text{for all } (t,x) \in \mathbb{R}^{d+1}_{0,T}; \\ u(T,x) = g(T,x), & \text{for all } x \in \mathbb{R}^d, \end{cases} \tag{3.6}$$

with $|u(t,x)| \leq c(1+|x|_d^2)$ for all $(t,x) \in \mathbb{R}^{d+1}_{0,T}$ and a suitable $c > 0$. ■

Notice that the conditions $u \geq g$ and $|\nabla u|_d \leq f$ hold for all (t,x) because of the embedding (1.6). Thus, the two sets $\mathcal{I}$ and $\mathcal{C}$ are open in $[0,T) \times \mathbb{R}^d$.

Lemma 3.3 *A function $u \in W^{1,2,p}_{\ell oc}(\mathbb{R}^{d+1}_{0,T})$ solves Problem A if and only if u solves the variational inequalities in (1.3) a.e. on $\mathbb{R}^{d+1}_{0,T}$ with quadratic growth.*

Proof. For simplicity of exposition, we recall that a solution of (1.3) is a function u belonging to a suitable Sobolev space $W^{1,2,p}_{\ell oc}(\mathbb{R}^{d+1}_{0,T})$, it has quadratic growth and satisfies a.e. on $\mathbb{R}^{d+1}_{0,T}$

$$\begin{aligned} &\min\{\max\{\partial_t u + \mathcal{L}u - ru + h, g - u\}, f - |\nabla u|_d\} = 0, \\ &\max\{\min\{\partial_t u + \mathcal{L}u - ru + h, f - |\nabla u|_d\}, g - u\} = 0, \end{aligned} \tag{3.7}$$

with terminal condition $u(T,x) = g(T,x)$ for all $x \in \mathbb{R}^d$.

We prove now that a solution of (3.7) satisfies the six conditions in (3.6). Let u be a solution of (3.7), the sixth equation of (3.6) holds. From the first line of (3.7) we have that

$$\max\{\partial_t u + \mathcal{L}u - ru + h, g - u\} \geq 0 \quad \text{and} \quad f - |\nabla u|_d \geq 0 \tag{3.8}$$

hold almost everywhere on $\mathbb{R}^{d+1}_{0,T}$. If one of the two inequalities is strict, then the other must be an equality. The last part in (3.8) implies that the fifth equation in (3.6) holds almost everywhere and it also holds everywhere by embedding (1.6). Let $(t,x) \in \mathcal{I}$, we have that $f(t,x) - |\nabla u(t,x)|_d > 0$ is strict and we obtain from the left-hand side of (3.8) that

$$\partial_t u + \mathcal{L}u - ru + x \leq 0 \quad \text{and} \quad g - u \leq 0$$

hold a.e. on $\mathcal{I}$, i.e., we have that the third equation of (3.6) holds. Similarly, we use the second line of (3.7) in order to prove the second and the fourth equation of (3.6). It is immediate from the previous arguments that the first equation of (3.6) holds almost everywhere. To complete the proof, let $(t,x) \in \mathcal{C} \cap \mathcal{I}$ with $t \neq T$; since the intersection of open sets is open, then we can find a $\rho > 0$ such that $\mathcal{O}_\rho \coloneqq (t-\rho, t+\rho) \times B_\rho(x) \subset \mathcal{C} \cap \mathcal{I}$ where $B_\rho(x)$ is the d-dimensional ball centred in x. Thus u is a strong solution of the problem

$$\begin{cases} \partial_t v + \mathcal{L}v - rv = -h, & \text{on } \mathcal{O}_\rho, \\ v(s,y) = u(s,y), & \text{for } (s,y) \in \partial_P \mathcal{O}_\rho, \end{cases}$$

with $\partial_P \mathcal{O}_\rho \coloneqq ((t-\rho, t+\rho) \times \partial B_\rho(x)) \cup (\{t+\rho\} \times \overline{B}_\rho(x))$. By [29, Thm. 3.4.9] we have that there exists a unique classical solution v of the boundary value problem above because $h \in C^\alpha(\overline{\mathcal{O}}_\rho)$, and $\mathcal{L}$ is sufficiently regular (we will specify this regularity in Assumption 3.4). Since v is also a strong solution, then $v - u \in W^{1,2,p}(\mathcal{O}_\rho)$ is a strong solution of $\partial_t w + \mathcal{L}w - rw = 0$ in $\mathcal{O}_\rho$ with $w = 0$ at the boundary. It follows that $\|v-u\|_{W^{1,2,p}(\mathcal{O}_\rho)} = 0$ by the the theory of strong solutions for linear parabolic PDEs (see [8, 8, Thm. 2.6.5 and Rem. 2.6.4]). It follows we can choose a $C^{1,2,\alpha}_{Loc}$-representative of u which satisfies $(\partial_t u + \mathcal{L}u - ru + h)(t,x) = 0$ everywhere on $\mathcal{C} \cap \mathcal{I}$.

It is straightforward to prove that a solution of Problem A is a solution of (3.7) thanks also to the arguments above. □

In the next paragraphs, we will present the heuristic argument which lead us to the set of equations (3.6). We start from a sufficiently regular function that solves (3.6) and we prove that it is the value function of the game.

The heuristic argument

Let $u \in C^{1,2}(\mathbb{R}^{d+1}_{0,T})$ be a solution of Problem A, we want to prove that

$$u(t,x) \le \underline{v}(t,x),$$
$$u(t,x) \ge \overline{v}(t,x).$$

Let $(n,\nu) \in \mathcal{A}_t$ and consider the controlled process defined in (3.1). Apply Itô's formula to the function

$$e^{-rs}u(t+s, X_s^{[n,\nu]})$$

up to the time $\tau \in \mathcal{T}_t$. We get

$$\begin{aligned} u(t,x) =& e^{-r\tau}u(t+\tau, X_\tau^{[n,\nu]}) - \int_0^\tau e^{-rs}(\partial_t + \mathcal{L} - r)u(t+s, X_s^{[n,\nu]})\,\mathrm{d}s \\ &- \int_0^\tau e^{-rs}\sigma(X_s^{[n,\nu]})\nabla u(t+s, X_s^{[n,\nu]})\,\mathrm{d}W_s \\ &- \int_0^\tau e^{-rs}\langle \nabla u(t+s, X_s^{[n,\nu]}), n_s\rangle\,\mathrm{d}\nu_s^c \\ &- \sum_{0\le s\le\tau} e^{-rs}\int_{\nu_{s-}}^{\nu_s}\langle \nabla u(t+s, X_{s-}^{[n,\nu]} + \lambda n_s), n_s\rangle\,\mathrm{d}\lambda. \end{aligned}$$

We apply the expectation to the equation above and we obtain

$$\begin{aligned} u(t,x) = \mathsf{E}_x\Bigg[& e^{-r\tau}u(t+\tau, X_\tau^{[n,\nu]}) - \int_0^\tau e^{-rs}(\partial_t + \mathcal{L} - r)u(t+s, X_s^{[n,\nu]})\,\mathrm{d}s \\ &- \int_0^\tau e^{-rs}\langle \nabla u(t+s, X_s^{[n,\nu]}), n_s\rangle\,\mathrm{d}\nu_s^c \\ &- \sum_{0\le s\le\tau} e^{-rs}\int_{\nu_{s-}}^{\nu_s}\langle \nabla u(t+s, X_{s-}^{[n,\nu]} + \lambda n_s), n_s\rangle\,\mathrm{d}\lambda\Bigg], \end{aligned} \tag{3.9}$$

because the stochastic integral is a martingale.

We suppose that there exists a control pair $(n^*, \nu^*) \in \mathcal{A}_t$ such that: if the process starts inside the region $\{|\nabla u|_d < f\}$, the control pair acts only when the process hits the boundary of this set. The control pair at that point is an impulse in the opposite direction of the gradient of the value function with a force equal to the norm of thegradient; if it starts outside, the control makes an initial jump from outside to inside the region $\{|\nabla u|_d < f\}$ and it leaves the process evolving as above. Notice that the control acts whenever $|\nabla u|_d = f$ and the process stays inside $\overline{\{|\nabla u|_d < f\}}$ for all the interval time $(0, T-t]$. The existence of this control is related to the Skorohod reflection problem which requires regularity of the value function in order to be solved.

Using this control in (3.9) we get:

$$\begin{aligned} u(t,x) = \mathsf{E}_x\Bigg[& e^{-r\tau} u(t+\tau, X_\tau^{[n^*,\nu^*]}) - \int_0^\tau e^{-rs}(\partial_t + \mathcal{L} - r) u(t+s, X_s^{[n^*,\nu^*]})\, \mathrm{d}s \\ & - \int_0^\tau e^{-rs} \langle \nabla u(t+s, X_s^{[n^*,\nu^*]}), n_s^* \rangle \, \mathrm{d}\nu_s^{c,*} \\ & - \sum_{0 \le s \le \tau} e^{-rs} \int_{\nu_{s-}}^{\nu_s} \langle \nabla u(t+s, X_{s-}^{[n,\nu]} + \lambda n_s), n_s \rangle \, \mathrm{d}\lambda \Bigg] \\ = \mathsf{E}_x\Bigg[& e^{-r\tau} u(t+\tau, X_\tau^{[n^*,\nu^*]}) - \int_0^\tau e^{-rs}(\partial_t + \mathcal{L} - r) u(t+s, X_s^{[n^*,\nu^*]})\, \mathrm{d}s \\ & + \int_0^\tau e^{-rs} f(t+s, X_s^{[n^*,\nu^*]}) \, \mathrm{d}\nu_s^{c,*} \\ & + \sum_{0 \le s \le \tau} e^{-rs} \int_{\nu_{s-}}^{\nu_s} f(t+s, X_{s-}^{[n,\nu]} + \lambda n_s) \, \mathrm{d}\lambda \Bigg]. \end{aligned}$$

Using the third and fourth line of (3.6), we have from the above equation

$$\begin{aligned} u(t,x) \ge \mathsf{E}_x\Bigg[& e^{-r\tau} g(t+\tau, X_\tau^{[n^*,\nu^*]}) + \int_0^\tau e^{-rs} h(t+s, X_s^{[n^*,\nu^*]})\, \mathrm{d}s \\ & + \int_0^\tau e^{-rs} f(t+s, X_s^{[n^*,\nu^*]}) \circ \mathrm{d}\nu_s^* \Bigg] \\ = \mathcal{J}_{t,x} & (n^*, \nu^*, \tau) \end{aligned}$$

The inequality holds for any choice of $\tau \in \mathcal{T}_t$, thus

$$u(t,x) \ge \sup_{\tau \in \mathcal{T}_t} \mathcal{J}_{t,x}(n^*, \xi^*, \tau) \ge \overline{v}(t,x).$$

We prove now that $u(t,x) \le \underline{v}(t,x)$. Recalling the last line of (3.6), we consider the stopping time $\tau_* = \tau_*(t,x,n,\nu)$:

$$\tau_* := \inf\big\{ s \in [0, T-t] : u(t+s, X_s^{[n,\nu]}) = g(t+s, X_s^{[n,\nu]}) \big\}.$$

From (3.9), using the second line of (3.6), we get

$$
\begin{aligned}
u(t,x) \le \mathsf{E}_x \Bigg[& e^{-r\tau_*} g(t+\tau_*, X^{[n,\nu]}_{\tau_*}) + \int_0^\tau e^{-rs} h(t+s, X^{[n,\nu]}_s)\,\mathrm{d}s \\
& - \int_0^{\tau_*} e^{-rs} \langle \nabla u(t+s, X^{[n,\nu]}_s), n_s\rangle \,\mathrm{d}\nu^c_s \\
& - \sum_{0\le s\le \tau_*} e^{-rs} \int_{\nu_{s-}}^{\nu_s} \langle \nabla u(t+s, X^{[n,\nu]}_{s-} + \lambda n_s), n_s\rangle \,\mathrm{d}\lambda \Bigg].
\end{aligned}
$$

We use now the fifth line of (3.6) and

$$
\begin{aligned}
u(t,x) \le \mathsf{E}_x \Bigg[& e^{-r\tau_*} g(t+\tau_*, X^{[n,\nu]}_{\tau_*}) + \int_0^\tau e^{-rs} h(t+s, X^{[n,\nu]}_s)\,\mathrm{d}s \\
& + \int_0^{\tau_*} e^{-rs} f(t+s, X^{[n,\nu]}_s) \,\mathrm{d}\nu^c_s \\
& + \sum_{0\le s\le \tau_*} e^{-rs} \int_{\nu_{s-}}^{\nu_s} f(t+s, X^{[n,\nu]}_{s-} + \lambda n_s) \,\mathrm{d}\lambda \Bigg] \\
= \mathcal{J}_{t,x}(n,\nu,\tau_*) &
\end{aligned}
$$

The inequality holds for every choice of $(n,\nu) \in \mathcal{A}_t$ and we have

$$
\begin{aligned}
u(t,x) &\le \inf_{(n,\nu)\in\mathcal{A}} \mathcal{J}_{t,x}(n,\nu,\tau_*) \\
&\le \underline{v}(t,x).
\end{aligned}
$$

Thus we have that $u(t,x) = \underline{v}(t,x) = \overline{v}(t,x)$. ■

Next we give assumptions under which we obtain our main result (Theorem 3.6).

Assumption 3.4 (Controlled SDE) *The functions b and σ are continuously differentiable and locally Lipschitz on $\mathbb{R}^d$. Moreover, there is $D_1 > 0$ such that*

$$
|b(x)|_d + |\sigma(x)|_{d\times d'} \le D_1(1+|x|_d), \ \textit{for all } x \in \mathbb{R}^d.
$$

Recalling $a = \sigma\sigma^\top$, for any bounded set $B \subset \mathbb{R}^d$ there is $\theta_B > 0$ such that $a(\cdot)$ is *locally elliptic:*

$$\langle \zeta, a(x)\zeta\rangle \geq \theta_B |\zeta|_d^2 \qquad \text{for any } \zeta \in \mathbb{R}^d \text{ and all } x \in \overline{B}. \tag{3.10}$$

Assumption 3.5 (Functions f, g, h) *For the functions $f, g, h : \mathbb{R}^{d+1}_{0,T} \to [0,\infty)$ the following hold:*

(i) $f^2, g \in C^{1,2,\alpha}_{\ell oc}(\mathbb{R}^{d+1}_{0,T})$ *and* $h \in C^{0,1,\alpha}_{\ell oc}(\mathbb{R}^{d+1}_{0,T})$ *for some* $\alpha \in (0,1)$;

(ii) $t \mapsto f(t,x)$ *is non-increasing for each* $x \in \mathbb{R}^d$ *and* $0 \leq f(t,x) \leq c(1+|x|_d^p)$ *for some* $c, p > 0$;

(iii) there is $K_0 \in (0,\infty)$ *such that for all* $0 \leq s < t \leq T$ *and all* $x \in \mathbb{R}^{d+1}_{0,T}$

$$h(t,x) - h(s,x) \leq K_0(t-s) \quad \text{and} \quad g(t,x) - g(s,x) \leq K_0(t-s); \tag{3.11}$$

(iv) there is $K_1 \in (0,\infty)$ *such that*

$$0 \leq g(t,x) + h(t,x) \leq K_1(1+|x|_d^2), \ \text{for } (t,x) \in \mathbb{R}^{d+1}_{0,T}; \tag{3.12}$$

(v) f *and* g *are such that*

$$|\nabla g(t,x)|_d \leq f(t,x), \ \text{for all } (t,x) \in \mathbb{R}^{d+1}_{0,T}; \tag{3.13}$$

(vi) there is $K_2 \in (0,\infty)$ *such that*

$$\Theta(t,x) := \big(h + \partial_t g + \mathcal{L}g - rg\big)(t,x) \geq -K_2, \ \text{for all } (t,x) \in \mathbb{R}^{d+1}_{0,T}. \tag{3.14}$$

Condition (3.11) is immediately satisfied if h and g are time-homogeneous (as it is often the case in investment problems, see, e.g., [14], [15], [19]) or if the maps $t \mapsto \big(h(t,x), g(t,x)\big)$ are non-increasing for all $x \in \mathbb{R}^d$. Otherwise, that condition amounts to setting a maximum growth rate on $t \mapsto \big(h(t,x), g(t,x)\big)$ as time increases. Condition (3.12) is sufficient to guarantee that the value function of the game has at most a quadratic growth. Intuitively, if the controller decides to not use any control, the game evolves according the uncontrolled SDE and the expected payoff has quadratic growth since g and h have. Condition (3.13) is sufficient to guarantee that the stopping payoff satisfies the gradient constraint in (3.6) and therefore, from a probabilistic point of view, the stopper can stop at any point in the state-space: strictly speaking, this condition is only needed in the contact set $\{u = g\}$. However, the contact set is unknown *a priori* so it is difficult in our generality to formulate an assumption involving that set's properties; on a more technical level, it will be shown in Lemma 3.32 that (3.13) implies that the controller should never exert a jump at the time the stopper ends the game. Condition (3.14) guarantees that there is no region in the state space where the controller (minimiser) can push the process and obtain arbitrarily large (negative) running gains.

Example: Assumption 3.5 describes conditions on the functions f, g, h. We present now an example where these conditions are satisfied. Consider $d = 1$ and let $(0, \infty)$ be the state space (see Remark 3.8). Let the underlying process be a geometric Brownian motion with $b(x) = c_1 x$ and $\sigma(x) = c_2 x$ for some c_1, c_2 positive constants. In singular control problems, a common function cost is a constant function: $f \equiv 1$. For the terminal payoff we use $g(t, x) = x$, this is a function independent of time and its derivatives are

$$g_x(t, x) = 1 \quad \text{and} \quad g_{xx}(t, x) = 0.$$

We obtain that (3.13) is satisfied and

$$(\partial_t + \mathcal{L} - r)g = c_1 x - rx.$$

Thus, condition (3.14) is satisfied if $r \leq c_1$ or, for example, if we take a function h with quadratic growth in x such as $h(t, x) = t + c_3 x^2$. This function depends on time and $h(t, x) - h(s, x) \leq (t - s)$.

The next theorem is the main result of the paper and its proof is distilled in the following sections through a number of technical results and estimates.

Theorem 3.6 *The game described above admits a value (i.e., (3.5) holds) and the value function v of the game is the maximal solution to Problem A. Moreover, for any given $(t, x) \in \mathbb{R}^{d+1}_{0,T}$ and any admissible control $(n, \nu) \in \mathcal{A}_t$, the stopping* *time* $\tau_* = \tau_*(t,x;n,\nu) \in \mathcal{T}_t$ *defined under* P_x *as*

$$\tau_* := \inf\big\{s \geq 0 \,\big|\, v(t+s, X_s^{[n,\nu]}) = g(t+s, X_s^{[n,\nu]})\big\} \wedge (T-t), \tag{3.15}$$

is optimal for the stopper.

Remark 3.7: Uniqueness of the solution to Problem A remains an open question. Methods used in, e.g., [62], do not apply due to the presence of obstacle and gradient constraints. Existence of an optimal control pair $(n^*, \nu^*) \in \mathcal{A}_t$ is also subtle and cannot be addressed in the generality of our setting. Even in standard singular control problems (not games) abstract existence results rely on compactness arguments in the space of increasing processes under more stringent assumptions (e.g., convexity or concavity) on the functions f, g, h, b and σ (see, e.g., [12], [34], [49], [64]). ■

Remark 3.8: Our choice to work with $X^{[n,\nu]} \in \mathbb{R}^d$ is purely for simplicity of exposition. It will be clear from our arguments of proof that other types of unbounded domains as, e.g., orthants of $\mathbb{R}^d$, are equally covered by our analysis, provided that the controlled process $X^{[n,\nu]}$ cannot leave the domain in finite time. In particular, for $d = 1$ our results apply to $X^{[n,\nu]} \in (0, \infty)$, which is of specific interest for economic applications with geometric Brownian motion or certain CIR dynamics. ■

3.2 Penalised Problem and A Priori Estimates

In this section we first introduce a class of penalised problems and illustrate their connection with a class of ZSGs of control (Section 3.2.1). Then we provide important *a priori* estimates on the growth and gradient of the solution of such penalised problems (Sections 3.2.2 and 3.2.3) and, finally, we prove existence and uniqueness of the solution (Section 3.2.4).

3.2.1 A penalised problem

For technical reasons related to solvability of the penalised problem and the probabilistic representation of its solution we choose to work on a sequence of bounded domains $(\mathcal{O}_m)_{m\in\mathbb{N}} \subset \mathbb{R}^{d+1}_{0,T}$. Recall that $B_m \subset \mathbb{R}^d$ is the open ball of radius m centred in the origin and set $\mathcal{O}_m := [0,T) \times B_m$ with parabolic boundary $\partial_P\mathcal{O}_m = ([0,T) \times \partial B_m) \cup (\{T\} \times \overline{B}_m)$. To simplify notation, we use $\|\cdot\|_m = \|\cdot\|_{C^0(\overline{\mathcal{O}}_m)}$.

Let $(\xi_m)_{m\in\mathbb{N}} \subset C^\infty_c(\mathbb{R}^d)$ be such that for each $m \in \mathbb{N}$ we have:

(i) $0 \le \xi_m \le 1$ on $\mathbb{R}^d$, with $\xi_m = 1$ on B_m and $\xi_m = 0$ on $\mathbb{R}^d \setminus B_{m+1}$;

(ii) there is $C_0 > 0$ independent of $m \in \mathbb{N}$ such that

$$|\nabla \xi_m|_d^2 \le C_0 \xi_m \text{ on } \mathbb{R}^d. \tag{3.16}$$

An example of such functions is provided in Appendix B.1 for completeness. We define

$$g_m(t,x) := \xi_{m-1}(x) g(t,x) \quad \text{and} \quad h_m(t,x) := \xi_{m-1}(x) h(t,x), \quad \text{for } (t,x) \in \mathbb{R}^{d+1}_{0,T}.$$

Clearly $g_m = h_m = 0$ on $\mathbb{R}^{d+1}_{0,T} \setminus \mathcal{O}_m$ while $g_m = g$ and $h_m = h$ on $\mathcal{O}_{m-1}$. We also define a version f_m of the function f so that $f_m = f$ on $\mathcal{O}_{m-1}$ and the condition

$$|\nabla g_m(t,x)|_d \le f_m(t,x), \quad \text{for } (t,x) \in \overline{\mathcal{O}}_m, \tag{3.17}$$

is preserved. For $(t,x) \in \mathbb{R}^{d+1}_{0,T}$ we let

$$f_m(t,x) := \Big(f^2(t,x) + \|g\|_m^2 |\nabla \xi_{m-1}(x)|_d^2 + 2\big(g\xi_{m-1}\langle \nabla \xi_{m-1}, \nabla g\rangle\big)(t,x)\Big)^{\frac{1}{2}}$$

and notice that on $\overline{\mathcal{O}}_m$

$$\begin{aligned}|\nabla g_m|_d^2 =& \sum_{i=1}^d \big(\xi_{m-1}\partial_{x_i} g + g\partial_{x_i}\xi_{m-1}\big)^2 \\ =& \xi_{m-1}^2|\nabla g|_d^2 + g^2|\nabla \xi_{m-1}|_d^2 + 2\xi_{m-1} g\langle \nabla\xi_{m-1}, \nabla g\rangle \\ \le& f_m^2,\end{aligned}$$

where the inequality follows by the assumption $|\nabla g|_d \le f$ and $|\xi_m| \le 1$. Since $\nabla \xi_{m-1} = \mathbf{0}$ on B_{m-1}, we have $f_m = f$ on $\mathcal{O}_{m-1}$. Notice that $f_m^2 \in C^{0,1,\alpha}(\overline{\mathcal{O}}_m)$ by Assumption 3.5 but it does not vanish on the boundary of $\mathcal{O}_m$. By construction, it is clear that

$$g_m \to g, \quad h_m \to h \quad \text{and} \quad f_m \to f, \quad \text{as } m \to \infty,$$

uniformly on any compact $\mathcal{K} \subset \mathbb{R}^{d+1}_{0,T}$.

Let us now state the penalised problem. Fix $(\varepsilon, \delta) \in (0,1)^2$ and $m \in \mathbb{N}$. Let $\psi_\varepsilon \in C^2(\mathbb{R})$ be a non-negative, convex function such that $\psi_\varepsilon(y) = 0$ for $y \le 0$, $\psi_\varepsilon(y) > 0$ for $y > 0$, $\psi_\varepsilon' \ge 0$ and $\psi_\varepsilon(y) = \frac{y-\varepsilon}{\varepsilon}$ for $y \ge 2\varepsilon$. An example of such functions is provided in Appendix B.2 for completeness. We also denote $(y)^+ := \max\{0, y\}$ for $y \in \mathbb{R}$. In order to use the results within this chapter also in Chapter 4 we introduce $\mathcal{H} : \mathbb{R}^d \to \mathbb{R}$ as $\mathcal{H}(p) := |p|_d^2$. Notice that $\nabla\mathcal{H}(p) = 2p$ and $D^2\mathcal{H}(p) = 2I_{d\times d}$. Indeed, $\mathcal{H}$ is a uniform convex function and this will be important in Chapter 4.

Problem B. Find $u = u_m^{\varepsilon,\delta}$ with $u \in C^{1,2,\alpha}(\overline{\mathcal{O}}_m)$, for $\alpha \in (0,1)$ as in Assumption 3.5, that solves:

$$\begin{cases} (\partial_t + \mathcal{L} - r)u = -h_m - \frac{1}{\delta}(g_m - u)^+ + \psi_\varepsilon\big(\mathcal{H}(\nabla u) - f_m^2\big), & \text{on } \mathcal{O}_m, \\ u(t,x) = g_m(t,x), & (t,x) \in \partial_P \mathcal{O}_m. \end{cases} \tag{3.18}$$

■

There are two useful probabilistic interpretations of a solution to Problem B, which we are going to illustrate next. Given $t \in [0,T]$, define the control classes

$$\mathcal{A}_t^\circ := \left\{ (n,\nu) \,\middle|\, \begin{array}{l} (n_s)_{s\in[0,\infty)} \text{ is progressively measurable, } \mathbb{R}^d\text{-valued,} \\ \text{with } |n_s|_d = 1, \text{ P-a.s. for all } s \in [0,\infty);\ (\nu_s)_{s\in[0,\infty)} \text{ is} \\ \mathbb{F}\text{-adapted, real valued, non-decreasing and absolutely} \\ \text{continuous in time, P-a.s., with } \mathsf{E}[|\nu_{T-t}|^2] < \infty \end{array} \right\}$$

and

$$\mathcal{T}_t^\delta := \left\{ w \,\middle|\, \begin{array}{l} (w_s)_{s\in[0,\infty)} \text{ is progressively measurable,} \\ \text{with } 0 \le w_s \le \frac{1}{\delta}, \text{ P-a.s. for all } s \in [0, T-t] \end{array} \right\}.$$

It is obvious that $\mathcal{A}_t^\circ \subset \mathcal{A}_t$. For $(t,x) \in \mathbb{R}^{d+1}_{0,T}$ and $y \in \mathbb{R}^d$ we define the Hamiltonian

$$H_m^\varepsilon(t,x,y) := \sup_{p\in\mathbb{R}^d} \big\{ \langle y, p\rangle - \psi_\varepsilon\big(\mathcal{H}(p) - f_m^2(t,x)\big)\big\}. \tag{3.19}$$

The function H_m^ε is non-negative (pick $p = \mathbf{0}$). Thanks to (3.17), choosing $p = -\nabla g_m(t,x)$ we have

$$H_m^\varepsilon(t,x,y) \ge -\langle y, \nabla g_m(t,x)\rangle, \quad \text{for all } (t,x,y) \in [0,T]\times\mathbb{R}^d\times\mathbb{R}^d. \tag{3.20}$$

For any admissible pair $(n,\nu) \in \mathcal{A}_t^\circ$, we consider the controlled dynamics

$$X_s^{[n,\nu]} = x + \int_0^s \big[b(X_u^{[n,\nu]}) + n_u\dot{\nu}_u\big]\mathrm{d}u + \int_0^s \sigma(X_u^{[n,\nu]})\mathrm{d}W_u, \qquad \text{for } 0 \le s \le T-t,$$

where now the process ν is absolutely continuous with respect the Lebesgue measure of time, P-a.s. Thus for a fixed $T > 0$ and for almost every $\omega \in \Omega$ there exists a measurable function $(\dot{\nu}_s)_{s\in[0,\infty)}$ (which depends on ω) such that

$$\int_0^t \mathbb{1}_A \,\mathrm{d}\nu_s = \int_0^t \mathbb{1}_A \dot{\nu}_s \,\mathrm{d}s$$

for all measurable sets $A \subseteq [0,T]$. Sometimes, $\dot{\nu}$ is called the Radon-Nikodym derivative of $\mathrm{d}\nu_s$ with respect to $\mathrm{d}s$. We define $\rho_m = \rho_{\mathcal{O}_m}(t,x;n,\nu)$ as

$$\rho_m = \inf\{s \ge 0 \,|\, X_s^{[n,\nu]} \notin B_m\} \wedge (T-t). \tag{3.21}$$

Instead, for $w \in \mathcal{T}_t^\delta$, we introduce a controlled discount factor

$$R_s^w := \exp\Big(- \int_0^s \big[r + w_\lambda\big]\,\mathrm{d}\lambda\Big).$$

For $(t,x) \in \overline{\mathcal{O}}_m$ and a treble $[(n,\nu),w] \in \mathcal{A}_t^\circ \times \mathcal{T}_t^\delta$ let us consider an expected payoff:

$$\mathcal{J}_{t,x}^{\varepsilon,\delta,m}(n,\nu,w) := \mathsf{E}_x\Big[R_{\rho_m}^w g_m(t+\rho_m, X_{\rho_m}^{[n,\nu]}) + \int_0^{\rho_m} R_s^w h_m(t+s, X_s^{[n,\nu]})\,\mathrm{d}s \\ + \int_0^{\rho_m} R_s^w \big[w_s g_m + H_m^\varepsilon(\cdot, n_s\dot{\nu}_s)\big](t+s, X_s^{[n,\nu]})\,\mathrm{d}s\Big]. \tag{3.22}$$

The associated upper and lower value read, respectively,

$$\overline{v}_m^{\varepsilon,\delta}(t,x) = \inf_{(n,\nu)\in\mathcal{A}_t^\circ} \sup_{w\in\mathcal{T}_t^\delta} \mathcal{J}_{t,x}^{\varepsilon,\delta,m}(n,\nu,w) \quad \text{and}$$
$$\underline{v}_m^{\varepsilon,\delta}(t,x) = \sup_{w\in\mathcal{T}_t^\delta} \inf_{(n,\nu)\in\mathcal{A}_t^\circ} \mathcal{J}_{t,x}^{\varepsilon,\delta,m}(n,\nu,w),$$

so that $\underline{v}_m^{\varepsilon,\delta} \le \overline{v}_m^{\varepsilon,\delta}$. A solution of Problem B coincides with the value function of this ZSG.

Proposition 3.9 *Let $u_m^{\varepsilon,\delta}$ be a solution of Problem B. Then*

$$u_m^{\varepsilon,\delta}(t,x) = \overline{v}_m^{\varepsilon,\delta}(t,x) = \underline{v}_m^{\varepsilon,\delta}(t,x), \quad \textit{for all } (t,x) \in \overline{\mathcal{O}}_m. \tag{3.23}$$

Proof. For simplicity denote $u = u_m^{\varepsilon,\delta}$. By definition $\overline{v}_m^{\varepsilon,\delta} = \underline{v}_m^{\varepsilon,\delta} = u = g_m$ on $\partial_P\mathcal{O}_m$. Fix $(t,x) \in \mathcal{O}_m$ and an arbitrary treble $[(n,\nu),w] \in \mathcal{A}_t^\circ \times \mathcal{T}_t^\delta$. Applying Itô's formula to $R_{\rho_m}^w u(t+\rho_m, X_{\rho_m}^{[n,\nu]})$ we have

$$\begin{aligned} R_{\rho_m}^w u(t+\rho_m, X_{\rho_m}^{[n,\nu]}) = u(t,x) &+ \int_0^{\rho_m} R_s^w(\partial_t + \mathcal{L} - r)u(t+s, X_s^{[n,\nu]})\,\mathrm{d}s \\ &- \int_0^{\rho_m} R_s^w w_s u(t+s, X_s^{[n,\nu]})\,\mathrm{d}s \\ &+ \int_0^{\rho_m} R_s^w \big\langle n_s\dot{\nu}_s, \nabla u(t+s, X_s^{[n,\nu]})\big\rangle\,\mathrm{d}s \\ &- \int_0^{\rho_m} R_s^w \nabla u(t+s, X_s^{[n,\nu]})\sigma(X_s^{[n,\nu]})\,\mathrm{d}W_s. \end{aligned}$$

Notice that the function u belongs to $C^{1,2,\alpha}(\overline{\mathcal{O}}_m)$ and thus Itô's formula applies up to time ρ_m. Applying expectation to both sides on the equation above and arranging terms we get

$$\begin{aligned} u(t,x) = \mathsf{E}_x\bigg[& R^w_{\rho_m} u(t+\rho_m, X^{[n,\nu]}_{\rho_m}) - \int_0^{\rho_m} R^w_s(\partial_t + \mathcal{L} - r)u(t+s, X^{[n,\nu]}_s)\,\mathrm{d}s \\ & - \int_0^{\rho_m} R^w_s w_s u(t+s, X^{[n,\nu]}_s)\,\mathrm{d}s + \int_0^{\rho_m} R^w_s \big\langle n_s\dot{\nu}_s, \nabla u(t+s, X^{[n,\nu]}_s)\big\rangle\,\mathrm{d}s \\ & - \int_0^{\rho_m} R^w_s \nabla u(t+s, X^{[n,\nu]}_s)\sigma(X^{[n,\nu]}_s)\,\mathrm{d}W_s\bigg]. \end{aligned}$$

The last term on the right-hand side above is a martingale because the functions inside the integral are bounded and therefore the expectation is zero. Using that u solves (3.18) we have

$$\begin{aligned} u(t,x) = \mathsf{E}_x\bigg[& R^w_{\rho_m} g_m(t+\rho_m, X^{[n,\nu]}_{\rho_m}) + \int_0^{\rho_m} R^w_s h_m(t+s, X^{[n,\nu]}_s)\,\mathrm{d}s \\ & + \int_0^{\rho_m} R^w_s \big[\tfrac{1}{\delta}\big(g_m - u\big)^+ - \psi_\varepsilon\big(\mathcal{H}(\nabla u) - f_m^2\big)\big](t+s, X^{[n,\nu]}_s)\,\mathrm{d}s \\ & + \int_0^{\rho_m} R^w_s \big[w_s u(t+s, X^{[n,\nu]}_s) - \big\langle n_s\dot{\nu}_s, \nabla u(t+s, X^{[n,\nu]}_s)\big\rangle\big]\,\mathrm{d}s\bigg]. \end{aligned} \tag{3.24}$$

We prove first that $u(t,x) \le \underline{v}^{\varepsilon,\delta}_m(t,x)$. By definition of the Hamiltonian we have (choosing $p = -\nabla u(t+s, X^{[n,\nu]}_s)$ in (3.19))

$$\begin{aligned} -\big\langle n_s\dot{\nu}_s, \nabla u(t+s, X^{[n,\nu]}_s)\big\rangle & - \psi_\varepsilon\big(\mathcal{H}(\nabla u) - f_m^2\big)(t+s, X^{[n,\nu]}_s) \\ & \le H^\varepsilon_m(t+s, X^{[n,\nu]}_s, n_s\dot{\nu}_s). \end{aligned} \tag{3.25}$$

Moreover, choosing $w = w^* \in \mathcal{T}^\delta_t$ defined as

$$w^*_s := \begin{cases} 0 & \text{if } u(t+s, X^{[n,\nu]}_s) > g_m(t+s, X^{[n,\nu]}_s), \\ \frac{1}{\delta} & \text{if } u(t+s, X^{[n,\nu]}_s) \le g_m(t+s, X^{[n,\nu]}_s), \end{cases} \tag{3.26}$$

we also have

$$\tfrac{1}{\delta}\big(g_m - u\big)^+(t+s, X^{[n,\nu]}_s) + w^*_s u(t+s, X^{[n,\nu]}_s) = w^*_s g_m(t+s, X^{[n,\nu]}_s). \tag{3.27}$$

The process $w^* \in \mathcal{T}^\delta_t$ if it is progressively measurable. For all $\lambda \in [0, \rho_m]$, we just need to prove that the pre-image of

$$\begin{aligned} D_\lambda &:= \{(s,\omega) \in [0,\lambda] \times \Omega \,|\, w^*_s(\omega) = 0\} \\ &= \{(s,\omega) \in [0,\lambda] \times \Omega \,|\, (u - g_m)(t+s, X^{[n,\nu]}_s(\omega)) > 0\} \end{aligned}$$

is a measurable set. Defining $Y_\lambda(s,\omega; n,\nu) = (u - g_m)(t+s, X^{[n,\nu]}_s(\omega))$ for $(s,\omega) \in [0,\lambda] \times \Omega$, we have that Y_λ is continuous because it is composition of a continuous function $u - g_m$ with a progressively measurable process $X^{[n,\nu]}_\lambda$, thus the pre-image of an open set is a measurable set, i.e., $Y^{-1}_\lambda((0,\infty)) = D_\lambda$ is measurable.

Then, plugging (3.25) and (3.27) into (3.24) we arrive at

$$\begin{aligned} u(t,x) &= \mathsf{E}_x \bigg[R^{w^*}_{\rho_m} g_m(t+\rho_m, X^{[n,\nu]}_{\rho_m}) + \int_0^{\rho_m} R^{w^*}_s h_m(t+s, X^{[n,\nu]}_s)\,\mathrm{d}s \\ &\qquad + \int_0^{\rho_m} R^{w^*}_s \big[w^*_s g(t+s, X^{[n,\nu]}_s) + H^\varepsilon_m(t+s, X^{[n,\nu]}_s, n_s\dot{\nu}_s)\rangle \big]\,\mathrm{d}s \bigg] \\ &= \mathcal{J}^{\varepsilon,\delta,m}_{t,x}(n,\nu,w^*). \end{aligned}$$

Since the pair (n,ν) was arbitrary, then we have

$$u(t,x) \le \inf_{(n,\nu)\in\mathcal{A}^\circ_t} \mathcal{J}^{\varepsilon,\delta,m}_{t,x}(n,\nu,w^*),$$

and therefore $u(t,x) \le \underline{v}^{\varepsilon,\delta}_m(t,x)$. Next we are going to prove that $u \ge \overline{v}^{\varepsilon,\delta}_m$.

For any $w \in \mathcal{T}^\delta_t$ it is not hard to see that

$$\tfrac{1}{\delta}\big(g_m - u\big)^+(t+s, X^{[n,\nu]}_s) + w_s u(t+s, X^{[n,\nu]}_s) \ge w_s g_m(t+s, X^{[n,\nu]}_s), \tag{3.28}$$

since $0 \le w_s \le \frac{1}{\delta}$ for all $s \in [0, T-t]$. Assume, it is possible to find a pair $(n^*,\nu^*) \in \mathcal{A}^\circ_t$ such that, putting $X^* = X^{[n^*,\nu^*]}$ we obtain

$$\begin{aligned} -\big\langle n^*_s\dot{\nu}^*_s, \nabla u(t+s, X^*_s)\big\rangle - \psi_\varepsilon\big(\mathcal{H}(\nabla u) - f^2_m\big)(t+s, X^*_s) \\ = H^\varepsilon_m(t+s, X^*_s, n^*_s\dot{\nu}^*_s). \end{aligned} \tag{3.29}$$

Then, plugging (3.28) and (3.29) into (3.24), we have $u(t,x) \ge \mathcal{J}^{\varepsilon,\delta,m}_{t,x}(n^*,\nu^*,w)$.

Since $w \in \mathcal{T}_t^\delta$ is arbitrary we have

$$u(t,x) \geq \sup_{w \in \mathcal{T}_t^\delta} \mathcal{J}_{t,x}^{\varepsilon,\delta,m}(n^*,\nu^*,w) \geq \overline{v}_m^{\varepsilon,\delta}(t,x).$$

It remains to find the pair (n^*,ν^*). By concavity, the supremum in $H_m^\varepsilon(t,x,y)$ is uniquely attained at a point $p = p(t,x,y) \in \mathbb{R}^d$ identified by the first-order condition $y = \psi_\varepsilon'(\mathcal{H}(p) - f_m^2(t,x))\nabla\mathcal{H}(p) = \psi_\varepsilon'(\mathcal{H}(p) - f_m^2(t,x))2p$. Taking

$$\begin{aligned} n_s^* &:= \begin{cases} -\frac{\nabla u_m^{\varepsilon,\delta}(t+s,X_s^*)}{|\nabla u_m^{\varepsilon,\delta}(t+s,X_s^*)|_d}, & \text{if } \nabla u_m^{\varepsilon,\delta}(t+s,X_s^*) \neq \mathbf{0}, \\ \text{any unit vector}, & \text{if } \nabla u_m^{\varepsilon,\delta}(t+s,X_s^*) = \mathbf{0}, \end{cases} \\ \dot{\nu}_s^* &:= 2\psi_\varepsilon'\big(\mathcal{H}(\nabla u_m^{\varepsilon,\delta}(t+s,X_s^*)) - f_m^2(t+s,X_s^*)\big)|\nabla u_m^{\varepsilon,\delta}(t+s,X_s^*)|_d, \end{aligned} \tag{3.30}$$

with

$$\begin{aligned} X_{s\wedge\rho_m}^* = x &+ \int_0^{s\wedge\rho_m} b(X_\lambda^*) - \big(2\psi_\varepsilon'(\mathcal{H}(\nabla u_m^{\varepsilon,\delta}) - f_m^2)\nabla u_m^{\varepsilon,\delta}\big)(t+\lambda,X_\lambda^*)\,\mathrm{d}\lambda \\ &+ \int_0^{s\wedge\rho_m} \sigma(X_\lambda^*)\,\mathrm{d}W_\lambda, \end{aligned} \tag{3.31}$$

for $s \in [0,T-t]$, we have that (3.29) holds (we restored the notation $u_m^{\varepsilon,\delta}$ for future reference). It remains to check that $(X_{s\wedge\rho_m}^*)_{s\in[0,T]}$ is actually well-defined.

Since $u_m^{\varepsilon,\delta} \in C^{1,2,\alpha}(\overline{\mathcal{O}}_m)$ and $\psi_\varepsilon \in C^2(\mathbb{R})$, then both the drift and diffusion coefficients of the controlled SDE are Lipschitz in space (recall Assumption 3.4). Hence (3.31) admits a unique strong solution. Moreover, both n^* and ν^* are progressively measurable and since $0 \leq \psi_\varepsilon' \leq 1/\varepsilon$ then also $0 \leq \dot{\nu}_s^* \leq 2\varepsilon^{-1}\|\nabla u_m^{\varepsilon,\delta}\|_m < \infty$. Therefore $(n^*,\nu^*) \in \mathcal{A}_t^\circ$. □

Remark 3.10: From the proof of Proposition 3.9 we see that w^* defined in (3.26) is optimal for the maximiser and (n^*,ν^*) defined in (3.30) is optimal for the minimiser in the ZSG with payoff (3.22). ■

We also show that $u_m^{\varepsilon,\delta}$ is the value function of a control problem with a recursive structure.

Proposition 3.11 *Let $u_m^{\varepsilon,\delta}$ be a solution of Problem B. Then, for $(t,x) \in \overline{\mathcal{O}}_m$,*

$$u_m^{\varepsilon,\delta}(t,x) = \inf_{(n,\nu)\in\mathcal{A}_t^\circ} \mathsf{E}_x\Big[R_{\rho_m}^{\delta^{-1}} g_m(t+\rho_m, X_{\rho_m}^{[n,\nu]}) + \int_0^{\rho_m} R_s^{\delta^{-1}} h_m(t+s, X_s^{[n,\nu]})\,\mathrm{d}s \\ + \int_0^{\rho_m} R_s^{\delta^{-1}}\big[\tfrac{1}{\delta}(g_m \vee u_m^{\varepsilon,\delta}) + H_m^\varepsilon(\cdot, n_s\dot{\nu}_s)\big](t+s, X_s^{[n,\nu]})\,\mathrm{d}s\Big], \tag{3.32}$$

and the pair (n^*,ν^*) from (3.30) is optimal.

Proof. For simplicity denote $u = u_m^{\varepsilon,\delta}$. Since $\frac{1}{\delta}(g_m - u)^+ + \frac{1}{\delta}u = \frac{1}{\delta}g_m \vee u$, then taking $w \equiv \frac{1}{\delta}$ in (3.24) and using (3.25) we get

$$u(t,x) \le \inf_{(n,\nu)\in\mathcal{A}_t^\circ} \mathsf{E}_x\Big[R_{\rho_m}^{\delta^{-1}} g_m(t+\rho_m, X_{\rho_m}^{[n,\nu]}) \\ + \int_0^{\rho_m} R_s^{\delta^{-1}}\big[h_m + \tfrac{1}{\delta}(g_m \vee u) + H_m^\varepsilon(\cdot, n_s\dot{\nu}_s)\big](t+s, X_s^{[n,\nu]})\,\mathrm{d}s\Big].$$

The equality is obtained by substituting in (3.24) the controls $w \equiv \frac{1}{\delta}$ and (n^*,ν^*) defined in (3.30). Recalling the notation $X^* = X^{[n^*,\nu^*]}$ and (3.29) we obtain (3.32) and optimality of (n^*,ν^*). □

From the probabilistic representation of $u_m^{\varepsilon,\delta}$ in (3.23) we establish uniqueness in Problem B.

Corollary 3.12 *There is at most one solution to Problem B.*

3.2.2 Quadratic growth and stability

Here we establish growth and stability results for $u_m^{\varepsilon,\delta}$.

Lemma 3.13 *Let $u_m^{\varepsilon,\delta}$ be a solution of Problem B. Then, there is a constant $K_3 > 0$ independent of ε, δ, m such that*

$$0 \le u_m^{\varepsilon,\delta}(t,x) \le K_3(1+|x|_d^2), \qquad \textit{for all } (t,x) \in \overline{\mathcal{O}}_m. \tag{3.33}$$

Proof. Since f_m, g_m and h_m are non-negative, then by (3.23) we get $u_m^{\varepsilon,\delta} \ge 0$. For the upper bound we pick the control pair $(n,\nu) = (e_1, 0)$ (where e_1 is the unit vector with 1 in the first entry) in (3.22) and we call $X^0 = X^{[e_1,0]}$ for simplicity.

Notice that $H^{\varepsilon}_m(t+s, X^0_s, 0) = 0$, then

$$
\begin{aligned}
&u^{\varepsilon,\delta}_m(t,x)\\
&\le \sup_{w\in\mathcal{T}^\delta} \mathsf{E}_x\Big[\int_0^{\rho_m} R^w_s\big[h_m + w_s g_m\big](t+s, X^0_s)\,\mathrm{d}s + R^w_{\rho_m} g_m(t+\rho_m, X^0_{\rho_m})\Big]\\
&\le \sup_{w\in\mathcal{T}^\delta} \mathsf{E}_x\Big[\int_0^{\rho_m} R^w_s\Big[K_1(1+w_s)\Big(1+\sup_{0\le\lambda\le s}\big|X^0_\lambda\big|^2_d\Big)\Big]\,\mathrm{d}s + R^w_{\rho_m} K_1\Big(1+\big|X^0_{\rho_m}\big|^2_d\Big)\Big]\\
&\le K_1 \sup_{w\in\mathcal{T}^\delta} \mathsf{E}_x\bigg[\Big[1+\sup_{0\le s\le T-t} e^{-rs}\big|X^0_s\big|^2_d\Big]\Big[e^{-\int_0^{\rho_m} w_\lambda \mathrm{d}\lambda} + \int_0^{\rho_m} e^{-\int_0^s w_\lambda \mathrm{d}\lambda}\big[1+w_s\big]\,\mathrm{d}s\Big]\bigg],
\end{aligned}
$$

where the second inequality is using the quadratic growth of h_m and g_m (see (3.12)); the last inequality is using the definition of R^w with

$$
R^w_s(1+x) = e^{-\int_0^s w_\lambda \mathrm{d}\lambda} e^{-rs}(1+x) \le e^{-\int_0^s w_\lambda \mathrm{d}\lambda}(1+e^{-rs}x), \quad \forall x\in[0,\infty).
$$

For P_x-a.e. ω we have the simple bound

$$
\begin{aligned}
&e^{-\int_0^{\rho_m(\omega)} w_\lambda \mathrm{d}\lambda} + \int_0^{\rho_m(\omega)} e^{-\int_0^s w_\lambda(\omega)\mathrm{d}\lambda}\big[1+w_s(\omega)\big]\,\mathrm{d}s\\
&\qquad = e^{-\int_0^{\rho_m(\omega)} w_\lambda \mathrm{d}\lambda} + \int_0^{\rho_m(\omega)} e^{-\int_0^s w_\lambda(\omega)\mathrm{d}\lambda}\,\mathrm{d}s + \Big[-e^{-\int_0^s w_\lambda(\omega)\mathrm{d}\lambda}\Big]_0^{\rho_m(\omega)}\\
&\qquad \le T+1.
\end{aligned}
$$

Therefore

$$
u^{\varepsilon,\delta}_m(t,x) \le K_1(T+1)\mathsf{E}_x\Big[1+\sup_{0\le s\le T-t} e^{-rs}\big|X^0_s\big|^2_d\Big] \le K_3(1+|x|^2_d),
$$

by standard estimates for SDEs with coefficients with linear growth ([48, Cor. 2.5.10]) and the constant $K_3>0$ depends only on T, D_1 and K_1 in Assumptions 3.4 and 3.5. □

Remark 3.14: In particular, (3.33) implies that for any $m \ge m_0 \in \mathbb{N}$ and $\varepsilon,\delta\in(0,1)$

$$
\big\|u^{\varepsilon,\delta}_m\big\|_{m_0} \le K_3(1+|m_0|^2) =: M_1(m_0),
$$

where we recall $\|\cdot\|_{m_0} = \|\cdot\|_{C^0(\overline{\mathcal{O}}_{m_0})}$. ■

The next result relies upon standard PDE arguments. Its proof is in Appendix B.3 for completeness.

Lemma 3.15 *Let $u^{\varepsilon,\delta}_m$ be a solution of Problem B. Let $u^n \in C^\infty(\overline{\mathcal{O}}_m)$ and $\chi_n \in C^\infty(\mathbb{R})$ be such that $\chi_n(0) = 0$, $\chi_n' \geq 0$, $(\chi_n)_{n\in\mathbb{N}}$ are equi-Lipschitz, and*

$$\|u^n - u^{\varepsilon,\delta}_m\|_{C^{1,2,\gamma}(\overline{\mathcal{O}}_m)} + \|\chi_n - (\cdot)^+\|_{C^0(\mathbb{R})} \leq \tfrac{1}{n}, \quad n \in \mathbb{N},$$

for some $\gamma \in (0, \alpha)$. Then, there exists a unique solution $w^n \in C^{1,2,\alpha}(\overline{\mathcal{O}}_m)$ of

$$\begin{cases} (\partial_t + \mathcal{L} - r)w^n = -h_m - \frac{1}{\delta}\chi_n\big(g_m - u^{\varepsilon,\delta}_m\big) \\ \qquad\qquad\qquad +\psi_\varepsilon(\mathcal{H}(\nabla u^n) - f^2_m - \frac{1}{n}), & \text{on } \mathcal{O}_m, \\ w^n(t,x) = g_m(t,x), & (t,x) \in \partial_P\mathcal{O}_m. \end{cases} \tag{3.34}$$

Moreover, $w^n \in C^{1,3,\alpha}_{Loc}(\mathcal{O}_m)$ and $w^n \to u^{\varepsilon,\delta}_m$ in $C^{1,2,\beta}(\overline{\mathcal{O}}_m)$ as $n \to \infty$, for all $\beta \in (0, \alpha)$.

3.2.3 Gradient bounds

Our next goal is to find a bound for the norm of the gradient of $u^{\varepsilon,\delta}_m$ uniformly in ε, δ. We start by considering an estimate on the parabolic boundary $\partial_P\mathcal{O}_m$ that will be later used to bound $u^{\varepsilon,\delta}_m$ on the whole $\overline{\mathcal{O}}_m$.

Lemma 3.16 *Let $u^{\varepsilon,\delta}_m$ be a solution of Problem B. Then, there is $M_2 = M_2(m) > 0$ such that*

$$\sup_{(t,x)\in\partial_P\mathcal{O}_m} |\nabla u^{\varepsilon,\delta}_m(t,x)|_d \leq M_2, \quad \text{for all } \varepsilon, \delta \in (0,1). \tag{3.35}$$

Proof. For simplicity we denote $u = u^{\varepsilon,\delta}_m$. If $t = T$ we have $u(T,x) = g_m(T,x)$ for $x \in \overline{B}_m$ and the bound is trivial because $\nabla u(T,x) = \nabla g_m(T,x)$.

Next, let $t \in [0,T)$. Notice that $u|_{\partial B_m} = g_m|_{\partial B_m} = 0$. Fix $x \in B_m$ and $y \in \partial B_m$. Then

$$\begin{aligned} 0 \leq u(t,x) - u(t,y) &= u(t,x) - g_m(t,y) \\ &\leq \|\nabla g_m\|_m |x-y|_d + u(t,x) - g_m(t,x), \end{aligned} \tag{3.36}$$

where we recall the definition $\|\cdot\|_m = \|\cdot\|_{C^0(\overline{\mathcal{O}}_m)}$.

For arbitrary $(n,\nu)\in\mathcal{A}^\circ_t$ and $w\in\mathcal{T}^\delta_t$, Dynkin's formula gives

$$g_m(t,x) = \mathsf{E}_x\bigg[R^w_{\rho_m} g_m(t+\rho_m, X^{[n,\nu]}_{\rho_m}) - \int_0^{\rho_m} R^w_s \langle n_s\dot{\nu}_s, \nabla g_m(t+s, X^{[n,\nu]}_s)\rangle\,\mathrm{d}s \\ - \int_0^{\rho_m} R^w_s\big[\partial_t g_m + \mathcal{L}g_m - rg_m - w_s g_m\big](t+s, X^{[n,\nu]}_s)\,\mathrm{d}s\bigg]. \quad (3.37)$$

Then, setting $\Theta_m = \partial_t g_m + \mathcal{L}g_m - rg_m + h_m$ and recalling (3.23), we can write

$$u(t,x) - g_m(t,x) \\ = \inf_{(n,\nu)\in\mathcal{A}^\circ_t}\sup_{w\in\mathcal{T}^\delta_t}\mathsf{E}_x\bigg[\int_0^{\rho_m} R^w_s\Big(\big[\Theta_m + \langle n_s\dot{\nu}_s, \nabla g_m\rangle\big] + H^\varepsilon_m(\cdot, n_s\dot{\nu}_s)\Big)(t+s, X^{[n,\nu]}_s)\,\mathrm{d}s\bigg].$$

Picking $(n,\nu)=(e_1,0)$ and recalling that $H^\varepsilon_m(\cdot,0)=0$ and $\rho_m=\rho_{\mathcal{O}_m}$ we obtain the upper bound

$$u(t,x) - g_m(t,x) \le \sup_{w\in\mathcal{T}^\delta_t}\mathsf{E}_x\bigg[\int_0^{\rho_{\mathcal{O}_m}} R^w_s\Theta_m(t+s, X^{[e_1,0]}_s)\,\mathrm{d}s\bigg] \le \big\|\Theta_m\big\|_m \mathsf{E}_x\big[\rho_{\mathcal{O}_m}\big].$$

Combining the latter with (3.36) we obtain

$$0 \le u(t,x) - u(t,y) \le \|\nabla g_m\|_m |x-y|_d + \|\Theta_m\|_m \mathsf{E}_x\big[\rho_{\mathcal{O}_m}\big].$$

Since $\rho_{\mathcal{O}_m} = \rho_{\mathcal{O}_m}(t,x;e_1,0)$ is associated to the control pair $(n,\nu)=(e_1,0)$, then

$$\rho_{\mathcal{O}_m} = \inf\{s\ge 0\,|\, X^{[e_1,0]}_s \notin B_m\}\wedge(T-t) =: \tau_m\wedge(T-t),$$

and clearly $\mathsf{E}_x[\rho_{\mathcal{O}_m}] \le \mathsf{E}_x[\tau_m] =: \pi(x)$. It is well-known that $\pi\in C^2(\overline{B}_m)$ and it solves

$$\mathcal{L}\pi(x) = -1 \quad \text{for } x\in B_m \quad \text{with} \quad \pi(x)=0 \quad \text{for } x\in\partial B_m,$$

by uniform ellipticity of $\mathcal{L}$ on B_m (see [30, Thm. 6.14]). That is sufficient to conclude

$$\mathsf{E}_x\big[\rho_{\mathcal{O}_m}\big] \le \pi(x) = \pi(x) - \pi(y) \le L_{\pi,m}|x-y|_d,$$

for some constant $L_{\pi,m} > 0$ depending only on the coefficients of $\mathcal{L}$ and the radius m. Then, for all $t \in [0,T)$ we have $0 \le u(t,x) - u(t,y) \le M_2|x-y|_d$,with $M_2 = \|\nabla g_m\|_m + \|\Theta_m\|_m L_{\pi,m}$. This implies (3.35) because $u \in C^{1,2,\alpha}(\overline{\mathcal{O}}_m)$. $\square$

Using Lemma 3.16 we can also provide a bound on $|\nabla u_m^{\varepsilon,\delta}|_d$ in the whole domain $\overline{\mathcal{O}}_m$. It is useful to recall that a function $\varphi \in C^{1,2}(\overline{\mathcal{O}}_m)$ attaining a maximum at a point $(t_0,x_0) \in \mathcal{O}_m$ also satisfies

$$\mathcal{L}\varphi(t_0,x_0) + \partial_t\varphi(t_0,x_0) \le 0, \tag{3.38}$$

by the maximum principle (see [29, Lemma 2.1]). Since $(t_0,x_0) \in \mathcal{O}_m$, then $\nabla\varphi(t_0,x_0) = \mathbf{0}$ and, when $t_0 \in (0,T)$, also $\partial_t\varphi(t_0,x_0) = 0$. We also provide a probabilistic proof of the maximum principle in Appendix (see Lemma B.3).

Proposition 3.17 *Let $u_m^{\varepsilon,\delta}$ be a solution of Problem B. Then, there is $M_3 = M_3(m)$ such that*

$$\sup_{(t,x)\in\overline{\mathcal{O}}_m} |\nabla u_m^{\varepsilon,\delta}(t,x)|_d \le M_3, \quad \textit{for all } \varepsilon,\delta \in (0,1). \tag{3.39}$$

Proof. This proof refines and extends arguments from [69, Lemma A.2] (see also [45, Lemma 2.8]). For simplicity we denote $u = u_m^{\varepsilon,\delta}$. Let $\lambda_m \in (0,\infty)$ be a constant depending on m but independent of ε, δ, which will be chosen later. Let $v^\lambda \in C^{0,1,\alpha}(\overline{\mathcal{O}}_m)$ be defined as

$$v^\lambda(t,x) := |\nabla u(t,x)|_d^2 - \lambda u(t,x)$$

for some $\lambda \in (0,\lambda_m]$. Recalling $M_1 = M_1(m)$ from Remark 3.14 we have for any $\lambda \in (0,\lambda_m]$

$$\sup_{(t,x)\in\overline{\mathcal{O}}_m} |\nabla u(t,x)|_d^2 \le \sup_{(t,x)\in\overline{\mathcal{O}}_m} v^\lambda(t,x) + \lambda_m M_1(m). \tag{3.40}$$

Let $(t^\lambda,x^\lambda) \in \overline{\mathcal{O}}_m$ be a maximum point for v^λ. Two situations may arise: either $(t^\lambda,x^\lambda) \in \mathcal{O}_m$ or $(t^\lambda,x^\lambda) \in \partial_P\mathcal{O}_m$. If $(t^\lambda,x^\lambda) \in \partial_P\mathcal{O}_m$, then by Lemmas

3.13 and 3.16 we have

$$v^\lambda(t^\lambda,x^\lambda)\le|\nabla u(t^\lambda,x^\lambda)|_d^2\le M_2^2 \implies \sup_{(t,x)\in\overline{\mathcal{O}}_m}|\nabla u(t,x)|_d^2\le M_2^2+\lambda_m M_1,$$

where the implication follows by (3.40). Thus, if

$$\Lambda_m:=\{\lambda\in(0,\lambda_m]:(t^\lambda,x^\lambda)\in\partial_P\mathcal{O}_m\}\neq\varnothing, \tag{3.41}$$

it is sufficient to pick $\lambda\in\Lambda_m$ and (3.39) holds.

Let us now assume $\Lambda_m=\varnothing$ (i.e., $(t^\lambda,x^\lambda)\in\mathcal{O}_m$ for all $\lambda\in(0,\lambda_m]$).With no loss of generality:

$$|\nabla u(t^\lambda,x^\lambda)|_d>1 \quad\text{and}\quad \langle\nabla u(t^\lambda,x^\lambda),\nabla g_m(t^\lambda,x^\lambda)\rangle-|\nabla u(t^\lambda,x^\lambda)|_d^2<0. \tag{3.42}$$

If either condition fails, then (3.39) trivially holds. Likewise, we assume

$$\psi_\varepsilon'\big(\mathcal{H}(\nabla u(t^\lambda,x^\lambda))-f_m^2(t^\lambda,x^\lambda)\big)>1 \tag{3.43}$$

because ψ_ε is convex and $\psi_\varepsilon'(r)=\frac{1}{\varepsilon}>1$ for $r\ge 2\varepsilon$. So, if (3.43) fails, it must be $|\nabla u(t^\lambda,x^\lambda)|_d^2=\mathcal{H}(\nabla u(t^\lambda,x^\lambda))<f_m^2(t^\lambda,x^\lambda)+2\varepsilon$ and (3.39) holds because f_m is bounded on $\overline{\mathcal{O}}_m$.

We would like to compute $\partial_t v+\mathcal{L}v$ but the term containing $(\,\cdot\,)^+$ in the PDE for u is not continuously differentiable and, therefore, it is not clear that $|\nabla u|_d^2$ admits classical derivatives. That is why we resort to an approximation procedure. Let u^n, χ_n and w^n be defined as in Lemma 3.15. Recall that $w^n\in C^{1,3,\alpha}_{Loc}(\mathcal{O}_m)\cap C^{1,2,\alpha}(\overline{\mathcal{O}}_m)$ and $w^n\to u$ in $C^{1,2,\gamma}(\overline{\mathcal{O}}_m)$ for all $\gamma\in(0,\alpha)$. Define

$$v^{\lambda,n}(t,x):=|\nabla w^n(t,x)|_d^2-\lambda u(t,x),$$

so that $v^{\lambda,n}\in C^{1,2,\alpha}_{Loc}(\mathcal{O}_m)\cap C^{0,1,\alpha}(\overline{\mathcal{O}}_m)$. Clearly, $v^{\lambda,n}\to v^\lambda$ uniformly on $\overline{\mathcal{O}}_m$.

Let $(t_n^\lambda,x_n^\lambda)_{n\in\mathbb{N}}$ be a sequence with $(t_n^\lambda,x_n^\lambda)\in\arg\max_{\overline{\mathcal{O}}_m}v^{\lambda,n}$ for $n\in\mathbb{N}$. Since $\overline{\mathcal{O}}_m$ is compact, the sequence admits a subsequence $(t_{n_k}^\lambda,x_{n_k}^\lambda)_{k\in\mathbb{N}}$ converging to some $(\tilde t,\tilde x)\in\overline{\mathcal{O}}_m$. It is not hard to show that $(\tilde t,\tilde x)\in\arg\max_{\overline{\mathcal{O}}_m}v^\lambda$ (we provide the full argument in Appendix for completeness). Then, with no loss of generality we can assume $(t^\lambda,x^\lambda)=(\tilde t,\tilde x)$. That implies that we can choose $(t_{n_k}^\lambda,x_{n_k}^\lambda)_{k\in\mathbb{N}}\subset\mathcal{O}_m$.

More precisely, by continuity of v^λ, for any $\eta > 0$ there exist bounded open sets $U_{\lambda,\eta} \subset \mathcal{O}_m$ and $V_{\lambda,\eta} \subset B_m$ such that $(t^\lambda, x^\lambda) \in U_{\lambda,\eta} \cup \big(\{0\} \times V_{\lambda,\eta}\big)$ and $v^\lambda(t,x) > v^\lambda(t^\lambda, x^\lambda) - \eta$ for all $(t,x) \in U_{\lambda,\eta} \cup \big(\{0\} \times V_{\lambda,\eta}\big)$ (by convention, if $t^\lambda \neq 0$ we take $V_{\lambda,\eta} = \varnothing$). Moreover, for $k \in \mathbb{N}$ sufficiently large we have $(t^\lambda_{n_j}, x^\lambda_{n_j})_{j\geq k} \subset U_{\lambda,\eta} \cup \big(\{0\} \times V_{\lambda,\eta}\big)$.

From now on we simply relabel our subsequence by $(t^\lambda_n, x^\lambda_n)_{n\in\mathbb{N}}$ with a slight abuse of notation. Since $(t^\lambda_n, x^\lambda_n)$ is a maximum point of $v^{\lambda,n}$ from the maximum principle (see (3.38)) we get

$$\mathcal{L}v^{\lambda,n}(t^\lambda_n, x^\lambda_n) + \partial_t v^{\lambda,n}(t^\lambda_n, x^\lambda_n) \leq 0. \tag{3.44}$$

Next we compute explicitly all terms in (3.44) and to simplify notation we drop the argument $(t^\lambda_n, x^\lambda_n)$. Denoting $\partial_t v = v_t$, $\partial_{x_i} v = v_{x_i}$ and $\partial_{x_i x_j} v = v_{x_i x_j}$, we obtain

$$\begin{aligned}
v^{\lambda,n}_t &= 2\big\langle \nabla w^n, \nabla w^n_t \big\rangle - \lambda u_t; \\
v^{\lambda,n}_{x_i} &= 2\langle \nabla w^n, \nabla w^n_{x_i} \rangle - \lambda u_{x_i}, && 1 \leq i \leq d; \\
v^{\lambda,n}_{x_i x_j} &= 2\langle \nabla w^n_{x_i}, \nabla w^n_{x_j} \rangle + 2\langle \nabla w^n, \nabla w^n_{x_i x_j} \rangle - \lambda u_{x_i x_j}, && 1 \leq i,j \leq d.
\end{aligned} \tag{3.45}$$

Substituting in (3.44) gives:

$$0 \geq \sum_{i,j=1}^d a_{ij} \langle \nabla w^n_{x_i}, \nabla w^n_{x_j} \rangle + 2 \sum_{k=1}^d w^n_{x_k} \big(\partial_t w^n_{x_k} + \mathcal{L} w^n_{x_k}\big) - \lambda(\partial_t u + \mathcal{L} u). \tag{3.46}$$

From uniform ellipticity (3.10) on B_m and denoting $\theta_{B_m} = \theta$ we have

$$\begin{aligned}
\sum_{i,j=1}^d a_{ij} \langle \nabla w^n_{x_i}, \nabla w^n_{x_j} \rangle &= \sum_{k=1}^d \sum_{i,j=1}^d a_{ij} w^n_{x_k x_i} w^n_{x_k x_j} \\
&\geq \sum_{k=1}^d \theta |\nabla w^n_{x_k}|^2_d \\
&\geq \theta |D^2 w^n|^2_{d\times d}.
\end{aligned} \tag{3.47}$$

To study the second term in (3.46) we introduce the differential operator $\mathcal{L}_{x_k}$:

$$(\mathcal{L}_{x_k}\varphi)(x) = \tfrac{1}{2}\mathrm{tr}\big(a_{x_k}(x) D^2\varphi(x)\big) + \langle b_{x_k}(x), \nabla\varphi(x) \rangle, \quad \text{for } \varphi \in C^2(\mathbb{R}^d), \tag{3.48}$$

where $a_{x_k} \in \mathbb{R}^{d\times d}$ is the matrix with entries $(\partial_{x_k} a_{ij})_{i,j=1}^d$ and $b_{x_k} \in \mathbb{R}^d$ the vector with entries $(\partial_{x_k} b_i)_{i=1}^d$. Differentiating with respect to x_k the PDE in (3.34) and rearranging terms we get

$$\begin{aligned}\partial_t w^n_{x_k} + \mathcal{L} w^n_{x_k} = & \, r w^n_{x_k} + \psi'_\varepsilon(\bar\zeta_n)\cdot(\mathcal{H}(\nabla u^n) - f^2_m)_{x_k} - \partial_{x_k} h_m \\ & - \tfrac{1}{\delta}\chi'_n(g_m - u)\cdot(g_m - u)_{x_k} - \mathcal{L}_{x_k} w^n,\end{aligned} \tag{3.49}$$

where we set $\bar\zeta_n := (\mathcal{H}(\nabla u^n) - f^2_m)(t^\lambda_n, x^\lambda_n) - \frac{1}{n}$ for the argument of ψ'_ε.

In the third term of (3.46) we substitute (3.18) and, combining with (3.47) and (3.49), we obtain

$$\begin{aligned} 0 \ge \theta |D^2 w^n|^2_{d\times d} + 2\Big[& r|\nabla w^n|^2_d + \psi'_\varepsilon(\bar\zeta_n)\langle \nabla w^n, \nabla(\mathcal{H}(\nabla u^n) - f^2_m)\rangle - \langle \nabla w^n, \nabla h_m\rangle \\ & - \tfrac{1}{\delta}\chi'_n(g_m - u)\langle \nabla w^n, \nabla(g_m - u)\rangle - \sum_{k=1}^d w^n_{x_k}\mathcal{L}_{x_k} w^n \Big] \\ & - \lambda\big(ru + \psi_\varepsilon(\zeta_n) - \tfrac{1}{\delta}(g_m - u)^+ - h_m\big),\end{aligned} \tag{3.50}$$

where we set $\zeta_n := (\mathcal{H}(\nabla u) - f^2_m)(t^\lambda_n, x^\lambda_n)$ for the argument of ψ_ε.

Let us denote $\hat w^n = u - w^n$. Then $\|\hat w^n\|_{C^{1,2,\gamma}(\overline{\mathcal{O}}_m)} \to 0$ as $n \to \infty$ because $w^n \to u$ in $C^{1,2,\gamma}(\overline{\mathcal{O}}_m)$, for all $\gamma \in (0, \alpha)$. We claim that

$$\begin{aligned} & 2\Big[r|\nabla w^n|^2_d - \langle \nabla w^n, \nabla h_m\rangle - \tfrac{1}{\delta}\chi'_n(g_m - u)\langle \nabla w^n, \nabla(g_m - u)\rangle - \sum_{k=1}^d w^n_{x_k}\mathcal{L}_{x_k} w^n \Big] \\ & - \lambda\big(ru + \psi_\varepsilon(\zeta_n) - \tfrac{1}{\delta}(g_m - u)^+ - h_m\big) \\ & \ge -C_1|\nabla u|^2_d - \theta|D^2 w^n|^2_{d\times d} - C_2 - \lambda r M_1 - \lambda\psi'_\varepsilon(\bar\zeta_n)\mathcal{H}(\nabla u) - R_n, \end{aligned} \tag{3.51}$$

for M_1 as in Remark 3.14, constants $C_1 = C_1(m) > 0$, $C_2 = C_2(m) > 0$ depending only on m and with

$$0 \le R_n \le \kappa_{\delta,m}\Big(\|\hat w^n\|_{C^{0,1,\gamma}(\overline{\mathcal{O}}_m)} + \lambda\big\|\psi'_\varepsilon(\zeta_n) - \psi'_\varepsilon(\bar\zeta_n)\big\|_m \Big),$$

where $\kappa_{\delta,m} > 0$ depends on δ, $\|\nabla u\|_m$ and $\|\nabla g_m\|_m$. Clearly $R_n \to 0$ as $n \to \infty$. For the ease of exposition the proof of (3.51) is given separately at the end of this proof.

Plugging (3.51) into (3.50) we obtain

$$\begin{aligned} 0 \geq\, & 2\psi'_\varepsilon(\bar{\zeta}_n)\langle \nabla w^n, \nabla(\mathcal{H}(\nabla u^n) - f_m^2)\rangle - C_1|\nabla u|_d^2 - C_2 \\ & - \lambda r M_1 - \lambda \psi'_\varepsilon(\bar{\zeta}_n)\mathcal{H}(\nabla u) - R_n. \end{aligned}$$

By (3.43), $\psi'_\varepsilon(\bar{\zeta}_n) \geq 1$ for large n. Then, multiplying both sides of the inequality by -1 we obtain

$$0 \leq \psi'_\varepsilon(\bar{\zeta}_n)\Big(\lambda\mathcal{H}(\nabla u) - 2\langle \nabla w^n, \nabla(\mathcal{H}(\nabla u^n) - f_m^2)\rangle + C_1|\nabla u|_d^2 + C_2 + \lambda r M_1 + R_n\Big).$$

That implies

$$0 \leq C_1|\nabla u|_d^2 + \lambda\mathcal{H}(\nabla u) - 2\langle \nabla w^n, \nabla(\mathcal{H}(\nabla u^n) - f_m^2)\rangle + C_2 + \lambda r M_1 + R_n. \tag{3.52}$$

We claim that

$$-2\langle \nabla w^n, \nabla(\mathcal{H}(\nabla u^n) - f_m^2)\rangle \leq -2\lambda\mathcal{H}(\nabla u) + 2|\nabla u|_d|\nabla f_m^2|_d + \tilde{R}_n, \tag{3.53}$$

where $\tilde{R}_n = \tilde{R}_n(m, \varepsilon, \delta)$ goes to zero as $n \to \infty$. Again, for ease of exposition the proof of (3.53) is given separately at end of this proof, after the one for (3.51). Collecting $C_1|\nabla u|_d^2$ and $\lambda\mathcal{H}(\nabla u)$ we have that (3.52) becomes

$$0 \leq (C_1 - \lambda)|\nabla u|_d^2 + 2|\nabla u|_d|\nabla f_m^2|_d + C_2 + \lambda r M_1 + R_n + \tilde{R}_n. \tag{3.54}$$

By definition of f_m it is not hard to verify that $|\nabla f_m^2|_d \leq C_3$ for a constant $C_3 = C_3(m) > 0$ independent of ε and δ. Then from the inequality above we obtain

$$(\lambda - C_1)|\nabla u|_d^2 \leq 2C_3|\nabla u|_d + C_2 + \lambda r M_1 + R_n + \tilde{R}_n.$$

Choosing $\lambda = \bar{\lambda} := C_1 + 1$ and recalling our shorthand notation $\nabla u = \nabla u(t_n^\lambda, x_n^\lambda)$, we have

$$|\nabla u(t_n^{\bar{\lambda}}, x_n^{\bar{\lambda}})|_d^2 \leq 2C_3|\nabla u(t_n^{\bar{\lambda}}, x_n^{\bar{\lambda}})|_d + C_2 + \bar{\lambda} r M_1 + R_n + \tilde{R}_n. \tag{3.55}$$

Thanks to (3.42) we have $|\nabla u(t^{\bar{\lambda}}, x^{\bar{\lambda}})|_d > 1$. Thus, with no loss of generality we can assume that $|\nabla u(t,x)|_d \geq 1$ for all $(t,x) \in \overline{U}_{\bar{\lambda},\eta} \cup \big(\{0\}\times\overline{V}_{\bar{\lambda},\eta}\big)$ and recall also that $(t_n^{\bar{\lambda}}, x_n^{\bar{\lambda}})_{n\in\mathbb{N}} \subset U_{\bar{\lambda},\eta} \cup \big(\{0\}\times V_{\bar{\lambda},\eta}\big)$. Thus, dividing by $|\nabla u(t_n^{\bar{\lambda}}, x_n^{\bar{\lambda}})|_d$ in (3.55) we get

$$|\nabla u(t_n^{\bar{\lambda}}, x_n^{\bar{\lambda}})|_d \leq 2C_3 + C_2 + \bar{\lambda} r M_1 + R_n + \tilde{R}_n. \tag{3.56}$$

From (3.56) and the definition of $U_{\bar{\lambda},\eta} \cup (\{0\} \times V_{\bar{\lambda},\eta})$

$$\begin{aligned}\sup_{(t,x)\in\overline{\mathcal{O}}_m} v^{\bar{\lambda}}(t,x) &= v^{\bar{\lambda}}(t^{\bar{\lambda}}, x^{\bar{\lambda}})\\ &\leq v(t_n^{\bar{\lambda}}, x_n^{\bar{\lambda}}) + \eta\\ &\leq (2C_3 + C_2 + \bar{\lambda} r M_1 + R_n + \tilde{R}_n)^2 + \eta.\end{aligned}$$

Letting $n \to \infty$ we obtain

$$\sup_{(t,x)\in\overline{\mathcal{O}}_m} v^{\bar{\lambda}}(t,x) \leq (2C_3 + C_2 + \bar{\lambda} r M_1)^2 + \eta.$$

By the arbitrariness of η and since (3.40) holds for any $\lambda \in (0, \lambda_m]$, taking $\lambda_m = \bar{\lambda}$ we have

$$\sup_{(t,x)\in\overline{\mathcal{O}}_m} |\nabla u(t,x)|_d^2 \leq (2C_3 + C_2 + \bar{\lambda} r M_1)^2 + \bar{\lambda} M_1.$$

Hence, the proposition holds with $M_3 = ((2C_3 + C_2 + \bar{\lambda} r M_1)^2 + \bar{\lambda} M_1)^{1/2}$ independent of ε and δ. □

Proof of (3.51). Recalling $\hat{w}^n = u - w^n$ we first notice

$$\begin{aligned}|\nabla w^n|_d^2 &= \big(|\nabla u|_d^2 - 2\langle \nabla u, \nabla \hat{w}^n\rangle + |\nabla \hat{w}^n|_d^2\big)\\ &\leq |\nabla u|_d^2 + \|\nabla \hat{w}^n\|_m \big(2\|\nabla u\|_m + \|\nabla \hat{w}^n\|_m\big).\end{aligned} \tag{3.57}$$

The first term on the left-hand side of (3.51) is positive. Let us look at the second and third term on the left-hand side of (3.51). For the former we have

$$\langle \nabla w^n, \nabla h_m\rangle \leq |\nabla w^n|_d |\nabla h_m|_d \leq \tfrac{1}{2}|\nabla w^n|_d^2 + \tfrac{1}{2}|\nabla h_m|_d^2. \tag{3.58}$$

For the latter, notice

$$\begin{aligned}\langle \nabla w^n, \nabla(g_m - u)\rangle &= \langle \nabla(u-\hat{w}^n), \nabla(g_m-u)\rangle \\ &= \langle \nabla u, \nabla g_m\rangle - |\nabla u|_d^2 - \langle \nabla \hat{w}^n, \nabla(g_m-u)\rangle \\ &\le \langle \nabla \hat{w}^n, \nabla(u-g_m)\rangle \\ &\le \|\hat{w}^n\|_{C^{0,1,\gamma}(\overline{\mathcal{O}}_m)}(\|\nabla u\|_m + \|\nabla g_m\|_m),\end{aligned}$$

where the first inequality holds by (3.42) in $U_{\lambda,\eta} \cup (\{0\}\times V_{\lambda,\eta})$. Since $0 \le \chi_n' \le 2$, then

$$\tfrac{1}{\delta}\chi_n'(g_m-u)\langle \nabla w^n, \nabla(g_m-u)\rangle \le \tfrac{2}{\delta}\|\hat{w}^n\|_{C^{0,1,\gamma}(\overline{\mathcal{O}}_m)}(\|\nabla u\|_m + \|\nabla g_m\|_m), \quad (3.59)$$

and this term can be collected into R_n in (3.51).

The fourth term on the left-hand side of (3.51), recalling (3.48) we have

$$\begin{aligned}\sum_{k=1}^d w^n_{x_k}\mathcal{L}_{x_k}w^n &= \tfrac{1}{2}\sum_{i,j=1}^d \langle \nabla w^n, \nabla a_{ij}\rangle w^n_{x_ix_j} + \sum_{i=1}^d \langle \nabla w^n, \nabla b_i\rangle w^n_{x_i} \\ &\le \tfrac{d^2}{2}A_m|\nabla w^n|_d|D^2w^n|_{d\times d} + dA_m|\nabla w^n|_d^2,\end{aligned} \quad (3.60)$$

where we used Cauchy-Schwartz inequality and set

$$A_m := \max_{i,j}\Big(\big\|\nabla a_{ij}\big\|_{C^0(\overline{B}_m)} + \big\|\nabla b_i\big\|_{C^0(\overline{B}_m)}\Big). \quad (3.61)$$

Using that $d^2A_m|\nabla w^n|_d|D^2w^n|_{d\times d} \le \theta^{-1}d^4A_m^2|\nabla w^n|_d^2 + \theta|D^2w^n|^2_{d\times d}$, with $\theta = \theta_{B_m}$ as in (3.10), and combining (3.58), (3.60) and (3.57), we have

$$\begin{aligned}&\langle \nabla w^n, \nabla h_m\rangle + \sum_{k=1}^d w^n_{x_k}\mathcal{L}_{x_k}w^n \le \tfrac{1}{2}C_1|\nabla w^n|_d^2 + \tfrac{1}{2}\theta|D^2w^n|^2_{d\times d} + \tfrac{1}{2}C_2 \\ &\le \tfrac{1}{2}\Big[C_1|\nabla u|_d^2 + \theta|D^2w^n|^2_{d\times d} + C_2 + C_1\|\nabla \hat{w}^n\|_m\big(2\|\nabla u\|_m + \|\nabla \hat{w}^n\|_m\big)\Big],\end{aligned} \quad (3.62)$$

with

$$C_1 = C_1(m) := 1 + d^4A_m^2\theta^{-1} + 2dA_m \quad \text{and} \quad C_2 = C_2(m) := \|\nabla h_m\|_m^2. \quad (3.63)$$

The expression involving $\|\nabla \hat{w}^n\|_m$ can be collected into R_n in (3.51).

It remains to find an upper bound for $\lambda\big(ru+\psi_\varepsilon(\zeta_n)-\frac{1}{\delta}(g_m-u)^+-h_m\big)$. Since $h_m\geq 0$ and $(g_m-u)^+\geq 0$ and taking $M_1=M_1(m)>0$ as in Remark 3.14 we have

$$\lambda\big(ru+\psi_\varepsilon(\zeta_n)-\tfrac{1}{\delta}(g_m-u)^+-h_m\big)\leq\lambda\big(rM_1+\psi_\varepsilon(\zeta_n)\big). \tag{3.64}$$

By convexity of ψ_ε and since $\psi_\varepsilon(0)=0$, we have

$$\begin{aligned}\psi_\varepsilon(\zeta_n)&\leq\psi'_\varepsilon(\zeta_n)(\mathcal{H}(\nabla u)-f_m^2)\\&\leq\psi'_\varepsilon(\zeta_n)\mathcal{H}(\nabla u)\\&\leq\psi'_\varepsilon(\bar\zeta_n)\mathcal{H}(\nabla u)+\big|\psi'_\varepsilon(\zeta_n)-\psi'_\varepsilon(\bar\zeta_n)\big|\mathcal{H}(\nabla u)\\&\leq\psi'_\varepsilon(\bar\zeta_n)\mathcal{H}(\nabla u)+\big\|\psi'_\varepsilon(\zeta_n)-\psi'_\varepsilon(\bar\zeta_n)\big\|_m\|\nabla u\|_m^2.\end{aligned}\tag{3.65}$$

Combining (3.59), (3.62), (3.64) and (3.65) we obtain (3.51). □

Proof of (3.53). By Cauchy-Schwartz inequality and recalling that $\hat w^n=u-w^n$ we have

$$\begin{aligned}&-2\left\langle\nabla w^n,\nabla(\mathcal{H}(\nabla u^n)-f_m^2)\right\rangle\\&\quad\leq-2\langle\nabla w^n,\nabla(\mathcal{H}(\nabla u^n))\rangle+2|\nabla w^n|_d|\nabla f_m^2|_d\\&\quad\leq-2\langle\nabla w^n,\nabla(\mathcal{H}(\nabla u^n))\rangle+2|\nabla u|_d|\nabla f_m^2|_d+2\|\nabla\hat w^n\|_m\|\nabla f_m^2\|_m.\end{aligned}\tag{3.66}$$

The k-th entry of the vector $\nabla(\mathcal{H}(\nabla u^n))$ reads $\big(\mathcal{H}(\nabla u^n)\big)_{x_k}=\langle\nabla\mathcal{H}(\nabla u^n),\nabla u^n_{x_k}\rangle$ and therefore the first term on the right-hand side of (3.66) can be written

$$\begin{aligned}w^n_{x_k}\big(\mathcal{H}(\nabla u^n)\big)_{x_k}&=w^n_{x_k}\langle\nabla\mathcal{H}(\nabla u^n),\nabla u^n_{x_k}\rangle\\&=w^n_{x_k}\langle\nabla\mathcal{H}(\nabla u^n)-\nabla\mathcal{H}(\nabla u),\nabla u^n_{x_k}\rangle\\&\quad+w^n_{x_k}\langle\nabla\mathcal{H}(\nabla u),\nabla u^n_{x_k}-\nabla u_{x_k}\rangle+w^n_{x_k}\langle\nabla\mathcal{H}(\nabla u),\nabla u_{x_k}\rangle\\&\geq-C_{\nabla\mathcal{H}}\|u^n-u\|_{C^{1,2,\gamma}(\overline{\mathcal{O}}_m)}\|\nabla u\|_m\big(\|D^2u^n\|_m+\|\nabla u\|_m\big)\\&\quad+2w^n_{x_k}\langle\nabla\mathcal{H}(\nabla u),\nabla u_{x_k}\rangle,\end{aligned}\tag{3.67}$$

where $C_{\nabla\mathcal{H}}$ is the Lipschitz constant of $\nabla\mathcal{H}$, i.e., 2. The last term above is estimated as

$$\begin{aligned} w^n_{x_k}\langle \nabla\mathcal{H}(\nabla u), \nabla u_{x_k}\rangle &= w^n_{x_k}\langle \nabla\mathcal{H}(\nabla u), \nabla(\hat{w}^n_{x_k} + w^n_{x_k})\rangle \\ &= w^n_{x_k}\langle \nabla\mathcal{H}(\nabla u), \nabla\hat{w}^n_{x_k}\rangle + w^n_{x_k}\langle \nabla\mathcal{H}(\nabla u), \nabla w^n_{x_k}\rangle \\ &\geq - C_{\nabla\mathcal{H}}\|\hat{w}^n\|_{C^{1,2,\gamma}(\overline{\mathcal{O}}_m)}\|\nabla w^n\|_m\|\nabla u\|_m \\ &\quad + w^n_{x_k}\langle \nabla\mathcal{H}(\nabla u), \nabla w^n_{x_k}\rangle. \end{aligned} \tag{3.68}$$

For the last term above, recall that all expressions are evaluated at $(t^\lambda_n, x^\lambda_n)$, which is a stationary point for $v^{\lambda,n}$ in the spatial coordinates. Thus, $2\langle \nabla w^n, \nabla w^n_{x_i}\rangle = \lambda u_{x_i}$ by (3.45). Summing over all k's in (3.68) we get

$$\begin{aligned} \sum_{k=1}^d w^n_{x_k}\langle \nabla\mathcal{H}(\nabla u), \nabla w^n_{x_k}\rangle &= \sum_{k=1}^d\sum_{i=1}^d w^n_{x_k}(\nabla\mathcal{H}(\nabla u))_i w^n_{x_k x_i} \\ &= \sum_{i=1}^d (\nabla\mathcal{H}(\nabla u))_i\langle \nabla w^n, \nabla w^n_{x_i}\rangle \\ &= \sum_{i=1}^d (\nabla\mathcal{H}(\nabla u))_i \lambda u_{x_i} \\ &= \lambda\langle \nabla\mathcal{H}(\nabla u), \nabla u\rangle \\ &\geq \lambda\mathcal{H}(\nabla u), \end{aligned} \tag{3.69}$$

where the last inequality is justified because $\mathcal{H}(\cdot)$ is a convex function such that $\mathcal{H}(0) = 0$, thus $\mathcal{H}(p) \leq \langle \nabla\mathcal{H}(p), p\rangle$. Plugging (3.69) back into (3.66) we get

$$\begin{aligned} &-2\langle \nabla w^n, \nabla(\mathcal{H}(\nabla u^n) - f^2_m)\rangle \\ &\quad\leq - 2\lambda\mathcal{H}(\nabla u) + 2|\nabla u|_d|\nabla f^2_m| + 2\|\nabla\hat{w}^n\|_m\|\nabla f^2_m\|_m \\ &\qquad - C_{\nabla\mathcal{H}}d\|\hat{w}^n\|_{C^{1,2,\gamma}(\overline{\mathcal{O}}_m)}\|\nabla u\|_m\big(\|\nabla w^n\|_m + \|D^2 u\|_m\big) \\ &\qquad - C_{\nabla\mathcal{H}}d\|u^n - u\|_{C^{1,2,\gamma}(\overline{\mathcal{O}}_m)}\|\nabla u\|_m\big(\|D^2 u^n\|_m + \|\nabla u\|_m\big) \end{aligned} \tag{3.70}$$

Since $w^n \to u$ and $u^n \to u$ in $C^{1,2,\gamma}(\overline{\mathcal{O}}_m)$ as $n \to \infty$ for all $\gamma \in (0, \alpha)$ (Lemma 3.15), then all the terms in (3.66) and (3.70) depending on $\hat{w}^n$ and $u^n - u$ can be collected in a remainder $\tilde{R}_n = \tilde{R}_n(m, \varepsilon, \delta)$ that goes to zero as $n \to \infty$. So, (3.66) and (3.70) give us

$$-2\langle \nabla w^n, \nabla(\mathcal{H}(\nabla u^n) - f_m^2)\rangle \le -2\lambda\mathcal{H}(\nabla u) + 2|\nabla u|_d |\nabla f_m^2|_d + \tilde{R}_n,$$

as claimed in (3.53). □

3.2.4 Solution to the Penalised Problem

In this section we prove that Problem B admits a unique solution. The proof is based on a fixed point argument that requires the *a priori* estimates derived in Sections 3.2.2 and 3.2.3 as well as the next bound.

Proposition 3.18 *Let $u_m^{\varepsilon,\delta}$ be a solution of Problem B. For any $\beta \in (0,1)$ there is $M_4 = M_4(m,\varepsilon,\delta,\beta)$ such that*

$$\|u_m^{\varepsilon,\delta}\|_{C^{0,1,\beta}(\overline{\mathcal{O}}_m)} \le M_4. \tag{3.71}$$

Proof. Let $u = u_m^{\varepsilon,\delta}$ for simplicity. Then, u can be seen as the unique solution φ of the linear PDE

$$\partial_t\varphi + \mathcal{L}\varphi - r\varphi = -h_m - \tfrac{1}{\delta}(g_m - u)^+ + \psi_\varepsilon(\mathcal{H}(\nabla u) - f_m^2), \quad \text{on } \mathcal{O}_m$$

with boundary conditions $\varphi(t,x) = 0$ for $x \in \partial B_m$ and $\varphi(T,x) = g_m(T,x)$ for all $x \in B_m$. The theory of strong solutions for linear parabolic PDEs (see [8, Thm. 2.6.5 and Rem. 2.6.4]) gives

$$\begin{aligned}\|u\|_{W^{1,2,p}(\mathcal{O}_m)} \le C\Big(&\big\|h_m + \tfrac{1}{\delta}(g_m - u)^+ - \psi_\varepsilon(\mathcal{H}(\nabla u) - f_m^2)\big\|_{L^p(\mathcal{O}_m)} \\ &+ \|g_m\|_{W^{1,2,p}(\mathcal{O}_m)}\Big)\end{aligned} \tag{3.72}$$

for any $p \in (1,\infty)$, with C a constant independent of ε and δ. Denoting $|\mathcal{O}_m|$ the volume of the set $\mathcal{O}_m$, thanks to Proposition 3.17 and $u \ge 0$ we have

$$\begin{aligned}\|u\|_{W^{1,2,p}(\mathcal{O}_m)} \le\, & C|\mathcal{O}_m|^{\frac{1}{p}}\Big(\big\|h_m + \tfrac{1}{\delta}g_m\big\|_m + \tfrac{1}{\varepsilon}M_3^2\Big) \\ &+ C\|g_m\|_{W^{1,2,p}(\mathcal{O}_m)} < \infty,\end{aligned}$$

having also used that $\psi_\varepsilon(x) \le \frac{1}{\varepsilon}x$ for all $x \ge 0$.

Since p is arbitrary, then Sobolev embedding (1.6) guarantees that for any $\beta \in (0,1)$ there exists a constant $M_4 = M_4(m, \varepsilon, \delta, \beta)$ such that (3.71) holds. □

The next theorem requires two ingredients, a result on the existence and uniqueness for linear parabolic PDE on bounded domain and a fixed point theorem for Banach spaces. The first ingredient comes from [29, Thm. 3.3.7] whose assumptions are:

(i) the coefficients of $\mathcal{L}$ and r are α-Hölder continuous in $\overline{\mathcal{O}}_m$,

(ii) the second-order operator a is uniformly elliptic in $\overline{B}_m$,

(iii) the right-hand side of (3.73) is α-Hölder in $\overline{\mathcal{O}}_m$,

(iv) a so-called compatible conditions which is presented in the proof of the theorem for clarity,

(v) for every point $(t, x) \in \partial B_m$, there exists a neighbourhood V such that all points $(s, y) \in V \cap \partial B_m$ can be represented as

$$y_i = h(t, y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_d) \quad \text{for some } i \in \{1, \ldots, d\},$$

and the function h and its derivatives, $\partial_t h, \partial_{x_j} h, \partial_{tt} h, \partial_{tx_j} h, \partial_{x_j x_k} h$, are α-Hölder continuous in $\overline{V \cap (\{T\} \times \partial B_m)}$ for all $1 \leq j, k \leq d$ with $j, k \neq i$.

In our case, Assumptions (i), (ii) and (iii) are satisfied by Assumption 3.5. Assumption (iv) is presented in the theorem below and Assumption (v) holds because we have that for all $(s, y) \in [0, T] \times \partial B_m$ we can find i such that $|y_i| \geq \frac{1}{\sqrt{d}}$. Defining $V_y := \{x \in \mathbb{R}^d : |x - y|_d < \frac{1}{\sqrt{2d}}\}$, we have that for $(t, x) \in [0, T] \times (V_y \cap \partial B_m) =: V_y^T$ we can use the function

$$x_i = h(t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d) := \big(1 - \sum_{j \neq i} x_j^2\big)^{\frac{1}{2}}.$$

We have that $h \geq \frac{\sqrt{2d-1}}{\sqrt{2d}}$ everywhere in V_y^T and it is independent of time. Thus, the derivatives with respect to time are sufficiently regular, the derivatives with respect to the space variables only are continuously differentiable because they are a ratio of infinitely differentiable functions with a non-zero denominator everywhere.

The second ingredient is a Banach space. Let us define the subset of $C^{0,1,\alpha}(\overline{\mathcal{O}}_m)$ as:

$$C_*^{0,1,\alpha}(\overline{\mathcal{O}}_m) := \left\{ \varphi \in C^{0,1,\alpha}(\overline{\mathcal{O}}_m) \middle| \begin{array}{l} \varphi = 0 \text{ on } \{T\} \times \partial B_m \text{ and} \\ \nabla\varphi = \mathbf{0} \text{ on } \{T\} \times \partial B_m \end{array} \right\}.$$

Lemma 3.19 *The space $(C_*^{0,1,\alpha}(\overline{\mathcal{O}}_m), \|\cdot\|_{C^{0,1,\alpha}(\overline{\mathcal{O}}_m)})$ is a Banach space.*

Proof. The space $C_*^{0,1,\alpha}(\overline{\mathcal{O}}_m)$ is a subset of $C^{0,1,\alpha}(\overline{\mathcal{O}}_m)$ which is a Banach space with respect to $\|\cdot\|_{C^{0,1,\alpha}(\overline{\mathcal{O}}_m)}$. Thus, it is enough to show that $C_*^{0,1,\alpha}(\overline{\mathcal{O}}_m)$ is closed under the same norm. This implies that $(C_*^{0,1,\alpha}(\overline{\mathcal{O}}_m), \|\cdot\|_{C^{0,1,\alpha}(\overline{\mathcal{O}}_m)})$ is a Banach space

Let $(\varphi_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $C_*^{0,1,\alpha}(\overline{\mathcal{O}}_m)$ convergent to $\varphi \in C^{0,1,\alpha}(\overline{\mathcal{O}}_m)$. Since $\varphi_n = 0$ and $\nabla\varphi_n = \mathbf{0}$ for all $n \in \mathbb{N}$ and $\varphi_n \to \varphi$ in $C^{0,1,\alpha}(\overline{\mathcal{O}}_m)$ we have that $\varphi \in C_*^{0,1,\alpha}(\overline{\mathcal{O}}_m)$. □

Remark 3.20: Defining $C_*^{1,2,\alpha}(\overline{\mathcal{O}}_m) := (C^{1,2,\alpha} \cap C_*^{0,1,\alpha})(\overline{\mathcal{O}}_m)$, we have that also the pair $(C_*^{1,2,\alpha}(\overline{\mathcal{O}}_m), \|\cdot\|_{C^{1,2,\alpha}(\overline{\mathcal{O}}_m)})$ is a Banach space. ■

Theorem 3.21 *There exists a unique solution $u_m^{\varepsilon,\delta}$ of Problem B.*

Proof. Uniqueness is by Corollary 3.12. Existence will be proved refining arguments from [45, Prop. 1.2]. Fix $\varepsilon, \delta \in (0,1)$ and $m \in \mathbb{N}$.

Given $\varphi \in C_*^{0,1,\alpha}(\overline{\mathcal{O}}_m)$ we consider the linear partial differential equation for $w = w^\varphi$:

$$\begin{cases} \partial_t w + \mathcal{L}w - rw = -h_m - \frac{1}{\delta}(g_m - \varphi)^+ + \psi_\varepsilon\big(\mathcal{H}(\nabla\varphi) - f_m^2\big), & \text{on } \mathcal{O}_m, \\ w(t,x) = g_m(t,x), & (t,x) \in \partial_P\mathcal{O}_m. \end{cases} \tag{3.73}$$

For $x \in \partial B_m$ the compatibility condition

$$\lim_{s\uparrow T}(\partial_t g_m + \mathcal{L}g_m - rg_m)(s,x) = \Big[-h_m - \tfrac{1}{\delta}\big(g_m - \varphi\big)^+ + \psi_\varepsilon\big(\mathcal{H}(\nabla\varphi) - f_m^2\big)\Big](T,x),$$

holds, with both sides of the expression being equal to zero. Indeed, on the left-hand side, properties of the cut-off function $\xi_{m-1} \in C_c^\infty(B_m)$ guarantee $g_m = \partial_{x_i} g_m = \partial_{x_i,x_j} g_m = 0$ on $[0,T] \times \partial B_m$, hence also $\partial_t g_m = 0$. On the right-hand side of the equation, we use that $h_m = \varphi = \partial_{x_i}\varphi = 0$ on $\{T\} \times \partial B_m$. Therefore (3.73) admits a unique solution in $C^{1,2,\alpha}(\overline{\mathcal{O}}_m)$ by [29, Thm. 3.3.7]. The boundary condition of the PDE implies $w = 0$ and $\nabla w = \mathbf{0}$ on $\{T\} \times \partial B_m$. Hence $w \in C_*^{1,2,\alpha}(\overline{\mathcal{O}}_m)$.

Define the operator $\Gamma : C_*^{0,1,\alpha}(\overline{\mathcal{O}}_m) \to C_*^{0,1,\alpha}(\overline{\mathcal{O}}_m)$ that maps φ to the solution of the PDE (3.73), i.e., $\Gamma[\varphi] = w^\varphi$. Next, we show that Γ has a fixed point by Schaefer's fixed point theorem (stated in Appendix for completeness). So Problem B admits a solution.

We have $\Gamma[\varphi] \in C_*^{1,2,\alpha}(\overline{\mathcal{O}}_m) \implies \Gamma[\varphi] \in C_*^{0,1,\beta}(\overline{\mathcal{O}}_m)$ for all $\beta \in (0,1)$ by (1.6). We must prove that Γ is continuous and compact in $C_*^{0,1,\alpha}(\overline{\mathcal{O}}_m)$. Consider a sequence $(\varphi_n)_{n\in\mathbb{N}} \subset C_*^{0,1,\alpha}(\overline{\mathcal{O}}_m)$ such that $\varphi_n \to \varphi$ in $C_*^{0,1,\alpha}(\overline{\mathcal{O}}_m)$. Let $F : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ be defined as

$$F(q,p) \coloneqq \psi_\varepsilon\big(\mathcal{H}(p) - f_m^2\big) - \tfrac{1}{\delta}(g_m - q)^+.$$

Clearly $|\varphi|$ and $|\nabla\varphi|_d$ are bounded on $\overline{\mathcal{O}}_m$. Since F is locally Lipschitz, then standard estimates for Hölder norms allow to prove $F(\varphi_n, \nabla\varphi_n) \to F(\varphi, \nabla\varphi)$ in $C^\gamma(\overline{\mathcal{O}}_m)$ as $n \to \infty$ for any $\gamma \in (0,\alpha)$. By Lemma B.1 in Appendix, $\Gamma[\varphi_n] \to \Gamma[\varphi]$ in $C_*^{1,2,\gamma'}(\overline{\mathcal{O}}_m)$ for any $\gamma' \in (0,\gamma)$, as $n \to \infty$. Sobolev embedding (1.6) implies $\Gamma[\varphi_n] \to \Gamma[\varphi]$ in $C_*^{0,1,\beta}(\overline{\mathcal{O}}_m)$ for any $\beta \in (0,1)$, hence continuity of Γ.

For compactness, notice that Γ maps bounded sets of $C_*^{0,1,\alpha}(\overline{\mathcal{O}}_m)$ into bounded sets of $C_*^{1,2,\alpha}(\overline{\mathcal{O}}_m)$ by [29, Thms. 3.2.6 and 3.3.7]. Since bounded sets in $C_*^{1,2,\alpha}(\overline{\mathcal{O}}_m)$ are bounded in $C_*^{0,1,\beta}(\overline{\mathcal{O}}_m)$ for all $\beta \in (0,1)$ (see (1.6)), then $C_*^{1,2,\alpha}(\overline{\mathcal{O}}_m) \hookrightarrow C_*^{0,1,\alpha}(\overline{\mathcal{O}}_m)$ is a compact embedding.

It remains to prove that the set

$$\mathcal{B} \coloneqq \Big\{\varphi \in C_*^{0,1,\alpha}(\overline{\mathcal{O}}_m) \Big| \rho\Gamma[\varphi] = \varphi \text{ for some } \rho \in [0,1]\Big\}$$

is bounded in $C_*^{0,1,\alpha}(\overline{\mathcal{O}}_m)$. If $\rho = 0$, then $\varphi = 0$. If $\rho\Gamma[\varphi] = \varphi$ for some $\rho \in (0,1]$, then φ satisfies

$$\begin{cases} \partial_t\varphi + \mathcal{L}\varphi - r\varphi = \rho\big[-h_m - \frac{1}{\delta}\big(g_m - \varphi\big)^+ + \psi_\varepsilon\big(\mathcal{H}(\nabla\varphi) - f_m^2\big)\big], & \text{on } \mathcal{O}_m, \\ \varphi = \rho g_m, & \text{on } \partial_P\mathcal{O}_m. \end{cases} \tag{3.74}$$

The PDE in (3.74) is the same as the one in (3.18) but with h_m, ψ_ε, g_m, $\frac{1}{\delta}$ replaced by ρh_m, $\rho\psi_\varepsilon$, ρg_m, $\frac{\rho}{\delta}$. Then, all the results from this section and Proposition 3.18 apply to φ. In particular, $\|\varphi\|_{C^{0,1,\alpha}(\overline{\mathcal{O}}_m)} \leq M_4$ uniformly for all $\rho \in (0,1]$, with M_4 as in Proposition 3.18. Finally, Schaefer's fixed point theorem guarantees existence of the solution of Problem B, for every treble (ε, δ, m). □

3.3 Penalised Problem on Unbounded Domain and Further Estimates

We refine our *a priori* estimates on the solution of Problem B. First we shall make all bounds independent of m so that we can construct a solution to a penalised problem on unbounded domain as $m \to \infty$. Then we shall find bounds independent of ε and δ so as to pass to the limit for $\varepsilon, \delta \downarrow 0$ and obtain a solution to Problem A.

3.3.1 Estimates independent of m

First we bound $\nabla u_m^{\varepsilon,\delta}$ independently of m on each compact.

Proposition 3.22 *Fix $m_0 \in \mathbb{N}$ and $q \geq m_0 + 2$. Let $u_q^{\varepsilon,\delta}$ be the unique solution of Problem* B *on $\mathcal{O}_q$. Then, there is $N_1 = N_1(m_0)$ independent of ε, δ and q, such that*

$$\sup_{(t,x)\in\overline{\mathcal{O}}_{m_0}} |\nabla u_q^{\varepsilon,\delta}(t,x)|_d \leq N_1. \tag{3.75}$$

Proof. For notational simplicity set $u = u_q^{\varepsilon,\delta}$, $\xi = \xi_{m_0}$ and $\|\cdot\|_0 = \|\cdot\|_{m_0+1}$. Since $q \geq m_0 + 2$, we have $f_q = f$, $g_q = g$ and $h_q = h$ on $\overline{\mathcal{O}}_{m_0+1}$. Let $\lambda_0 \in (0,\infty)$ be a constant depending on m_0 but independent of ε, δ, q, which will be chosen later. Let $v^\lambda \in C^{0,1,\alpha}(\overline{\mathcal{O}}_{m_0+1})$ be defined as

$$v^\lambda(t,x) := \xi(x)\big|\nabla u(t,x)\big|_d^2 - \lambda u(t,x), \qquad \text{for } (t,x) \in \overline{\mathcal{O}}_{m_0+1} \tag{3.76}$$

for some $\lambda \in (0, \lambda_0]$. We will use later that

$$\sup_{(t,x)\in\overline{\mathcal{O}}_{m_0}} |\nabla u(t,x)|_d^2 \le \sup_{(t,x)\in\overline{\mathcal{O}}_{m_0+1}} \xi(x)|\nabla u(t,x)|_d^2 \\ \le \sup_{(t,x)\in\overline{\mathcal{O}}_{m_0+1}} v^\lambda(t,x) + \lambda_0\|u\|_0, \tag{3.77}$$

where we also notice that $\|u\|_0 \le M_1$ with $M_1 = M_1(m_0+1)$ as in Remark 3.14.

Let $(t^\lambda, x^\lambda) \in \overline{\mathcal{O}}_{m_0+1}$ be a maximum point for v^λ. If $x^\lambda \in \partial B_{m_0+1}$ then $\xi(x^\lambda) = 0$ and $v(t^\lambda, x^\lambda) = -\lambda u(t^\lambda, x^\lambda) \le 0$. If $t^\lambda = T$, then $v^\lambda(T, x^\lambda) \le \xi(x^\lambda)|\nabla g(T, x^\lambda)|_d^2 \le \|f\|_0^2$ (see Assumption 3.5). Thus, in both cases (3.75) holds with $N_1 \ge (\|f\|_0^2 + \lambda_0 M_1)^{1/2}$. Defining Λ_{m_0+1} as in (3.41) but with λ_0 instead of λ_m, if $\Lambda_{m_0+1} \ne \varnothing$ the bound holds taking $\lambda \in \Lambda_{m_0+1}$. It remains to consider the case $\Lambda_{m_0+1} = \varnothing$, so that $(t^\lambda, x^\lambda) \in \mathcal{O}_{m_0+1}$ for all $\lambda \in (0, \lambda_0]$.

As in the proof of Proposition 3.17 we use the smooth approximation w^n of u, obtained from (3.34). Analogously, let us define $v^{\lambda,n} := \xi\big|\nabla w^n\big|_d^2 - \lambda u$ in $\overline{\mathcal{O}}_{m_0+1}$ and let $(t_n^\lambda, x_n^\lambda)_{n\in\mathbb{N}}$ be a sequence converging to (t^λ, x^λ) with $(t_n^\lambda, x_n^\lambda) \in \arg\max_{\overline{\mathcal{O}}_{m_0+1}} v^{\lambda,n}$. For any $\eta > 0$ there exists a neighbourhood $U_{\lambda,\eta} \cup \big(\{0\} \times V_{\lambda,\eta}\big)$ of (t^λ, x^λ) such that

$$v^\lambda(t,x) > v^\lambda(t^\lambda, x^\lambda) - \eta, \qquad \text{for all } (t,x) \in U_{\lambda,\eta} \cup \big(\{0\} \times V_{\lambda,\eta}\big) \tag{3.78}$$

and $(t_n^\lambda, x_n^\lambda) \in U_{\lambda,\eta} \cup \big(\{0\} \times V_{\lambda,\eta}\big)$ for sufficiently large n's (with the convention $V_{\lambda,\eta} = \varnothing$ if $t^\lambda \ne 0$).

Taking derivatives of $v^{\lambda,n}$ we have, for $1 \le i, j \le d$,

$$\begin{aligned}
v_t^{\lambda,n} &= 2\xi\langle \nabla w^n, \nabla w_t^n\rangle - \lambda u_t;\\
v_{x_i}^{\lambda,n} &= 2\xi\langle \nabla w^n, \nabla w_{x_i}^n\rangle - \lambda u_{x_i} + \xi_{x_i}|\nabla w^n|_d^2;\\
v_{x_i x_j}^{\lambda,n} &= 2\xi\langle \nabla w_{x_i}^n, \nabla w_{x_j}^n\rangle + 2\xi\langle \nabla w^n, \nabla w_{x_i x_j}^n\rangle\\
&\quad - \lambda u_{x_i x_j} + \xi_{x_i x_j}|\nabla w^n|_d^2 + 2\big(\xi_{x_i}\langle \nabla w^n, \nabla w_{x_j}^n\rangle + \xi_{x_j}\langle \nabla w^n, \nabla w_{x_i}^n\rangle\big).
\end{aligned}$$

Since $(t_n^\lambda, x_n^\lambda) \in \mathcal{O}_{m_0+1}$ then (3.38) gives $(\partial_t v^{\lambda,n} + \mathcal{L} v^{\lambda,n})(t_n^\lambda, x_n^\lambda) \le 0$. Substituting the expressions for the derivatives of $v^{\lambda,n}$, some tedious but straightforward calculations and symmetry of a_{ij} give

$$0 \geq 2\xi\langle \nabla w^n, (\partial_t + \mathcal{L})(\nabla w^n)\rangle - \lambda(\partial_t u + \mathcal{L}u) + \xi \sum_{i,j=1}^{d} a_{ij}\langle \nabla w^n_{x_i}, \nabla w^n_{x_j}\rangle \tag{3.79}$$
$$+(\mathcal{L}\xi)|\nabla w^n|_d^2 + 2\sum_{i,j=1}^{d} a_{ij}\xi_{x_i}\langle \nabla w^n, \nabla w^n_{x_j}\rangle,$$

where $(\partial_t + \mathcal{L})(\nabla w^n)$ denotes the vector with entries $(\partial_t + \mathcal{L})w^n_{x_k}$ for $k = 1, \ldots, d$. Here we omit the dependence on $(t^\lambda_n, x^\lambda_n)$ for notational convenience.

The expressions for $(\partial_t + \mathcal{L})w^n_{x_k}$ and $(\partial_t + \mathcal{L})u$ are the same as in (3.49) and (3.18), respectively. A lower bound for $\sum a_{ij}\langle \nabla w^n_{x_i}, \nabla w^n_{x_j}\rangle$ was also obtained in (3.47) but with θ therein replaced by $\theta = \theta_{B_{m_0+1}}$. Thus, from (3.79) we get

$$0 \geq \xi\theta|D^2 w^n|^2_{d\times d} \tag{3.80}$$
$$+2\xi\Big[r|\nabla w^n|_d^2 + \psi'_\varepsilon(\bar{\zeta}_n)\langle \nabla w^n, \nabla(\mathcal{H}(\nabla u^n) - f^2)\rangle - \tfrac{1}{\delta}\chi'_n(g-u)\langle \nabla w^n, \nabla(g-u)\rangle$$
$$-\langle \nabla w^n, \nabla h\rangle - \sum_{k=1}^{d} w^n_{x_k}\mathcal{L}_{x_k} w^n\Big] - \lambda\big(ru + \psi_\varepsilon(\zeta_n) - h - \tfrac{1}{\delta}(g-u)^+\big)$$
$$+(\mathcal{L}\xi)|\nabla w^n|_d^2 + 2\sum_{i,j=1}^{d} a_{ij}\xi_{x_i}\langle \nabla w^n, \nabla w^n_{x_j}\rangle,$$

where $\bar{\zeta}_n = (\mathcal{H}(\nabla u^n) - f^2)(t^\lambda_n, x^\lambda_n) - \frac{1}{n}$ and $\zeta_n = (\mathcal{H}(\nabla u) - f^2)(t^\lambda_n, x^\lambda_n)$. Up to a factor ξ, the expression above is the analogue of (3.50) but with two additional terms. As in (3.51) we have

$$2\xi\Big[r|\nabla w^n|_d^2 - \tfrac{1}{\delta}\chi'_n(g-u)\langle \nabla w^n, \nabla(g-u)\rangle - \langle \nabla w^n, \nabla h\rangle - \sum_{k=1}^{d} w^n_{x_k}\mathcal{L}_{x_k} w^n\Big]$$
$$- \lambda\big(ru + \psi_\varepsilon(\zeta_n) - \tfrac{1}{\delta}(g-u)^+ - h\big) \tag{3.81}$$
$$\geq -C_1\xi|\nabla u|_d^2 - \tfrac{\theta}{2}\xi|D^2 w^n|^2_{d\times d} - \xi C_2 - \lambda r M_1 - \lambda\psi'_\varepsilon(\bar{\zeta}_n)\mathcal{H}(\nabla u) - R_n,$$

where $R_n \to 0$ as $n \to \infty$ uniformly on $\overline{\mathcal{O}}_{m_0+1}$. We notice that, differently to (3.51), we have a factor $1/2$ multiplying $|D^2 w^n|^2_{d\times d}$. That of course is obtained by adjusting the constant $C_1 = (1 + 2d^4 A^2_{m_0+1}\theta^{-1} + 2dA_{m_0+1})$ (see (3.63)) with A_{m_0+1} as in (3.61) and $C_2 = \|\nabla h\|_0$.

The last term on the right-hand side of (3.80) can be easily bounded. Set

$\bar{a}_0 := \max_{i,j} \|a_{ij}\|_0$, which is finite by continuity of a_{ij}, and recall that $|\nabla\xi|_d^2 \le C_0\xi$ by (3.16). Then

$$
\begin{aligned}
2\sum_{i,j=1}^{d} a_{ij}\xi_{x_i}\langle \nabla w^n, \nabla w^n_{x_j}\rangle &\ge -2\bar{a}_0 d \sum_{j=1}^{d} |\nabla\xi|_d |\nabla w^n|_d |\nabla w^n_{x_j}|_d \\
&\ge -2\bar{a}_0 d^2 \sqrt{C_0\xi}|\nabla w^n|_d |D^2 w^n|_{d\times d} \\
&\ge -C_3|\nabla w^n|_d^2 - \xi\frac{\theta}{2}|D^2 w^n|^2_{d\times d},
\end{aligned}
\tag{3.82}
$$

where the final inequality is by $|ab| \le pa^2 + \frac{b^2}{p}$ with $a = 2\bar{a}_0 d^2\sqrt{C_0}|\nabla w^n|_d$, $b = \sqrt{\xi}|D^2 w^n|_{d\times d}$ and $p = 2/\theta$, and setting $C_3 = 8\theta^{-1}\bar{a}_0^2 d^4 C_0$. It is also easy to check that $\|\mathcal{L}\xi\|_0 \le \kappa$ for some $\kappa = \kappa(\|b\|_0, \|\sigma\|_0) > 0$, because the derivatives of ξ are bounded independently of m_0.

Plugging (3.81) and (3.82) into (3.80) and setting $C_4 = C_3 + \kappa$ we have

$$
\begin{aligned}
0 \ge\, & 2\xi\psi'_\varepsilon(\bar{\zeta}_n)\langle \nabla w^n, \nabla\big(\mathcal{H}(\nabla u^n) - f^2\big)\rangle - \xi C_1|\nabla u|_d^2 - C_4|\nabla w^n|_d^2 \\
& - \lambda\psi'_\varepsilon(\bar{\zeta}_n)\mathcal{H}(\nabla u) - \xi C_2 - \lambda r M_1 - R_n.
\end{aligned}
\tag{3.83}
$$

Using (3.57) we have $\xi C_1|\nabla u|_d^2 + C_4|\nabla w^n|_d^2 \le C_5|\nabla u|_d^2 + \tilde{R}_n$, where $\tilde{R}_n \to 0$ as $n \to \infty$, uniformly on $\overline{\mathcal{O}}_{m_0+1}$ and $C_5 = C_1 + C_4$. It remains to find a lower bound for $\xi\langle\nabla w^n, \nabla\big(\mathcal{H}(\nabla u^n) - f^2\big)\rangle$. By the same arguments as in (3.66), the bound in (3.67) continue to hold, up to the inclusion of the multiplicative factor ξ. We have an additional term in the final expression in (3.69), because now $v^{\lambda,n}_{x_k} = 0$ gives $2\xi\langle\nabla w^n, \nabla w^n_{x_k}\rangle = \lambda u_{x_k} - \xi_{x_k}|\nabla w^n|_d^2$. So, the extra term appearing in (3.69) reads $-u_{x_k}\xi_{x_k}|\nabla w^n|_d^2$ and, similarly to (3.57), we get

$$
-u_{x_k}\xi_{x_k}|\nabla w^n|_d^2 \ge -u_{x_k}\xi_{x_k}|\nabla u|_d^2 - |u_{x_k}||\xi_{x_k}|\|\nabla\hat{w}^n\|_0\big(2\|\nabla u\|_0 + \|\nabla\hat{w}^n\|_0\big).
$$

In summary, we have

$$
\begin{aligned}
\xi\langle\nabla w^n, \nabla\big(\mathcal{H}(\nabla u^n) - f^2\big)\rangle \ge\, & \lambda\mathcal{H}(\nabla u) - \xi|\nabla u|_d|\nabla f^2|_d \\
& - |\nabla u|_d^3|\nabla\xi|_d - \xi\hat{R}_n,
\end{aligned}
\tag{3.84}
$$

where $\hat{R}_n \to 0$ as $n \to \infty$, uniformly on $\overline{\mathcal{O}}_{m_0+1}$.

Substituting (3.84) into (3.83) and grouping together all terms that vanish as

$n \to \infty$ we obtain

$$0 \geq 2\psi'_\varepsilon(\bar{\zeta}_n)\big[\tfrac{\lambda}{2}\mathcal{H}(\nabla u) - \xi|\nabla u|_d|\nabla f^2|_d - |\nabla u|_d^3|\nabla\xi|_d\big] \\ - C_5|\nabla u|_d^2 - C_2 - \lambda r M_1 - \bar{R}_n,$$

where $\bar{R}_n \to 0$ as $n \to \infty$. Using that $\psi'_\varepsilon(\bar{\zeta}_n) \geq 1$ by the analogue of (3.43), for sufficiently large n, and multiplying the above expression by -1 we arrive at

$$0 \leq \psi'_\varepsilon(\bar{\zeta}_n)\big[-\lambda\mathcal{H}(\nabla u) + 2\xi|\nabla u|_d|\nabla f^2|_d + 2|\nabla u|_d^3|\nabla\xi|_d \\ + C_5|\nabla u|_d^2 + C_2 + \lambda r M_1 + \bar{R}_n\big].$$

Using $\mathcal{H}(\nabla u) = |\nabla u|_d^2$, $2\xi|\nabla u|_d|\nabla f^2|_d \leq |\nabla u|_d^2 + |\nabla f^2|_d^2$ and $|\nabla f^2|_d \leq \|\nabla f^2\|_0$, the above inequality leads to

$$\big(\lambda - 1 - 2|\nabla\xi|_d|\nabla u|_d - C_5\big)|\nabla u|_d^2 \leq \|\nabla f^2\|_0^2 + C_2 + \lambda r M_1 + \bar{R}_n.$$

Then, recalling that $|\nabla\xi|_d^2 \leq C_0\xi$ (see (3.16)) and setting

$$\bar{\lambda} = 2 + 2\sqrt{C_0}\big\|\sqrt{\xi}|\nabla u|_d\big\|_0 + C_5,$$

we obtain

$$|\nabla u|_d^2(t_n^{\bar{\lambda}}, x_n^{\bar{\lambda}}) \leq \|\nabla f\|_0^2 + C_2 + \bar{\lambda} r M_1 + \bar{R}_n. \tag{3.85}$$

The parameter $\bar{\lambda}$ is bounded from above independently of ∇u as follows. Let $c > 0$ be a constant that varies from one expression to the next, independent of $\bar{\lambda}$, ε, δ, but depending on d and the $C^0(\overline{\mathcal{O}}_{m_0+1})$-norms of b, σ, g, h, f^2, and their spatial gradient. From (3.76), (3.77) and (3.78), we obtain

$$\begin{aligned}\big\|\xi|\nabla u|_d^2\big\|_0 &\leq |\nabla u(t_n^{\bar{\lambda}}, x_n^{\bar{\lambda}})|_d^2 + \bar{\lambda}M_1 + \eta \\ &\leq c(1 + \bar{\lambda} + \bar{R}_n) + \eta \\ &\leq c(1 + \big\|\sqrt{\xi}|\nabla u|_d\big\|_0 + \bar{R}_n) + \eta,\end{aligned}$$

where the second inequality uses (3.85) and the final one the definition of $\bar{\lambda}$. Since $\|\xi|\nabla u|_d^2\|_0 = \|\sqrt{\xi}|\nabla u|_d\|_0^2$, then $\|\sqrt{\xi}|\nabla u|_d\|_0 \leq \max\{1, c(1 + \bar{R}_n)\} + \eta$. As $n \uparrow \infty$ and $\eta \downarrow 0$ we get (3.75) from (3.77) and (3.85), choosing $\lambda_0 = 2 + 2(1+c)\sqrt{C_0} + C_5$, thanks to the bound on $\bar{\lambda}$. □

The bound in Proposition 3.18 depends on m. The next result instead provides a uniform bound on any compact $\overline{\mathcal{O}}_{m_0}$ for $m_0 < m$. This can be achieved thanks to (3.75).

Proposition 3.23 *Fix $m_0 \in \mathbb{N}$ and $q \geq m_0 + 3$ and let $u_q^{\varepsilon,\delta}$ be the unique solution of Problem B on $\mathcal{O}_q$. For any $p \in (d+2, \infty)$ and $\beta = 1 - (d+2)/p$ there is $N_2 = N_2(m_0, \varepsilon, \delta, p)$ such that*

$$\|u_q^{\varepsilon,\delta}\|_{W^{1,2,p}(\mathcal{O}_{m_0})} + \|u_q^{\varepsilon,\delta}\|_{C^{0,1,\beta}(\overline{\mathcal{O}}_{m_0})} \leq N_2. \tag{3.86}$$

Proof. Define $\varphi(t,x) := \xi_{m_0}(x) u_q^{\varepsilon,\delta}(t,x)$. Since $u_q^{\varepsilon,\delta}$ solves (3.18) and $f_q = f$, $g_q = g$ and $h_q = h$ on $\mathcal{O}_{m_0+1}$, then φ solves

$$\partial_t \varphi + \mathcal{L}\varphi - r\varphi = \xi_{m_0}\big[-h - \tfrac{1}{\delta}(g - u_q^{\varepsilon,\delta})^+ + \psi_\varepsilon\big(\mathcal{H}(\nabla u_q^{\varepsilon,\delta}) - f^2\big)\big] + Q, \quad \text{on } \mathcal{O}_{m_0+1},$$

where $Q(t,x) = u_q^{\varepsilon,\delta}(t,x)(\mathcal{L}\xi_{m_0})(x) + 2\langle a(x)\nabla\xi_{m_0}(x), \nabla u_q^{\varepsilon,\delta}(t,x)\rangle$, and with boundary conditions $\varphi(t,x) = 0$ for $x \in \partial B_{m_0+1}$ and $\varphi(T,x) = \xi_{m_0}(x) g(T,x)$ for $x \in B_{m_0+1}$. As in (3.72) we have

$$\begin{aligned} \|\varphi\|_{W^{1,2,p}(\mathcal{O}_{m_0+1})} \leq C\Big(&\big\|\xi_{m_0}\big[h + \tfrac{1}{\delta}(g - u_q^{\varepsilon,\delta})^+ - \psi_\varepsilon(\mathcal{H}(\nabla u_q^{\varepsilon,\delta}) - f^2)\big] + Q\big\|_{L^p(\mathcal{O}_{m_0+1})} \\ &+ \|\xi_{m_0} g\|_{W^{1,2,p}(\mathcal{O}_{m_0+1})}\Big) \end{aligned} \tag{3.87}$$

for any $p \in (1,\infty)$ and with $C > 0$ independent of ε, δ and q. Denoting $|\mathcal{O}_{m_0+1}|$ the volume of $\mathcal{O}_{m_0+1}$ and using Proposition 3.22 we obtain

$$\begin{aligned} \|u_q^{\varepsilon,\delta}\|_{W^{1,2,p}(\mathcal{O}_{m_0})} &\leq \|\varphi\|_{W^{1,2,p}(\mathcal{O}_{m_0+1})} \\ &\leq C|\mathcal{O}_{m_0+1}|^{\frac{1}{p}}\big(\|h + \tfrac{1}{\delta}g + Q\|_{m_0+1} + \tfrac{1}{\varepsilon}(N_1)^2\big) \\ &\quad + C\|\xi_{m_0} g\|_{W^{1,2,p}(\mathcal{O}_{m_0+1})}, \end{aligned}$$

where the first inequality is due to $u_q^{\varepsilon,\delta} = \varphi$ on $\mathcal{O}_{m_0}$. Since Q is bounded on $\overline{\mathcal{O}}_{m_0+1}$ independently of q, the estimate above and Sobolev embedding (1.6) give us (3.86). □

3.3.2 Penalised problem on unbounded domain

Combining the results obtained so far we can prove existence and uniqueness of the solution to a penalised problem on $\mathbb{R}^{d+1}_{0,T}$.

Problem C. Find $u = u^{\varepsilon,\delta}$ with $u^{\varepsilon,\delta} \in (C^{1,2,\alpha}_{Loc} \cap W^{1,2,p}_{\ell oc})(\mathbb{R}^{d+1}_{0,T})$, for any $p \in (1,\infty)$ and $\alpha \in (0,1)$ as in Assumption 3.5, that solves:

$$\begin{cases} (\partial_t + \mathcal{L} - r)u = -h - \frac{1}{\delta}\big(g-u\big)^+ + \psi_\varepsilon\big(\mathcal{H}(\nabla u) - f^2\big), & \text{on } [0,T)\times\mathbb{R}^d, \\ u(T,x) = g(T,x), & \text{for all } x \in \mathbb{R}^d. \end{cases} \tag{3.88}$$

■

Theorem 3.24 *There exists a solution $u^{\varepsilon,\delta}$ of Problem C.*

Proof. Fix n and take $m > n+3$. From Proposition 3.23 we know that for any $\beta \in (0,1)$ and $p \in (1,\infty)$, the norms $\|u^{\varepsilon,\delta}_m\|_{W^{1,2,p}(\mathcal{O}_n)}$ and $\|u^{\varepsilon,\delta}_m\|_{C^{0,1,\beta}(\overline{\mathcal{O}}_n)}$ are bounded by a constant independent of m. By weak compactness in $W^{1,2,p}$ and Ascoli-Arzelá's theorem we can then extract a sequence $(u^{\varepsilon,\delta}_{m^n_k})_{k\in\mathbb{N}}$ and there exists a function $u^{\varepsilon,\delta;n} \in W^{1,2,p}(\mathcal{O}_n)$ (both possibly depending on the choice of $\mathcal{O}_n$) such that, as $k \to \infty$ (and $m^n_k \to \infty$), we obtain

$$\begin{aligned} &u^{\varepsilon,\delta}_{m^n_k} \to u^{\varepsilon,\delta;n} \quad \text{and} \quad \nabla u^{\varepsilon,\delta}_{m^n_k} \to \nabla u^{\varepsilon,\delta;n} \quad \text{in } C^\alpha(\overline{\mathcal{O}}_n), \\ &\partial_t u^{\varepsilon,\delta}_{m^n_k} \to \partial_t u^{\varepsilon,\delta;n} \quad \text{and} \quad D^2 u^{\varepsilon,\delta}_{m^n_k} \to D^2 u^{\varepsilon,\delta;n} \quad \text{weakly in } L^p(\mathcal{O}_n). \end{aligned} \tag{3.89}$$

Since the sequence $(u^{\varepsilon,\delta}_{m^n_k})_{k\in\mathbb{N}}$ is bounded also in the $W^{1,2,p}(\mathcal{O}_{n+1})$-norm (perhaps by a larger constant), then up to selecting a further subsequence we have convergence as in (3.89) but with n replaced by $n+1$. Therefore $u^{\varepsilon,\delta;n} = u^{\varepsilon,\delta;n+1}$ on $\mathcal{O}_n$. Using that $\overline{\mathcal{O}}_n \uparrow \mathbb{R}^{d+1}_{0,T}$ as $n \to \infty$ and iterating the extraction of further subsequences (if needed), we can uniquely define a limit function $u^{\varepsilon,\delta} \in (C^{0,1,\alpha}_{\ell oc} \cap W^{1,2,p}_{\ell oc})(\mathbb{R}^{d+1}_{0,T})$.

Fix n and take $m > n$. Multiply the PDE solved by $u^{\varepsilon,\delta}_m$ (see (3.18)) by a test function supported on $\overline{\mathcal{O}}_n$. Then passing to the limit along the subsequence constructed above, it is standard procedure to show that $u^{\varepsilon,\delta}$ satisfies the first equation in (3.88) in the a.e. sense on $\mathcal{O}_n$ thanks to locally uniform convergence on compacts of $u^{\varepsilon,\delta}_m$ and $\nabla u^{\varepsilon,\delta}_m$ and the weak convergence of $\partial_t u^{\varepsilon,\delta}_m$ and $D^2 u^{\varepsilon,\delta}_m$. Since, this can be done for any $\mathcal{O}_n$ and $u^{\varepsilon,\delta}(T,\cdot) = g(T,\cdot)$, then $u^{\varepsilon,\delta}$ solves (3.88) in the a.e. sense (it is a *strong* solution). It now remains to prove it is actually a classical solution.

Fix an arbitrary open bounded domain $\mathcal{O} \subset \mathbb{R}^{d+1}_{0,T}$ with smooth parabolic boundary $\partial_P \mathcal{O}$. Let $v \in C^{1,2,\alpha}_{Loc}(\mathcal{O})$ be the unique classical solution of the boundary value problem

$$\begin{cases} \partial_t v + \mathcal{L}v - rv = -h - \frac{1}{\delta}\big(g - u^{\varepsilon,\delta}\big)^+ + \psi_\varepsilon\big(\mathcal{H}(\nabla u^{\varepsilon,\delta}) - f^2\big), & \text{on } \mathcal{O}, \\ v(t,x) = u^{\varepsilon,\delta}(t,x), & \text{for } (t,x) \in \partial_P \mathcal{O}. \end{cases} \tag{3.90}$$

Existence and uniqueness of such v is guaranteed by [29, Thm. 3.4.9] because

$$-h - \tfrac{1}{\delta}\big(g - u^{\varepsilon,\delta}\big)^+ + \psi_\varepsilon\big(\mathcal{H}(\nabla u^{\varepsilon,\delta}) - f^2\big) \in C^\alpha(\overline{\mathcal{O}}),$$

and $\mathcal{L}$ is uniformly elliptic on $\overline{\mathcal{O}}$ with continuously differentiable coefficients (Assumption 3.4). Since v is also a strong solution, then $v - u^{\varepsilon,\delta} \in W^{1,2,p}(\mathcal{O})$ is a strong solution of $\partial_t w + \mathcal{L}w - rw = 0$ in $\mathcal{O}$ with $w = 0$ on $\partial_P \mathcal{O}$. It follows that $\|v - u^{\varepsilon,\delta}\|_{W^{1,2,p}(\mathcal{O})} = 0$ by the same estimate as in (3.72). By arbitrariness of $\mathcal{O}$ we can choose a $C^{1,2,\alpha}_{Loc}$-representative of $u^{\varepsilon,\delta}$, as claimed. □

We now give a probabilistic representation for $u^{\varepsilon,\delta}$ analogue of (3.23) but on unbounded domain. For $(n,\nu) \in \mathcal{A}^\circ_t$ and $w \in \mathcal{T}^\delta_t$ let us denote by $\mathcal{J}^{\varepsilon,\delta}_{t,x}(n,\nu,w)$ a payoff analogue of (3.22) but with ρ_m, g_m, h_m replaced by $T-t$, g, h, respectively, and with the Hamiltonian H^ε_m replaced by

$$H^\varepsilon(t,x,y) := \sup_{p \in \mathbb{R}^d} \big\{ \langle y,p \rangle - \psi_\varepsilon\big(\mathcal{H}(p) - f^2(t,x)\big) \big\}. \tag{3.91}$$

Notice that

$$t \mapsto H^\varepsilon(t,x,y) \text{ is non-increasing for all } (x,y) \in \mathbb{R}^d \times \mathbb{R}^d, \tag{3.92}$$

because $t \mapsto f(t,x)$ is non-increasing by Assumption 3.5. Moreover, taking

$p = \varepsilon y/2$ in H^ε gives

$$\begin{aligned} H^\varepsilon(t,x,y) &\geq \frac{\varepsilon}{2}|y|_d^2 - \psi_\varepsilon\big(\mathcal{H}(\tfrac{\varepsilon}{2}y) - f^2(t,x)\big) \\ &\geq \frac{\varepsilon}{2}|y|_d^2 - \psi_\varepsilon\big(\mathcal{H}(\tfrac{\varepsilon}{2}y)\big) \\ &= \frac{\varepsilon}{2}|y|_d^2 - \psi_\varepsilon\big(\tfrac{\varepsilon^2}{4}|y|_d^2\big) \\ &\geq \frac{\varepsilon}{4}|y|_d^2. \end{aligned} \tag{3.93}$$

Proposition 3.25 *Let $u^{\varepsilon,\delta}$ be a solution of Problem C. Then*

$$u^{\varepsilon,\delta}(t,x) = \inf_{(n,\nu)\in\mathcal{A}_t^\circ} \sup_{w\in\mathcal{T}_t^\delta} \mathcal{J}_{t,x}^{\varepsilon,\delta}(n,\nu,w) = \sup_{w\in\mathcal{T}_t^\delta} \inf_{(n,\nu)\in\mathcal{A}_t^\circ} \mathcal{J}_{t,x}^{\varepsilon,\delta}(n,\nu,w). \tag{3.94}$$

Proof. Fix $[(n,\nu),w] \in \mathcal{A}_t^\circ \times \mathcal{T}_t^\delta$. Since $u^{\varepsilon,\delta} \in (W^{1,2,p}_{\ell oc} \cap C^{1,2,\beta}_{Loc})(\mathbb{R}^{d+1}_{0,T})$, its time derivative and its spatial second order derivatives could explode at time T, we define $\rho_m^k \coloneqq \rho_m \wedge (T-t-k^{-1})$ and ρ_m as in (3.21). Since the stochastic integral is a martingale because the spatial derivatives are locally bounded, by an application of Dynkin's formula to $R^w_{\rho_m^k} u^{\varepsilon,\delta}(t+\rho_m^k, X^{[n,\nu]}_{\rho_m^k})$ combined with (3.88) gives

$$\begin{aligned} u^{\varepsilon,\delta}(t,x) = \mathsf{E}_x\Bigg[& R^w_{\rho_m^k} u^{\varepsilon,\delta}\big(t+\rho_m^k, X^{[n,\nu]}_{\rho_m^k}\big) \\ & + \int_0^{\rho_m^k} R^w_s \Big[h + \tfrac{1}{\delta}\big(g-u^{\varepsilon,\delta}\big)^+ - \psi_\varepsilon\big(\mathcal{H}(\nabla u^{\varepsilon,\delta}) - f^2\big)\Big]\big(t+s, X^{[n,\nu]}_s\big)\, \mathrm{d}s \\ & + \int_0^{\rho_m^k} R^w_s \Big[w_s u^{\varepsilon,\delta} - \big\langle n_s \dot\nu_s, \nabla u^{\varepsilon,\delta}\big\rangle\Big]\big(t+s, X^{[n,\nu]}_s\big)\, \mathrm{d}s\Bigg]. \end{aligned}$$

Since the process is localised inside the ball B_m, all the functions involved in the expectations are bounded, thus we can send $k \uparrow \infty$ and passing the limit under expectation we get

$$\begin{aligned} u^{\varepsilon,\delta}(t,x) = \mathsf{E}_x\Bigg[& R^w_{\rho_m} u^{\varepsilon,\delta}\big(t+\rho_m, X^{[n,\nu]}_{\rho_m}\big) \\ & + \int_0^{\rho_m} R^w_s \Big[h + \tfrac{1}{\delta}\big(g-u^{\varepsilon,\delta}\big)^+ - \psi_\varepsilon\big(\mathcal{H}(\nabla u^{\varepsilon,\delta}) - f^2\big)\Big]\big(t+s, X^{[n,\nu]}_s\big)\, \mathrm{d}s \\ & + \int_0^{\rho_m} R^w_s \Big[w_s u^{\varepsilon,\delta} - \big\langle n_s \dot\nu_s, \nabla u^{\varepsilon,\delta}\big\rangle\Big]\big(t+s, X^{[n,\nu]}_s\big)\, \mathrm{d}s\Bigg]. \end{aligned} \tag{3.95}$$

By definition of the Hamiltonian H^ε in (3.91) we have

$$\begin{aligned} u^{\varepsilon,\delta}(t,x) &\le \mathsf{E}_x\Big[R^w_{\rho_m} u^{\varepsilon,\delta}\big(t+\rho_m, X^{[n,\nu]}_{\rho_m}\big)\Big] \qquad (3.96)\\ &\quad + \mathsf{E}_x\Big[\int_0^{\rho_m} R^w_s\Big[h+\tfrac{1}{\delta}\big(g-u^{\varepsilon,\delta}\big)^+ + w_s u^{\varepsilon,\delta} + H^\varepsilon(\cdot, n_s\dot\nu_s)\Big]\big(t+s, X^{[n,\nu]}_s\big)\,\mathrm{d}s\Big]. \end{aligned}$$

Letting $m\uparrow\infty$ we have $\rho_m\uparrow T-t$, P_x-a.s. We can take the limit inside the second expectation by monotone convergence as all the terms under the integral are positive. By Lemma 3.13 the term under the first expectation has quadratic growth in $X^{[n,\nu]}$. Thanks to standard estimates for SDEs (see [48, Thm. 2.5.10]) there is a constant $c>0$ independent of m such that

$$\mathsf{E}_x\Big[\sup_{0\le s\le T-t}\big|X^{[n,\nu]}_s\big|^2_d\Big] \le c\big(1+|x|^2_d+\mathsf{E}_x[|\nu_{T-t}|^2]\big).$$

Then, dominated convergence and $u^{\varepsilon,\delta}(T,\,\cdot\,)=g(T,\,\cdot\,)$ give us

$$\begin{aligned} u^{\varepsilon,\delta}(t,x) &\le \mathsf{E}_x\Big[R^w_{T-t} g\big(T, X^{[n,\nu]}_{T-t}\big) \qquad (3.97)\\ &\quad + \int_0^{T-t} R^w_s\Big[h+\tfrac{1}{\delta}\big(g-u^{\varepsilon,\delta}\big)^+ + w_s u^{\varepsilon,\delta} + H^\varepsilon(\cdot, n_s\dot\nu_s)\Big]\big(t+s, X^{[n,\nu]}_s\big)\,\mathrm{d}s\Big]. \end{aligned}$$

By arguments as in the proof of Proposition 3.9, with $w^*\in\mathcal{T}^\delta_t$ defined as in (3.26) but with u and g_m replaced by $u^{\varepsilon,\delta}$ and g, respectively, we obtain $u^{\varepsilon,\delta}(t,x)\le\mathcal{J}^{\varepsilon,\delta}_{t,x}(n,\nu,w^*)$. Therefore

$$u^{\varepsilon,\delta}(t,x) \le \sup_{w\in\mathcal{T}^\delta_t}\ \inf_{(n,\nu)\in\mathcal{A}^\circ_t} \mathcal{J}^{\varepsilon,\delta}_{t,x}(n,\nu,w). \qquad (3.98)$$

As in Proposition 3.9, for the reverse inequality we set $X^*=X^{[n^*,\nu^*]}$ and denote

$$\begin{aligned} n^*_s &:= \begin{cases} -\dfrac{\nabla u^{\varepsilon,\delta}(t+s,X^*_s)}{|\nabla u^{\varepsilon,\delta}(t+s,X^*_s)|_d}, & \text{if } \nabla u^{\varepsilon,\delta}(t+s,X^*_s)\neq \mathbf{0},\\ \text{any unit vector}, & \text{if } \nabla u^{\varepsilon,\delta}(t+s,X^*_s)= \mathbf{0}, \end{cases} \qquad (3.99)\\ \dot\nu^*_s &:= 2\psi'_\varepsilon\big(\mathcal{H}(\nabla u^{\varepsilon,\delta}(t+s,X^*_s)) - f^2(t+s,X^*_s)\big)|\nabla u^{\varepsilon,\delta}(t+s,X^*_s)|_d. \end{aligned}$$

We claim here and will prove later that $(n^*, \nu^*) \in \mathcal{A}_t^\circ$ and the SDE for X^* admits a unique non-exploding strong solution. For (n^*, ν^*) equality holds in (3.96). As $m \uparrow \infty$ Fatou's lemma gives $u^{\varepsilon,\delta}(t,x) \geq \mathcal{J}_{t,x}(n^*, \nu^*, w)$, hence

$$u^{\varepsilon,\delta}(t,x) \geq \inf_{(n,\nu)\in\mathcal{A}_t^\circ} \sup_{w\in\mathcal{T}_t^\delta} \mathcal{J}_{t,x}(n,\nu,w). \tag{3.100}$$

Combining (3.98) and (3.100) we conclude.

It remains to check that $(n^*, \nu^*) \in \mathcal{A}_t^\circ$ and X^* is non-exploding. We use an argument from [61, Lemma 13.7]. Let $\zeta_m = \inf\{s \geq 0 : |X_s^*|_d \geq m\}$. On the random time-interval $[0, \zeta_m \wedge (T-t)]$ the process X^* is well-defined and the pair (n^*, ν^*) is adapted because $u^{\varepsilon,\delta} \in C_{Loc}^{1,2,\alpha}(\mathcal{O}_m) \cap C^{0,1,\alpha}(\overline{\mathcal{O}}_m)$. Notice that $\zeta_k \leq \zeta_{k+1}$ and it may occur $\zeta_\infty := \lim_{k\to\infty} \zeta_k < T - t$ with positive probability. Moreover $\rho_m = \zeta_m \wedge (T-t)$ in (3.95) and let us take $w \equiv 0$ therein. By construction, for $s \in [0, \zeta_m \wedge (T-t)]$

$$-\big\langle n_s^* \dot{\nu}_s^*, \nabla u^{\varepsilon,\delta}(t+s, X_s^*)\big\rangle - \psi_\varepsilon\big(|\nabla u^{\varepsilon,\delta}|_d^2 - f^2\big)(t+s, X_s^*) = H^\varepsilon(t+s, X_s^*, n_s^*\dot{\nu}_s^*).$$

Then, by positivity of all remaining terms in (3.95)

$$u^{\varepsilon,\delta}(t,x) \geq \mathsf{E}_x\Big[\int_0^{\zeta_m\wedge(T-t)} e^{-rs} H^\varepsilon(t+s, X_s^*, n_s^*\dot{\nu}_s^*)\mathrm{d}s\Big].$$

By positivity of H^ε and monotone convergence, we can let $m \uparrow \infty$ and preserve the inequality while the integral in time extends to $\zeta_\infty \wedge (T-t)$. Combining with (3.93) and Lemma 3.13 we have

$$\frac{\varepsilon}{4}\mathsf{E}_x\Big[\int_0^{\zeta_\infty\wedge(T-t)} e^{-rs}|\dot{\nu}_s^*|^2\,\mathrm{d}s\Big] \leq u^{\varepsilon,\delta}(t,x) \leq K_3(1+|x|_d^2). \tag{3.101}$$

Since $\dot{\nu}^* \geq 0$ then $s \mapsto \nu_s^*$ is non-decreasing and

$$|\nu^*_{\zeta_m\wedge(T-t)}|^2 = 2\int_0^{\zeta_m\wedge(T-t)} \nu_s^*\dot{\nu}_s^*\,\mathrm{d}s \leq \int_0^{\zeta_m\wedge(T-t)} |\nu_s^*|^2\,\mathrm{d}s + \int_0^{\zeta_m\wedge(T-t)} |\dot{\nu}_s^*|^2\,\mathrm{d}s,$$

where we used $2ab \leq a^2 + b^2$. By Gronwall's lemma and taking expectations we obtain

$$\mathsf{E}_x\Big[\big|\nu^*_{\zeta_m\wedge(T-t)}\big|^2\Big] \le e^T\mathsf{E}_x\Big[\int_0^{\zeta_m\wedge(T-t)}|\dot\nu^*_s|^2\,\mathrm{d}s\Big] \le e^{T(1+r)}\mathsf{E}_x\Big[\int_0^{\zeta_m\wedge(T-t)}e^{-rs}|\dot\nu^*_s|^2\,\mathrm{d}s\Big].$$

Combining with (3.101) and letting $m\to\infty$, Fatou's lemma gives us

$$\mathsf{E}_x\Big[\big|\nu^*_{\zeta_\infty\wedge(T-t)}\big|^2\Big] \le 4e^{T(1+r)}\varepsilon^{-1}K_3(1+|x|_d^2). \tag{3.102}$$

Linear growth of (b,σ) and well-posedness of $X^*_{s\wedge\zeta_m}$ give, by Markov inequality and standard bounds,

$$\begin{aligned}\mathsf{P}_x(\zeta_m<T-t) &\le \frac{1}{m^2}\mathsf{E}_x\Big[\sup_{s\in[0,\zeta_m\wedge(T-t)]}|X^*_s|_d^2\Big]\\ &\le \frac{C}{m^2}\Big(1+|x|_d^2+\mathsf{E}_x\Big[\big|\nu^*_{\zeta_m\wedge(T-t)}\big|^2\Big]\Big)\\ &\le \frac{C}{m^2}c(\varepsilon)(1+|x|_d^2),\end{aligned}$$

where $C>0$ depends only on T and D_1 from Assumption 3.4, and $c(\varepsilon)$ depends on the constants from (3.102). Since $\zeta_m\uparrow\zeta_\infty$, then $\mathsf{P}_x(\zeta_m<T-t)\downarrow\mathsf{P}_x(\zeta_\infty\le T-t)$ as $m\to\infty$ and by taking limits in the expression above we conclude $\mathsf{P}_x(\zeta_\infty\le T-t)=0$. Thus, X^*_s is well-defined for all $s\in[0,T-t]$ and $\mathsf{E}_x[|\nu^*_{T-t}|^2]<\infty$, by (3.102) implying $(n^*,\nu^*)\in\mathcal{A}^\circ_t$ as claimed. □

Proposition 3.25 implies that Problem C admits a *unique* solution and that the treble $[(n^*,\nu^*),w^*]$ is optimal in (3.94). By arguments as in the proof of Proposition 3.11 we also obtain the next result.

Proposition 3.26 *Let $u^{\varepsilon,\delta}$ be the unique solution of Problem C. Then*

$$\begin{aligned}u^{\varepsilon,\delta}(t,x) = \inf_{(n,\nu)\in\mathcal{A}^\circ_t}\mathsf{E}_x\Bigg[&R^{\delta^{-1}}_{T-t}\,g(T,X^{[n,\nu]}_{T-t})\\ &+\int_0^{T-t}R^{\delta^{-1}}_s\big[h+\tfrac{1}{\delta}g\vee u^{\varepsilon,\delta}+H^\varepsilon(\cdot,n_s\dot\nu_s)\big](t+s,X^{[n,\nu]}_s)\,\mathrm{d}s\Bigg],\end{aligned}$$

and the pair (n^*,ν^*) from (3.99) is optimal.

3.3.3 Refined estimates independent of ε and δ

Here we develop bounds for the penalty terms in the PDE of Problem C which are independent of ε and δ.

Lemma 3.27 *For K_2 as in (3.14) we have*

$$\tfrac{1}{\delta}\big\|\big(g-u^{\varepsilon,\delta}\big)^+\big\|_\infty \le K_2. \tag{3.103}$$

Proof. For any $(n,\nu)\in\mathcal{A}_t^\circ$, the function g has the same probabilistic representation as in (3.37) but with g_m, ρ_m and w replaced by g, $T-t$ and $1/\delta$, respectively. Combining that with the expression for $u^{\varepsilon,\delta}$ in Proposition 3.26, and recalling Θ defined in (3.14), we get

$$\begin{aligned}(u^{\varepsilon,\delta}-g)(t,x) = \inf_{(n,\nu)\in\mathcal{A}_t^\circ}\mathsf{E}_x\bigg[&\int_0^{T-t} R_s^{\delta^{-1}}\Big[\langle n_s\dot\nu_s,\nabla g\rangle + H^\varepsilon(\cdot,n_s\dot\nu_s)\Big](t+s,X_s^{[n,\nu]})\,\mathrm{d}s\\ &+\int_0^{T-t} R_s^{\delta^{-1}}\Big[\Theta+\tfrac{1}{\delta}g\vee u^{\varepsilon,\delta}-\tfrac{1}{\delta}g\Big](t+s,X_s^{[n,\nu]})\,\mathrm{d}s\bigg].\end{aligned}$$

As in (3.20), $[\langle n_s\dot\nu_s,\nabla g\rangle + H^\varepsilon(\cdot,n_s\dot\nu_s)](t+s,X_s^{[n,\nu]})\ge 0$, and observing that $g\vee u^{\varepsilon,\delta}-g\ge 0$ we get

$$\begin{aligned}u^{\varepsilon,\delta}(t,x)-g(t,x) &\ge \inf_{(n,\nu)\in\mathcal{A}_t^\circ}\mathsf{E}_x\bigg[\int_0^{T-t} R_s^{\delta^{-1}}\Theta\big(t+s,X_s^{[n,\nu]}\big)\,\mathrm{d}s\bigg]\\ &\ge -K_2\frac{\delta}{r\delta+1},\end{aligned}$$

where K_2 was defined in (3.14). The above implies $\frac{1}{\delta}(g(t,x)-u^{\varepsilon,\delta}(t,x))^+\le K_2$ as needed. □

Next we give an upper bound on $\partial_t u^{\varepsilon,\delta}$. In the lemma below we understand

$$\partial_t u^{\varepsilon,\delta}(T,x) := \lim_{s\to 0}\frac{u^{\varepsilon,\delta}(T,x)-u^{\varepsilon,\delta}(T-s,x)}{s}.$$

Lemma 3.28 *There is $K_4>0$ only depending on K_0 and K_2 from Assumption 3.5 such that*

$$\partial_t u^{\varepsilon,\delta}(t,x)\le K_4,\qquad \textit{for } (t,x)\in\mathbb{R}^{d+1}_{0,T}. \tag{3.104}$$

Proof. Let $u = u^{\varepsilon,\delta}$ for simplicity and take $T \geq t_2 > t_1 \geq 0$. Let $w^{(2)} \in \mathcal{T}^\delta_{t_2}$ be optimal for the value function $u(t_2, x)$ and let $(n^{(1)}, \nu^{(1)}) \in \mathcal{A}^\circ_{t_1}$ be optimal for the value function $u(t_1, x)$. Set $X^{(1)} = X^{[n^{(1)},\nu^{(1)}]}$ and notice that $w^{(1)}_s := w^{(2)}_s \mathbb{1}_{\{s\leq T-t_2\}}$ lies in $\mathcal{T}^\delta_{t_1}$ and $(n^{(1)}, \nu^{(1)})$ restricted to $[0, T-t_2]$ lies in $\mathcal{A}^\circ_{t_2}$. To simplify notation let us also set $[\Delta_{t_2,t_1} g](s, x) = g(t_2 + s, x) - g(t_1 + s, x)$ and analogously for $[\Delta_{t_2,t_1} h](s, x)$ and $[\Delta_{t_2,t_1} H^\varepsilon](s, x, y)$. Then

$$
\begin{aligned}
&u(t_2, x) - u(t_1, x)\\
&\leq \mathcal{J}^{\varepsilon,\delta}_{t_2,x}(n^{(1)}, \nu^{(1)}, w^{(2)}) - \mathcal{J}^{\varepsilon,\delta}_{t_1,x}(n^{(1)}, \nu^{(1)}, w^{(1)})\\
&\leq \mathsf{E}_x\Big[R^{w^{(2)}}_{T-t_2} g(T, X^{(1)}_{T-t_2}) - R^{w^{(1)}}_{T-t_1} g(T, X^{(1)}_{T-t_1}) \qquad (3.105)\\
&\qquad - \int_{T-t_2}^{T-t_1} R^{w^{(1)}}_s \Big[h + H^\varepsilon(\cdot, n^{(1)}_s \dot\nu^{(1)}_s)\Big](t_1 + s, X^{(1)}_s)\,\mathrm{d}s\\
&\qquad + \int_0^{T-t_2} R^{w^{(2)}}_s \Big[[\Delta_{t_2,t_1} h] + [\Delta_{t_2,t_1} H^\varepsilon](\cdot, n^{(1)}_s \dot\nu^{(1)}_s) + w^{(2)}_s [\Delta_{t_2,t_1} g]\Big](s, X^{(1)}_s)\,\mathrm{d}s\Big].
\end{aligned}
$$

By (3.92) the Hamiltonian H^ε is non-increasing in time, so $[\Delta_{t_2,t_1} H^\varepsilon] \leq 0$. Next, we apply Dynkin's formula to $R^{w^{(1)}}_{T-t_1} g(T, X^{(1)}_{T-t_1})$ on the time interval $[T-t_2, T-t_1]$, to obtain

$$
\begin{aligned}
&\mathsf{E}_x\Big[R^{w^{(1)}}_{T-t_1} g(T, X^{(1)}_{T-t_1})\Big]\\
&= \mathsf{E}_x\Big[R^{w^{(1)}}_{T-t_2} g(T - (t_2 - t_1), X^{(1)}_{T-t_2}) \qquad (3.106)\\
&\qquad + \int_{T-t_2}^{T-t_1} R^{w^{(1)}}_s \Big[\partial_t g + \mathcal{L} g - r g + \langle \nabla g, n^{(1)}_s \dot\nu^{(1)}_s\rangle\Big](t_1 + s, X^{(1)}_s)\mathrm{d}s\Big].
\end{aligned}
$$

Let us plug (3.106) into (3.105), recall that $R^{w^{(1)}}_{T-t_2} = R^{w^{(2)}}_{T-t_2}$ and use (3.20) with f, g replacing f_m, g_m:

$$
\begin{aligned}
u(t_2, x) - u(t_1, x) \leq \mathsf{E}_x\Big[&R^{w^{(2)}}_{T-t_2}\big(g(T, X^{(1)}_{T-t_2}) - g(T - (t_2 - t_1), X^{(1)}_{T-t_2})\big)\\
&- \int_{T-t_2}^{T-t_1} R^{w^{(1)}}_s \Theta(t_1 + s, X^{(1)}_s)\,\mathrm{d}s\\
&+ \int_0^{T-t_2} R^{w^{(2)}}_s \Big[[\Delta_{t_2,t_1} h](s, X^{(1)}_s) + w^{(2)}_s [\Delta_{t_2,t_1} g](s, X^{(1)}_s)\Big]\,\mathrm{d}s\Big],
\end{aligned}
$$

where we recall $\Theta = \partial_t g + \mathcal{L}g - rg + h$. Thanks to condition (3.11) on h and g and (3.14) on Θ

$$u(t_2,x) - u(t_1,x) \le \big(K_0(1+T) + K_2\big)(t_2 - t_1),$$

by evaluating explicitly $\int_0^{T-t_2} w_s^{(2)} R_s^{w^{(2)}} \mathrm{d}s$. Then, the claim holds with $K_4 = K_0(1+T) + K_2$. □

Our next goal is to find a uniform bound for the penalty term involving ψ_ε.

Lemma 3.29 *There is $M_5 = M_5(m)$, independent of ε and δ, such that*

$$\big\|\psi_\varepsilon\big(\mathcal{H}(\nabla u^{\varepsilon,\delta}) - f^2\big)\big\|_m \le M_5, \tag{3.107}$$

where we recall $\|\cdot\|_m = \|\cdot\|_{C^0(\overline{\mathcal{O}}_m)}$.

Proof. For notational simplicity we set $u = u^{\varepsilon,\delta}$ and $\xi = \xi_m$. Let

$$v(t,x) := \xi(x)\psi_\varepsilon\big(\mathcal{H}(\nabla u(t,x)) - f^2(t,x)\big), \qquad \text{for } (t,x) \in \overline{\mathcal{O}}_{m+1}. \tag{3.108}$$

Since v is continuous, then it attains a maximum on $\overline{\mathcal{O}}_{m+1}$. If such maximum is attained at a point $(t^*,x^*) \in \partial_P \mathcal{O}_{m+1}$ then $v(t^*,x^*) = 0$ because either $x^* \in \partial B_{m+1}$ and $\xi(x^*) = 0$ or $t^* = T$ and $|\nabla u(t^*,x^*)|_d = |\nabla g(T,x^*)|_d \le f(T,x^*)$ by (3.13). Thus, suppose the maximum is attained in $\mathcal{O}_{m+1}$.

We argue similarly to the proof of Proposition 3.17. For any $\eta > 0$ there exists a neighbourhood $U_\eta \cup \big(\{0\}\times V_\eta\big)$ of (t^*,x^*) such that $v(t,x) > v(t^*,x^*) - \eta$ for all $(t,x) \in U_\eta \cup \big(\{0\}\times V_\eta\big)$. With no loss of generality, there is $S < T$ and B an open ball with $\overline{B} \subset B_{m+1}$, so that $U_\eta \cup \big(\{0\}\times V_\eta\big) \subset \mathcal{O}_{S,B}$, where $\mathcal{O}_{S,B} = [0,S) \times B$. Let w^n be the solution of a PDE as in (3.88) but with $\nabla u^{\varepsilon,\delta}$ and $(\cdot)^+$ on the right-hand side of that equation replaced by smooth approximations ∇u^n and χ_n, and the function f^2 in the argument of ψ_ε replaced by $f^2 + \frac{1}{n}$. By arguments analogous to those in Lemma 3.15, $w^n \in C^{1,3,\alpha}_{Loc}(\mathcal{O}_{m+1})$ and $w^n \to u$ in $C^{1,2,\beta}(\overline{\mathcal{O}}_{S,B})$ and in $(C^{0,1,\beta} \cap W^{1,2,p})(\overline{\mathcal{O}}_m)$ as $n \to \infty$ for all $\beta \in (0,\alpha)$ (see Remark B.2). Define

$$v^n(t,x) := \xi(x)\psi_\varepsilon\big(\mathcal{H}(\nabla w^n(t,x)) - f^2(t,x) - \tfrac{1}{n}\big), \qquad \text{for } (t,x) \in \overline{\mathcal{O}}_{m+1}.$$

We have that v^n belongs to $C^{1,2,\alpha}_{Loc}(\mathcal{O}_{m+1}) \cap C^{0,1,\alpha}(\overline{\mathcal{O}}_{m+1})$, it is non-negative and it is equal to zero for $x \in \partial B_{m+1}$. Moreover $v^n \to v$ in $C^{0,1,\gamma}(\overline{\mathcal{O}}_{m+1})$ for all $\gamma \in (0,\alpha)$.

Let $(t_n^*, x_n^*)_{n\in\mathbb{N}}$ be such that $(t_n^*, x_n^*) \in \arg\max_{\overline{\mathcal{O}}_{m+1}} v^n$ and, with no loss of generality, assume $(t_n^*, x_n^*) \to (t^*, x^*)$. We also assume $|\nabla g(t^*, x^*)|_d - |\nabla u(t^*, x^*)|_d < 0$, as otherwise $f(t^*, x^*) \geq |\nabla g(t^*, x^*)|_d \geq |\nabla u(t^*, x^*)|_d$ implies $0 \leq v(t,x) \leq v(t^*, x^*) = 0$. By uniform convergence of ∇w^n to ∇u, we can also assume $|\nabla g|_d - |\nabla w^n|_d \leq 0$ on $U_\eta \cup (\{0\} \times V_\eta)$ for all $n \in \mathbb{N}$.

We denote $\bar{\zeta}_n := (\mathcal{H}(\nabla w^n) - f^2 - \frac{1}{n})(t_n^*, x_n^*)$ and taking derivatives of v^n we obtain

$$
\begin{aligned}
v_t^n &= \xi \psi_\varepsilon'(\bar{\zeta}_n)\big(\mathcal{H}(\nabla w^n) - f^2\big)_t \\
v_{x_i}^n &= \xi_{x_i} \psi_\varepsilon(\bar{\zeta}_n) + \xi \psi_\varepsilon'(\bar{\zeta}_n)\big(\mathcal{H}(\nabla w^n) - f^2\big)_{x_i} \\
v_{x_i x_j}^n &= \xi_{x_i x_j} \psi_\varepsilon(\bar{\zeta}_n) + \psi_\varepsilon'(\bar{\zeta}_n)\Big(\xi_{x_i}\big(\mathcal{H}(\nabla w^n) - f^2\big)_{x_j} + \xi_{x_j}\big(\mathcal{H}(\nabla w^n) - f^2\big)_{x_i}\Big) \\
&\quad + \xi \psi_\varepsilon''(\bar{\zeta}_n)\big(\mathcal{H}(\nabla w^n) - f^2\big)_{x_i}\big(\mathcal{H}(\nabla w^n) - f^2\big)_{x_j} \\
&\quad + \xi \psi_\varepsilon'(\bar{\zeta}_n)\big(\langle D^2\mathcal{H}(\nabla w^n)\nabla w_{x_j}^n, \nabla w_{x_i}^n\rangle + \langle \nabla \mathcal{H}(\nabla w^n), \nabla w_{x_i x_j}^n\rangle - \big(f^2\big)_{x_i x_j}\big).
\end{aligned}
\tag{3.109}
$$

By (3.38) we have $0 \geq (\partial_t v^n + \mathcal{L} v^n)(t_n^*, x_n^*)$. Since (t_n^*, x_n^*) is fixed, we omit it from the calculations that follow, for notational simplicity. Then using (3.109) and symmetry of a_{ij}

$$
\begin{aligned}
0 \geq\, &(\mathcal{L}\xi)\psi_\varepsilon(\bar{\zeta}_n) - \xi \psi_\varepsilon'(\bar{\zeta}_n)\big(\partial_t(f^2) + \mathcal{L}(f^2)\big) + \xi \psi_\varepsilon'(\bar{\zeta}_n)\big\langle \mathcal{H}(\nabla w^n), (\partial_t + \mathcal{L})(\nabla w^n)\big\rangle \\
&+ \tfrac{1}{2}\xi \psi_\varepsilon''(\bar{\zeta}_n)\big\langle a\nabla(\mathcal{H}(\nabla w^n) - f^2), \nabla(\mathcal{H}(\nabla w^n) - f^2)\big\rangle \\
&+ \psi_\varepsilon'(\bar{\zeta}_n)\big\langle a\nabla\xi, \nabla(\mathcal{H}(\nabla w^n) - f^2)\big\rangle \\
&+ \frac{1}{2}\xi \psi_\varepsilon'(\bar{\zeta}_n)\sum_{i,j=1}^d a_{ij}\big\langle D^2\mathcal{H}(\nabla w^n)\nabla w_{x_i}^n, \nabla w_{x_j}^n\big\rangle,
\end{aligned}
\tag{3.110}
$$

where $(\partial_t + \mathcal{L})(\nabla w^n)$ is the vector with entries $(\partial_t + \mathcal{L})w_{x_k}^n$ for $k = 1, \ldots, d$.

Using that $\mathcal{H}$ is diagonal with entries 2, we have

$$
\frac{1}{2}\big\langle D^2\mathcal{H}(\nabla w^n)\nabla w_{x_i}^n, \nabla w_{x_j}^n\big\rangle = \big\langle \nabla w_{x_i}^n, \nabla w_{x_j}^n\big\rangle \tag{3.111}
$$

and recalling that ψ_ε is non-decreasing and convex, so $\psi_\varepsilon', \psi_\varepsilon'' \geq 0$, the last term on the right-hand side of (3.110) is bounded from below by $\xi \psi_\varepsilon'(\bar{\zeta}_n)\theta\big|D^2 w^n\big|_{d\times d}^2$ as in (3.47) with θ equal to $\theta_{B_{m+1}}$. Uniform ellipticity (3.10) on B_{m+1} also gives

$$\tfrac{1}{2}\xi\psi''_\varepsilon(\bar{\zeta}_n)\big\langle a\nabla(\mathcal{H}(\nabla w^n)-f^2),\nabla(\mathcal{H}(\nabla w^n)-f^2)\big\rangle\ge 0.$$

Set $\bar{a}_m:=\max_{i,j}\|a_{ij}\|_{C^0(\overline{B}_{m+1})}$ and recall $|\nabla\xi|_d^2\le C_0\xi$ (see (3.16)). Then

$$\begin{aligned}&\big\langle a\nabla\xi,\nabla(\mathcal{H}(\nabla w^n)-f^2)\big\rangle\\&\quad\ge-\bar{a}_m d^2|\nabla\xi|_d\Big(2|\nabla w^n|_d|D^2w^n|_{d\times d}+|\nabla f^2|_d\Big)\\&\quad\ge-\xi\frac{\theta}{4}|D^2w^n|^2_{d\times d}-\frac{16}{\theta}\bar{a}_m^2d^4C_0|\nabla w^n|_d^2-\bar{a}_md^2\sqrt{C_0\xi}|\nabla f^2|_d,\end{aligned}\tag{3.112}$$

where we used $|ab|\le pa^2+b^2/p$ with $p=\frac{4}{\theta}$, $b=\sqrt{\xi}|D^2w^n|_{d\times d}$ and $a=2\bar{a}_md^2\sqrt{C_0}|\nabla w^n|_d$. Since $\nabla w^n\to\nabla u$ uniformly on $\overline{\mathcal{O}}_{m+1}$, then by Proposition 3.22 we can assume $|\nabla w^n|_d\le 1+N_1$ and obtain

$$\big\langle a\nabla\xi,\nabla(\mathcal{H}(\nabla w^n)-f^2)\big\rangle\ge-\xi\frac{\theta}{4}|D^2w^n|^2_{d\times d}-C_1,$$

with $C_1=16d^4\bar{a}_m^2C_0\theta^{-1}(1+N_1)^2+2\bar{a}_md^2\sqrt{C_0\xi}\|\nabla f^2\|_{m+1}$. Since $\mathcal{L}\xi$ and $(\partial_t+\mathcal{L})f^2$ are continuous on $\overline{\mathcal{O}}_{m+1}$ we have $|\mathcal{L}\xi|+|(\partial_t+\mathcal{L})f^2|\le C_2$ on $\overline{\mathcal{O}}_{m+1}$. Similarly to the first inequality in (3.65), $\psi_\varepsilon(\bar{\zeta}_n)\le\psi'_\varepsilon(\bar{\zeta}_n)\mathcal{H}(\nabla w^n)=\psi'_\varepsilon(\bar{\zeta}_n)|\nabla w^n|_d^2\le\psi'_\varepsilon(\bar{\zeta}_n)(1+N_1)^2$. Thus

$$\psi_\varepsilon(\bar{\zeta}_n)|\mathcal{L}\xi|\le\psi'_\varepsilon(\bar{\zeta}_n)(1+N_1)^2C_2=:\psi'_\varepsilon(\bar{\zeta}_n)C_3,$$

where $C_3=C_3(m)>0$.

We claim that, for any $\lambda>0$ there are constants $C_4=C_4(m)>0$ and $\kappa_{\delta,m}>0$ such that

$$\begin{aligned}&\xi\langle\nabla w^n,(\partial_t+\mathcal{L})(\nabla w^n)\rangle\\&\quad\ge-\tfrac{\theta}{4}\xi|D^2w^n|^2_{d\times d}-2\xi\lambda\theta\psi_\varepsilon^2(\zeta_n)-C_4(1+\lambda^{-1})-\kappa_{\delta,m}R_n,\end{aligned}\tag{3.113}$$

where $\zeta_n:=(|\nabla u^n|_d^2-f^2-\frac{1}{n})(t_n^*,x_n^*)$ and R_n is independent of (t_n^*,x_n^*) and such that $R_n\to 0$ as $n\to\infty$. The claim is proven separately at the end of this proof, for the sake of readability. Plugging all the above estimates into (3.110) and factoring out $\psi'_\varepsilon(\bar{\zeta}_n) \geq 1$ gives us

$$0 \leq -\tfrac{\theta}{2}\xi|D^2w^n|^2_{d\times d} + 2\lambda\theta\xi\psi^2_\varepsilon(\zeta_n) + C_1 + C_2 + C_3 + C_4(1+\lambda^{-1}) + \kappa_{\delta,m}R_n.$$

Letting $C_5 = C_5(m) > 0$ be a suitable constant the expression simplifies to

$$\tfrac{\theta}{2}\xi|D^2w^n|^2_{d\times d} \leq 2\lambda\theta\xi\psi^2_\varepsilon(\zeta_n) + C_5(1+\lambda^{-1}) + \kappa_{\delta,m}R_n.$$

We want to bound $\xi|D^2w^n|_{d\times d}$ by $\xi\psi_\varepsilon(\zeta_n)$. So we multiply both sides of the inequality above by ξ, take square root and use $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ for $a, b \geq 0$ and $|\xi| \leq 1$. That gives

$$\xi|D^2w^n|_{d\times d} \leq 2\sqrt{\lambda}\xi\psi_\varepsilon(\zeta_n) + \sqrt{2\theta^{-1}C_5(1+\lambda^{-1})} + \sqrt{2\theta^{-1}\kappa_{\delta,m}R_n}. \tag{3.114}$$

Recall that w^n solves

$$\partial_t w^n + \mathcal{L}w^n - rw^n = \psi_\varepsilon(\zeta_n) - h - \tfrac{1}{\delta}\chi_n(g-u), \quad \text{on } [0,T)\times\mathbb{R}^d.$$

Multiplying by ξ we can express $\xi\psi_\varepsilon(\zeta_n)$ in terms of the remaining functions in the equation above. Since $(t^*_n, x^*_n) \in \mathcal{O}_{S,B}$, $w^n \to u$ in $C^{1,2,\beta}(\overline{\mathcal{O}}_{S,B})$ and $\chi_n(g-u) \to (g-u)^+$ uniformly on compacts, we can assume with no loss of generality that on $\overline{\mathcal{O}}_{S,B}$ the following hold: $\frac{1}{\delta}\chi_n(g-u) \leq (1+K_2)$ by (3.103), $|\nabla w^n|_d \leq (1+N_1)$ by (3.75), $\partial_t w^n \leq \frac{1}{2} + K_4$ by (3.104) and $rw^n \geq -\frac{1}{2}$ because $u \geq 0$. The coefficients a and b in $\mathcal{L}$ are bounded on $\overline{B}_{m+1}$ by a constant A_{m+1} (slightly abusing notation). Thus,

$$\begin{aligned}\xi\psi_\varepsilon(\zeta_n) &= \xi\partial_t w^n - \xi rw^n + \xi\mathcal{L}w^n + \xi h + \xi\tfrac{1}{\delta}\chi_n(g-u)\\ &\leq 1 + K_4 + A_{m+1}(1+N_1) + \tfrac{1}{2}A_{m+1}\xi|D^2w^n|_{d\times d}\\ &\quad + \|h\|_{m+1} + (1+K_2).\end{aligned}$$

Substituting (3.114) and grouping together the constants we obtain, for some $C_6 = C_6(m) > 0$,

$$\Big(1 - \sqrt{\lambda}A_{m+1}\Big)\xi\psi_\varepsilon(\zeta_n) \leq C_6\sqrt{1+\lambda^{-1}+\kappa_{\delta,m}R_n}. \tag{3.115}$$

Then, choosing $\lambda = (4A^2_{m+1})^{-1}$ and recalling that all expressions are evaluated at

(t_n^*, x_n^*) we obtain

$$\xi(x_n^*)\psi_\varepsilon\big(\mathcal{H}(\nabla u^n(t_n^*, x_n^*)) - f^2(t_n^*, x_n^*) - \tfrac{1}{n}\big) \le 2C_6\sqrt{1 + 4A_{m+1}^2 + \kappa_{\delta,m}R_n}.$$

Taking limits as $n \to \infty$, using that $(t_n^*, x_n^*) \to (t^*, x^*)$, $R_n \to 0$ and $\nabla u^n \to \nabla u$ (uniformly on compacts), we have

$$\xi(x^*)\psi_\varepsilon\big(\mathcal{H}(\nabla u(t^*, x^*)) - f^2(t^*, x^*)\big) \le 2C_6\sqrt{1 + 4A_{m+1}^2} =: M_5.$$

Recalling the definition of v in (3.108) we can conclude:

$$\big\|\psi_\varepsilon\big(\mathcal{H}(\nabla u) - f^2\big)\big\|_m \le \sup_{(t,x)\in\overline{\mathcal{O}}_{m+1}} v(t,x) = v(t^*, x^*) \le M_5,$$

with $M_5 = M_5(m)$ independent of δ and ε. □

Proof of (3.113). Recall that $u = u^{\varepsilon,\delta}$ and that w^n solves

$$(\partial_t + \mathcal{L} - r)w^n = -h - \tfrac{1}{\delta}\chi_n\big(g - u^{\varepsilon,\delta}\big) + \psi_\varepsilon(\mathcal{H}(\nabla u^n) - f^2 - \tfrac{1}{n}), \text{ on } [0,T) \times \mathbb{R}^d.$$

Differentiating with respect to x_k, multiplying by ξ and evaluating at (t_n^*, x_n^*) we get

$$\begin{aligned}\xi(\partial_t w_{x_k}^n + \mathcal{L}w_{x_k}^n) = & -\xi\mathcal{L}_{x_k}w^n + \xi r w_{x_k}^n - \xi h_{x_k} - \xi\tfrac{1}{\delta}\chi_n'(g-u)(g-u)_{x_k} \\ & + \xi\psi_\varepsilon'(\zeta_n)(\mathcal{H}(\nabla u^n) - f^2 - \tfrac{1}{n})_{x_k},\end{aligned} \tag{3.116}$$

where $\zeta_n = (\mathcal{H}(\nabla u^n) - f^2 - \frac{1}{n})(t_n^*, x_n^*)$. We subtract the term

$$v_{x_k}^n = \xi_{x_k}\psi_\varepsilon(\bar{\zeta}_n) + \xi\psi_\varepsilon'(\bar{\zeta}_n)(\mathcal{H}(\nabla w^n) - f^2)_{x_k}$$

from both sides of (3.116), and we add and subtract $\xi_{x_k}\psi_\varepsilon(\zeta_n)$ on the right-hand side of (3.116). Then

$$\begin{aligned}\xi(\partial_t w_{x_k}^n + \mathcal{L}w_{x_k}^n) - v_{x_k}^n = & -\xi\mathcal{L}_{x_k}w^n + \xi r w_{x_k}^n - \xi h_{x_k} - \xi\tfrac{1}{\delta}\chi_n'(g-u)(g-u)_{x_k} \\ & - \xi_{x_k}\psi_\varepsilon(\zeta_n) + P_{n,k},\end{aligned} \tag{3.117}$$

where

$$
\begin{aligned}
P_{n,k} = &\xi_{x_k}\Big(\psi_\varepsilon(\zeta_n) - \psi_\varepsilon(\bar{\zeta}_n)\Big) \\
&+ \xi\Big(\psi'_\varepsilon(\zeta_n)\big(\mathcal{H}(\nabla u^n) - f^2\big)_{x_k} - \psi'_\varepsilon(\bar{\zeta}_n)\big(\mathcal{H}(\nabla w^n) - f^2\big)_{x_k}\Big).
\end{aligned}
$$

Recall that $\nabla\mathcal{H}(\nabla w^n) = 2\nabla w^n$ and that $(t_n^*, x_n^*) \in \mathcal{O}_{S,B}$ is a stationary point for v^n in the spatial coordinates, so $v^n_{x_k} = 0$ in (3.117) for each $1 \le k \le d$. Then

$$
\begin{aligned}
\xi\langle\, \nabla\mathcal{H}(\nabla w^n), (\partial_t + \mathcal{L})(\nabla w^n)\rangle = &2\xi\Big(- \sum_{k=1}^d w^n_{x_k}\mathcal{L}_{x_k} w^n + r|\nabla w^n|_d^2 \\
&- \langle \nabla w^n, \nabla h + \tfrac{1}{\delta}\chi'_n(g-u)\nabla(g-u)\rangle\Big) \\
&- 2\psi_\varepsilon(\zeta_n)\langle \nabla w^n, \nabla\xi\rangle + 2\sum_{k=1}^d w^n_{x_k} P_{n,k}.
\end{aligned} \tag{3.118}
$$

Since $(t_n^*, x_n^*) \in \mathcal{O}_{S,B}$, then denoting $\|\cdot\|_{S,B} = \|\cdot\|_{C^0(\overline{\mathcal{O}}_{S,B})}$ we have

$$
\begin{aligned}
2\sum_{k=1}^d w^n_{x_k} P_{n,k} \ge &- 2\|\psi_\varepsilon(\zeta_n) - \psi_\varepsilon(\bar{\zeta}_n)\|_{S,B}\|\nabla w^n\|_{S,B}\|\nabla\xi\|_{S,B} \\
&- 2\|\psi'_\varepsilon(\zeta_n) - \psi'_\varepsilon(\bar{\zeta}_n)\|_{S,B}\|\nabla w^n\|_{S,B}\|\nabla\big(|\nabla w^n|_d^2 - f^2\big)\|_{S,B} \\
&- 2\|\psi'_\varepsilon(\zeta_n)\|_{S,B}\|\nabla w^n\|_{S,B}\|\nabla(|\nabla u^n|_d^2 - |\nabla w^n|_d^2)\|_{S,B} =: \tilde{R}_n,
\end{aligned}
$$

where $\tilde{R}_n \to 0$ as $n \to \infty$ thanks to $C^{1,2,\beta}(\overline{\mathcal{O}}_{S,B})$-convergence of w^n and u^n to u, for $\beta \in (0,\alpha)$. By Cauchy-Schwarz inequality, recalling that $0 \le \chi'_n(\cdot) \le 2$ and using arguments as in (3.59) we have

$$
\begin{aligned}
&\langle \nabla w^n, \nabla h\rangle \le \|\nabla w^n\|_{m+1}\|\nabla h\|_{m+1} \le (N_1+1)\|\nabla h\|_{m+1}, \\
&\tfrac{1}{\delta}\chi'_n(g-u)\langle \nabla w^n, \nabla(g-u)\rangle \\
&\qquad \le \tfrac{2}{\delta}\|\nabla\hat{w}^n\|_{S,B}\big(\|\nabla g\|_{m+1} + \|\nabla u\|_{m+1}\big) =: \kappa_{\delta,m}\tilde{R}'_n,
\end{aligned}
$$

where $N_1 = N_1(m+1)$ is as in Proposition 3.22 and $\hat{w}^n = u - w^n$. Recall that $\|\nabla\hat{w}^n\|_{S,B} \to 0$ as $n \to \infty$, therefore $\tilde{R}'_n \to 0$ too. For the penultimate term on the right-hand side of (3.118), recalling $|\nabla\xi|_d \le \sqrt{C_0\xi}$ (see (3.16)), we have

$$\begin{aligned}2\psi_\varepsilon(\zeta_n)\langle\nabla w^n, \nabla\xi\rangle &\le 2|\nabla w^n|_d|\nabla\xi|_d\psi_\varepsilon(\zeta_n)\\ &\le 2(N_1+1)\sqrt{C_0\xi}\psi_\varepsilon(\zeta_n)\\ &\le \frac{(N_1+1)^2 2C_0}{\lambda\theta} + 2\xi\lambda\theta\psi_\varepsilon^2(\zeta_n),\end{aligned}$$

where in the last inequality we used $ab \le \frac{a^2}{p} + pb^2$ with $a = (N_1+1)\sqrt{2C_0}$, $b = \sqrt{2\xi}\psi_\varepsilon(\zeta_n)$, and $p = \lambda\theta$, with λ a constant to be chosen later and $\theta = \theta_{B_{m+1}}$ as in (3.10). For the first term on the right-hand side of (3.118) we argue as in (3.60) and obtain

$$\sum_{k=1}^{d} w^n_{x_k}\mathcal{L}_{x_k}w^n \le \tfrac{\theta}{8}|D^2w^n|^2_{d\times d} + C_1(N_1+1)^2, \tag{3.119}$$

where $C_1 := 8d^4A^2_{m+1}\theta^{-1} + 2dA_{m+1}$, the constant A_{m+1} is defined as in (3.61) and, differently from (3.62), we use $ab \le pa^2 + \frac{b^2}{p}$ with $p = \frac{\theta}{8}$, $a = |D^2w^n|_{d\times d}$ and $b = \frac{d^2}{2}A_{m+1}|\nabla w^n|_d$.

Combining these bounds we get

$$\begin{aligned}\xi\langle\nabla\mathcal{H}(\nabla w^n), (\partial_t+\mathcal{L})(\nabla w^n)\rangle \ge &-\xi\tfrac{\theta}{4}|D^2w^n|^2_{d\times d} - 2\xi\lambda\theta\psi^2_\varepsilon(\zeta_n)\\ &- C_4(1+\lambda^{-1}) - \kappa_{\delta,m}R_n,\end{aligned}$$

where we define $C_4 := 2C_1(N_1+1)^2 + (N_1+1)^2 2C_0\theta^{-1} + 2(N_1+1)\|\nabla h\|_{m+1}$ and we collect $\tilde{R}_n$ and $2\kappa_{\delta,m}\tilde{R}'_n$ in $\kappa_{\delta,m}R_n$ with an abuse of notation. □

The bounds on the penalty terms in the PDE for $u^{\varepsilon,\delta}$ enable the next estimate.

Theorem 3.30 *For any $p \in (1,\infty)$, there is $M_6 = M_6(m,p)$ such that*

$$\|u^{\varepsilon,\delta}\|_{W^{1,2,p}(\mathcal{O}_m)} \le M_6, \quad \text{for all } \varepsilon,\delta \in (0,1). \tag{3.120}$$

Proof. The proof repeats the exact same arguments as in the proof of Proposition 3.23 but applied to $\varphi = \xi_{m+1}u^{\varepsilon,\delta}$ rather than to $\varphi = \xi_{m_0}u^{\varepsilon,\delta}_q$. In addition, we use Lemmas 3.27 and 3.29 to obtain the upper bound for $\|\varphi\|_{W^{1,2,p}(\mathcal{O}_{m+1})}$ as in (3.87), which is therefore independent of ε, δ. □

3.4 The Variational Inequality

In this section, we finally prove our main result, i.e., Theorem 3.6. First we prove that Problem A admits a solution (Theorem 3.31), then we prove that such solution is the value function of our game (Theorem 3.33) and it is the maximal solution for Problem A.

Theorem 3.31 *There exists a solution u of Problem A.*

Proof. Let $(\varepsilon_k)_{k\in\mathbb{N}}$ be a decreasing sequence with $\varepsilon_k \to 0$. Fix $k, m \in \mathbb{N}$. Thanks to Theorem 3.30 and the compact embedding of $W^{1,2,p}(\mathcal{O}_m)$ into $C^{0,1,\beta}(\overline{\mathcal{O}}_m)$ for $\beta = 1-(d+2)/p$, we can extract a sequence $(u^{\varepsilon_k,\delta^m_{k,j}})_{j\in\mathbb{N}}$ converging to a limit $u^{\varepsilon_k;[m]}$ (possibly depending on $\mathcal{O}_m$) as $j \to \infty$, in the following sense:

$$\begin{aligned}
&u^{\varepsilon_k,\delta^m_{k,j}} \to u^{\varepsilon_k;[m]} \quad \text{and} \quad \nabla u^{\varepsilon_k,\delta^m_{k,j}} \to \nabla u^{\varepsilon_k;[m]} \quad \text{in } C^\alpha(\overline{\mathcal{O}}_m), \\
&\partial_t u^{\varepsilon_k,\delta^m_{k,j}} \to \partial_t u^{\varepsilon_k;[m]} \quad \text{and} \quad D^2 u^{\varepsilon_k,\delta^m_{k,j}} \to D^2 u^{\varepsilon_k;[m]} \quad \text{weakly in } L^p(\mathcal{O}_m).
\end{aligned} \tag{3.121}$$

Up to selecting further subsequences (if needed), we find analogous limits on $\mathcal{O}_{m+1} \subset \mathcal{O}_{m+2} \subset \dots$ so that $u^{\varepsilon_k;[m]} = u^{\varepsilon_k;[m+1]}$ on $\overline{\mathcal{O}}_m$, $u^{\varepsilon_k;[m+1]} = u^{\varepsilon_k;[m+2]}$ on $\overline{\mathcal{O}}_{m+1}$ and so on. Since $\overline{\mathcal{O}}_m \uparrow \mathbb{R}^{d+1}_{0,T}$ as $m \to \infty$, iterating this procedure we can define a limit function u^{ε_k} on $\mathbb{R}^{d+1}_{0,T}$.

The sequence $(u^{\varepsilon_k})_{k\in\mathbb{N}}$ satisfies the same bound as in (3.120). Therefore, by the same argument as above we can extract a further converging subsequence, which we denote still by $(u^{\varepsilon_k})_{k\in\mathbb{N}}$ with an abuse of notation. That is, there is a function u on $\mathbb{R}^{d+1}_{0,T}$ such that for any $m \in \mathbb{N}$

$$\begin{aligned}
&u^{\varepsilon_k} \to u \quad \text{and} \quad \nabla u^{\varepsilon_k} \to \nabla u \quad \text{in } C^\alpha(\overline{\mathcal{O}}_m), \\
&\partial_t u^{\varepsilon_k} \to \partial_t u \quad \text{and} \quad D^2 u^{\varepsilon_k} \to D^2 u \quad \text{weakly in } L^p(\mathcal{O}_m).
\end{aligned}$$

Finally, we can extract a diagonal subsequence $(u^{\varepsilon_i,\delta_i})_{i\in\mathbb{N}}$ that converges to u locally on $\mathbb{R}^{d+1}_{0,T}$ in the sense above as $(\varepsilon_i, \delta_i) \to 0$, simultaneously.

Next we prove that the limit function u is solution of Problem A. By construction, $u \in W^{1,2,p}_{\ell oc}(\mathbb{R}^{d+1}_{0,T})$. Thanks to (3.103), (3.107) and $C^{0,1,\alpha}$-convergence on compacts, in the limit as $\varepsilon, \delta \to 0$ we obtain

$$g(t,x) - u(t,x) \le 0 \quad \text{and} \quad |\nabla u(t,x)|_d - f(t,x) \le 0,$$

for all $(t,x) \in \mathbb{R}^{d+1}_{0,T}$.

Fix $(\bar{t},\bar{x}) \in \mathcal{C}$, i.e., $u(\bar{t},\bar{x}) > g(\bar{t},\bar{x})$. By continuity of u and g there is an open neighbourhood $\mathcal{O}$ of $(\bar{t},\bar{x})$ such that $u(t,x) > g(t,x)$ for all $(t,x) \in \mathcal{O}$. Uniform convergence on compacts of $u^{\varepsilon_i,\delta_i}$ to u also guarantees that $u^{\varepsilon_i,\delta_i} > g$ on $\mathcal{O}$, for sufficiently large i's. Then, for large i's (3.88) reads

$$\partial_t u^{\varepsilon_i,\delta_i} + \mathcal{L}u^{\varepsilon_i,\delta_i} - ru^{\varepsilon_i,\delta_i} = -h + \psi_\varepsilon\big(\mathcal{H}(\nabla u^{\varepsilon_i,\delta_i}) - f^2\big) \geq -h, \quad \text{on } \mathcal{O}.$$

Multiplying the equation above by $\phi \in C^\infty_c(\mathcal{O})$, $\phi \geq 0$ and letting $i \to \infty$ we obtain the second equation in (3.6). Analogously, let $(\bar{t},\bar{x}) \in \mathcal{I}$, i.e., $|\nabla u(\bar{t},\bar{x})|_d < f(\bar{t},\bar{x})$. Then, by continuity of ∇u and uniform convergence on compacts of $\nabla u^{\varepsilon_i,\delta_i} \to \nabla u$ we find an open neighbourhood $\mathcal{O}$ such that $|\nabla u|_d < f$ on $\mathcal{O}$ and $|\nabla u^{\varepsilon_i,\delta_i}|_d < f$ on $\mathcal{O}$ for sufficiently large i. In such neighbourhood (3.88) reads

$$\partial_t u^{\varepsilon_i,\delta_i} + \mathcal{L}u^{\varepsilon_i,\delta_i} - ru^{\varepsilon_i,\delta_i} = -h - \tfrac{1}{\delta}\big(g - u^{\varepsilon_i,\delta_i}\big)^+ \leq -h,$$

and by the same argument as above, using test functions, we can pass to the limit and obtain the third equation in (3.6). The case in which $(\bar{t},\bar{x}) \in \mathcal{I} \cap \mathcal{C}$ is now obvious and the first equation in (3.6) also holds for all (t,x) by standard PDE theory (e.g., as in the proof of Theorem 3.24). Finally, the terminal condition is trivially satisfied since $u^{\varepsilon,\delta}(T,x) = g(T,x)$ for all $(\varepsilon,\delta) \in (0,1)^2$.

Notice that u has at most quadratic growth by Lemma 3.13. □

To prove that a solution of Problem A is the value of our game, we need the next lemma.

Lemma 3.32 *Let $(t,x) \in \mathbb{R}^{d+1}_{0,T}$. For any $\tau \in \mathcal{T}_t$ we have*

$$\inf_{(n,\nu)\in\mathcal{A}_t} \mathcal{J}_{t,x}(n,\nu,\tau) = \inf_{(n,\nu)\in\mathcal{A}^\tau_t} \mathcal{J}_{t,x}(n,\nu,\tau),$$

where $\mathcal{A}^\tau_t := \{(n,\nu) \in \mathcal{A}_t \,|\, \nu_\tau = \nu_{\tau-},\ \mathsf{P}_x - a.s.\}$.

Proof. In the expression of $\mathcal{J}_{t,x}(n,\nu,\tau)$, for any treble $[(n,\nu),\tau]\in\mathcal{A}_t\times\mathcal{T}_t$ we have

$$\begin{aligned}&e^{-r\tau}g(t+\tau,X^{[n,\nu]}_{\tau})+\int_{[0,\tau]}e^{-rs}f(t+s,X^{[n,\nu]}_{s})\circ\mathrm{d}\nu_s\\&=e^{-r\tau}g(t+\tau,X^{[n,\nu]}_{\tau-})+\int_{[0,\tau)}e^{-rs}f(t+s,X^{[n,\nu]}_{s})\circ\mathrm{d}\nu_s\\&\quad+e^{-r\tau}\int_0^{\Delta\nu_\tau}\big(\langle\nabla g,n_\tau\rangle+f\big)(t+\tau,X^{[n,\nu]}_{\tau-}+\lambda n_\tau)\mathrm{d}\lambda\\&\geq e^{-r\tau}g(t+\tau,X^{[n,\nu]}_{\tau-})+\int_{[0,\tau)}e^{-rs}f(t+s,X^{[n,\nu]}_{s})\circ\mathrm{d}\nu_s,\end{aligned}$$

where the final inequality is due to (3.13). Therefore, the controller attains a lower payoff by avoiding a jump of the control at time τ. That concludes the proof. □

Theorem 3.33 *The game in* (3.4) *admits a value* v *which is also the maximal solution of Problem* A*. Moreover,* τ_* *defined in* (3.15) *is optimal for the stopper.*

Proof. Fix $(t,x)\in\mathbb{R}^{d+1}_{0,T}$, let $[(n,\nu),\tau]\in\mathcal{A}_t\times\mathcal{T}_t$ and recall ρ_m. By regularity of $u^{\varepsilon,\delta}$ (Theorem 3.24), letting $\rho^k_m=\rho_m\wedge(T-t-k^{-1})^+$ Itô's formula applies to $e^{-r(\tau\wedge\rho^k_m)}u^{\varepsilon,\delta}(t+\tau\wedge\rho^k_m,X^{[n,\nu]}_{\tau\wedge\rho^k_m})$. Using that $u^{\varepsilon,\delta}$ solves Problem C, taking expectations and letting $k\uparrow\infty$ we obtain

$$\begin{aligned}u^{\varepsilon,\delta}(t,x)=\mathsf{E}_x\Big[&e^{-r(\tau\wedge\rho_m)}u^{\varepsilon,\delta}(t+\tau\wedge\rho_m,X^{[n,\nu]}_{\tau\wedge\rho_m})\\&+\int_0^{\tau\wedge\rho_m}e^{-rs}\big[h+\tfrac{1}{\delta}\big(g-u^{\varepsilon,\delta}\big)^+-\psi_\varepsilon\big(\mathcal{H}(\nabla u^{\varepsilon,\delta})-f^2\big)\big](t+s,X^{[n,\nu]}_{s-})\mathrm{d}s\\&-\int_0^{\tau\wedge\rho_m}e^{-rs}\langle\nabla u^{\varepsilon,\delta}(t+s,X^{[n,\nu]}_{s-}),n_s\rangle\,\mathrm{d}\nu^c_s\\&-\sum_{0\leq s\leq\tau\wedge\rho_m}e^{-rs}\int_0^{\Delta\nu_s}\langle\nabla u^{\varepsilon,\delta}(t+s,X^{[n,\nu]}_{s-}+\lambda n_s),n_s\rangle\,\mathrm{d}\lambda\Big].\end{aligned}\tag{3.122}$$

We want to take limits as $\varepsilon,\delta\to 0$ and pass the limits under expectations. To do that we notice that $X^{[n,\nu]}_{s-}\in\overline{B}_m$ for all $s\in[0,\rho_m]$, P_x-a.s. Then the terms under the integral with respect to 'ds' are bounded thanks to Assumption 3.5, Lemma 3.27 and Lemma 3.29. Since $\nabla u^{\varepsilon,\delta}$ is also bounded by $N_1(m)$ (Proposition 3.22 with m_0 therein replaced by m), the integrals with respect to the control are bounded by $N_1(m)\nu_{T-t}$, which is square integrable by definition of $\mathcal{A}_t$. Finally, recall that $u^{\varepsilon,\delta}$ has quadratic growth by Lemma 3.13 and notice that

$$\mathsf{E}_x\Big[\sup_{0\le s\le T-t}\big|X^{[n,\nu]}_s\big|^2_d\Big]\le c\big(1+|x|^2_d+\mathsf{E}_x[|\nu_{T-t}|^2]\big),\tag{3.123}$$

by standard estimates for SDEs [48, Thm. 2.5.10], with $c>0$ independent of m, ε, δ. Then we are allowed to use dominated convergence and it remains to evaluate the limit.

For P_x-a.e. $\omega\in\Omega$ there is a compact $\mathcal{K}_\omega\subset\mathbb{R}^d$ such that $X^{[n,\nu]}_s(\omega)\in\mathcal{K}_\omega$ for all $s\in[0,\rho_m(\omega)]$, by right-continuity of the process and the fact that ν is square integrable. Then, uniform convergence of $(u^{\varepsilon,\delta},\nabla u^{\varepsilon,\delta})$ to $(u,\nabla u)$ on compacts implies

$$\lim_{\varepsilon,\delta\to0}\langle\nabla u^{\varepsilon,\delta}(t+s,X^{[n,\nu]}_{s-}),n_s\rangle(\omega)=\langle\nabla u(t+s,X^{[n,\nu]}_{s-}),n_s\rangle(\omega),$$
$$\lim_{\varepsilon,\delta\to0}\langle\nabla u^{\varepsilon,\delta}(t+s,X^{[n,\nu]}_{s-}+\lambda n_s),n_s\rangle(\omega)=\langle\nabla u(t+s,X^{[n,\nu]}_{s-}+\lambda n_s),n_s\rangle(\omega),$$

for all $s\in[0,\rho_m(\omega)]$ and all $\lambda\in[0,\Delta\nu_s(\omega)]$, for P_x-a.e. $\omega\in\Omega$. Moreover, for arbitrary $\eta>0$, choosing $\tau=\tau_\eta=\inf\{s\ge0\,|\,u(t+s,X^{[n,\nu]}_s)\le g(t+s,X^{[n,\nu]}_s)+\eta\}$ gives

$$\liminf_{\varepsilon,\delta\to0}\,(u^{\varepsilon,\delta}-g)(t+s,X^{[n,\nu]}_{s-})\ge\eta,\quad\text{for all } s\in[0,\tau_\eta\wedge\rho_m],\ \mathsf{P}_x\text{-a.s.}$$

Using the observations above, combined with dominated convergence and $\psi_\varepsilon\ge0$ we obtain

$$\begin{aligned}u(t,x)\le\mathsf{E}_x\Bigg[&e^{-r(\tau_\eta\wedge\rho_m)}u(t+\tau_\eta\wedge\rho_m,X^{[n,\nu]}_{\tau_\eta\wedge\rho_m})+\int_0^{\tau_\eta\wedge\rho_m}e^{-rs}h(t+s,X^{[n,\nu]}_s)\,\mathrm{d}s\\&-\int_0^{\tau_\eta\wedge\rho_m}e^{-rs}\langle\nabla u(t+s,X^{[n,\nu]}_{s-}),n_s\rangle\,\mathrm{d}\nu^c_s\\&-\sum_{0\le s\le\tau_\eta\wedge\rho_m}e^{-rs}\int_0^{\Delta\nu_s}\langle\nabla u(t+s,X^{[n,\nu]}_{s-}+\lambda n_s),n_s\rangle\,\mathrm{d}\lambda\Bigg].\end{aligned}$$

By $|\nabla u|_d \leq f$ and the definition of $\tau_\eta \wedge \rho_m$ we have

$$u(t,x) \leq \eta + \mathsf{E}_x\Big[e^{-r\tau_\eta} g(t+\tau_\eta, X^{[n,\nu]}_{\tau_\eta})\mathbb{1}_{\{\tau_\eta \leq \rho_m\}} + e^{-r\rho_m} u(t+\rho_m, X^{[n,\nu]}_{\rho_m})\mathbb{1}_{\{\tau_\eta > \rho_m\}}\Big]$$
$$+\mathsf{E}_x\Big[\int_0^{\tau_\eta\wedge\rho_m} e^{-rs} h(t+s, X^{[n,\nu]}_s)\,\mathrm{d}s + \int_{[0,\tau_\eta\wedge\rho_m]} e^{-rs} f(t+s, X^{[n,\nu]}_s)\circ \mathrm{d}\nu_s\Big].$$

Now we let $m \to \infty$. Clearly $\rho_m \uparrow T-t$, P_x-a.s. by (3.123). Since the bound in (3.123) is independent of m and functions u and g have at most quadratic growth we can apply the dominated convergence theorem to pass the limit inside the first expectation. We can also take the limit inside the second expectation by monotone convergence as all terms under the integral are non-negative. For P_x-a.e. $\omega \in \Omega$ we have $X^{[n,\nu]}_s(\omega) \in \mathcal{K}_\omega$ for all $s \in [0, T-t]$. Then

$$\begin{aligned}
&\lim_{m\to\infty} e^{-r\rho_m(\omega)} u\big(t+\rho_m(\omega), X^{[n,\nu]}_{\rho_m}(\omega)\big)\mathbb{1}_{\{\tau_\eta>\rho_m\}}(\omega)\\
&\quad= e^{-r(T-t)} u\big(T, X^{[n,\nu]}_{T-t}(\omega)\big)\mathbb{1}_{\{\tau_\eta\geq T-t\}}(\omega)\\
&\quad= e^{-r(T-t)} g\big(T, X^{[n,\nu]}_{T-t}(\omega)\big)\mathbb{1}_{\{\tau_\eta= T-t\}}(\omega), \quad \mathsf{P}_x\text{-a.e. } \omega\in\Omega,
\end{aligned}$$

because $\tau_\eta(\omega) \leq T-t$ and u is uniformly continuous on $[0,T]\times\mathcal{K}_\omega$. Hence, for $m \to \infty$ we obtain

$$\begin{aligned}
u(t,x) &\leq \eta + \mathsf{E}_x\Big[e^{-r\tau_\eta} g(t+\tau_\eta, X^{[n,\nu]}_{\tau_\eta}) + \int_0^{\tau_\eta} e^{-rs} h(t+s, X^{[n,\nu]}_s)\,\mathrm{d}s\\
&\qquad + \int_{[0,\tau_\eta]} e^{-rs} f(t+s, X^{[n,\nu]}_s)\circ\mathrm{d}\nu_s\Big] \qquad (3.124)\\
&= \eta + \mathcal{J}_{t,x}\,(n,\nu,\tau_\eta).
\end{aligned}$$

By arbitrariness of $(n,\nu)\in\mathcal{A}_t$ and sub-optimality of τ_η we have $u(t,x) \leq \eta + \underline{v}(t,x)$ by definition of lower value. Letting $\eta \to 0$ we get $u(t,x) \leq \underline{v}(t,x)$.

Next we prove $u \geq \overline{v}$. Since $\frac{1}{\delta}(g-u^{\varepsilon,\delta})^+ \geq 0$ that term can de dropped from (3.122) to obtain a lower bound for $u^{\varepsilon,\delta}$. We let $\delta \to 0$ in (3.122) along the sequence constructed in (3.121) while keeping ε fixed. As above, dominated convergence applies, and thanks to $u^\varepsilon \geq g$ (Lemma 3.27) we obtain

$$\begin{aligned} u^\varepsilon(t,x) \geq \mathsf{E}_x\Big[& e^{-r(\tau\wedge\rho_m)} g(t+\tau\wedge\rho_m, X^{[n,\nu]}_{\tau\wedge\rho_m}) \\ & + \int_0^{\tau\wedge\rho_m} e^{-rs}\big[h - \psi_\varepsilon\big(|\nabla u^\varepsilon|_d^2 - f^2\big)\big]\big(t+s, X^{[n,\nu]}_{s-}\big)\,\mathrm{d}s \\ & - \int_0^{\tau\wedge\rho_m} e^{-rs}\langle \nabla u^\varepsilon(t+s, X^{[n,\nu]}_{s-}), n_s\rangle \mathrm{d}\nu^c_s \\ & - \sum_{0\leq s\leq \tau\wedge\rho_m} e^{-rs}\int_0^{\Delta\nu_s} \langle \nabla u^\varepsilon(t+s, X^{[n,\nu]}_{s-} + n_s\lambda), n_s\rangle\,\mathrm{d}\lambda\Big]. \end{aligned} \tag{3.125}$$

We can now choose a control pair $(n,\nu) = (n^\varepsilon, \nu^\varepsilon)$ defined as in (3.99) but with $u^{\varepsilon,\delta}$ therein replaced by u^ε. Although $\nabla u^\varepsilon(t,\cdot)$ is not Lipschitz, it can be shown by standard localisation procedure and the use of [65, Thm. 1] that the associated controlled SDE admits a unique, non-exploding, strong solution $X^\varepsilon = X^{[n^\varepsilon,\nu^\varepsilon]}$ on $[0, T-t]$ (the proof is given in Appendix B.6 for completeness).

By construction, the pair $(n^\varepsilon, \nu^\varepsilon)$ satisfies

$$\begin{aligned} -\big\langle n^\varepsilon_s \dot{\nu}^\varepsilon_s, \nabla u^\varepsilon(t+s, X^\varepsilon_s)\big\rangle & - \psi_\varepsilon\big(\mathcal{H}(\nabla u^\varepsilon) - f^2\big)(t+s, X^\varepsilon_s) \\ & = H^\varepsilon(t+s, X^\varepsilon_s, n^\varepsilon_s \dot{\nu}^\varepsilon_s), \end{aligned} \tag{3.126}$$

with H^ε as in (3.91). Then, from (3.125) we obtain

$$\begin{aligned} u^\varepsilon(t,x) \geq \mathsf{E}_x\Big[& e^{-r(\tau\wedge\rho_m)} g(t+\tau\wedge\rho_m, X^\varepsilon_{\tau\wedge\rho_m}) \\ & + \int_0^{\tau\wedge\rho_m} e^{-rs}\big(h + H^\varepsilon(\cdot, n^\varepsilon_s\dot{\nu}^\varepsilon_s)\big)(t+s, X^\varepsilon_s)\mathrm{d}s\Big]. \end{aligned}$$

Taking $p = f(t+s, X^\varepsilon_s)n^\varepsilon_s$ in the Hamiltonian, letting $m \to \infty$ and using Fatou's lemma, we obtain

$$\begin{aligned} u^\varepsilon(t,x) \geq \mathsf{E}_x\Big[& e^{-r\tau} g\big(t+\tau, X^\varepsilon_\tau\big) + \int_0^\tau e^{-rs} h(t+s, X^\varepsilon_s)\mathrm{d}s \\ & + \int_{[0,\tau]} e^{-rs} f(t+s, X^\varepsilon_s) \circ \mathrm{d}\nu^\varepsilon_s\Big], \end{aligned}$$

where we notice that ν^ε is absolutely continuous so that the final integral is obvious.

By arbitrariness of τ we can take supremum over all stopping times. Since $(n^\varepsilon,\nu^\varepsilon)\in\mathcal{A}_t$, we can also take infimum over all admissible controls and continue with the same direction of inequalities. That is, $u^\varepsilon(t,x)\geq\overline{v}(t,x)$. Finally, letting $\varepsilon\to 0$ we obtain $u(t,x)\geq\overline{v}(t,x)$, as needed. Since $u\leq\underline{v}$ was proven above, we conclude $u=v=\overline{v}=\underline{v}$.

Next we prove optimality of τ_*. From (3.124) and the fact that $u=v$ we deduce that τ_η is η-optimal for the stopper. For an arbitrary $(n,\nu)\in\mathcal{A}_t$, letting $(\eta_m)_{m\in\mathbb{N}}$ with $\eta_m\downarrow 0$, the sequence $(\tau_{\eta_m})_{m\in\mathbb{N}}$ is a non-decreasing sequence P_x-a.s. We introduce an event and its complement:

$$
\begin{aligned}
B &:= \big\{\omega\in\Omega:\ \tau_{\eta_m}(\omega)<\tau_*(\omega),\forall m\in\mathbb{N}\big\}\quad\text{and}\\
B^c &:= \big\{\omega\in\Omega:\exists\, m\in\mathbb{N}\text{ such that }\tau_{\eta_k}(\omega)=\tau_*(\omega)\,\forall k\geq m\big\}.
\end{aligned}
$$

Thus, we have

$$
\lim_{m\to\infty}X^{[n,\nu]}_{\tau_{\eta_m}}(\omega)=\mathbb{1}_B(\omega)X^{[n,\nu]}_{\tau_0-}(\omega)+\mathbb{1}_{B^c}(\omega)X^{[n,\nu]}_{\tau_*}(\omega),
$$

where

$$
\tau_0=\inf\{s\geq 0\,|\,u(t+s,X^{[n,\nu]}_{s-})=g(t+s,X^{[n,\nu]}_{s-})\}\wedge(T-t).
$$

Letting $m\to\infty$ and applying dominated convergence to (3.124) along the sequence τ_{η_m} gives

$$
\begin{aligned}
u(t,x)\leq\mathsf{E}_x\bigg[&\mathbb{1}_B\bigg(e^{-r\tau_0}g\big(t+\tau_0,X^{[n,\nu]}_{\tau_0-}\big)+\int_0^{\tau_0}e^{-rs}h\big(t+s,X^{[n,\nu]}_s\big)\,\mathrm{d}s\\
&\qquad+\int_{[0,\tau_0)}e^{-rs}f\big(t+s,X^{[n,\nu]}_s\big)\circ\mathrm{d}\nu_s\bigg)\\
&+\mathbb{1}_{B^c}\bigg(e^{-r\tau_*}g\big(t+\tau_*,X^{[n,\nu]}_{\tau_*}\big)+\int_0^{\tau_*}e^{-rs}h\big(t+s,X^{[n,\nu]}_s\big)\,\mathrm{d}s\\
&\qquad+\int_{[0,\tau_*]}e^{-rs}f\big(t+s,X^{[n,\nu]}_s\big)\circ\mathrm{d}\nu_s\bigg)\bigg].
\end{aligned}
$$

The above equation holds for any $(n,\nu)\in\mathcal{A}_t$. We now take the infimum over all pairs $(n,\nu)\in\mathcal{A}_t$. Recalling Lemma 3.32, there is no loss of generality in restricting such infimum to the class of controls such that $\Delta\nu_{\tau_*}=0$. Since $B^c\subseteq\{\Delta\nu_{\tau_*}\neq 0\}$, then $\mathsf{P}(B^c)=0$ and it is not hard to check that $\tau_0=\tau_*$ and $X^{[n,\nu]}_{\tau_0-}=X^{[n,\nu]}_{\tau_*}$, P_x-a.s.

Therefore

$$u(t,x) \le \inf_{(n,\nu)\in\mathcal{A}_t} \mathsf{E}_x\Big[e^{-r\tau_*} g\big(t+\tau_*, X^{[n,\nu]}_{\tau_*}\big) + \int_0^{\tau_*} e^{-rs} h\big(t+s, X^{[n,\nu]}_s\big)\,\mathrm{d}s + \int_{[0,\tau_*]} e^{-rs} f\big(t+s, X^{[n,\nu]}_s\big)\circ \mathrm{d}\nu_s\Big] \le v(t,x).$$

Since $u = v$ then τ_* is optimal for the stopper.

It remains to prove that v is indeed the maximal solution of Problem A. Let w be another solution of Problem A. The same argument as in the proof of [27, Thm. 4.1, Ch. VIII] can be adapted to our proof. Consider a family of mollifiers $(\zeta_k)_{k\in\mathbb{N}} \subset C^\infty_c(\mathbb{R}^{d+1}_{0,T})$ and let $(w_k)_{k\in\mathbb{N}} \subset C^\infty_c(\mathbb{R}^{d+1}_{0,T})$ be the mollified family such that $w_k \to w$ and $\nabla w_k \to \nabla w$ uniformly on compact sets, with $\partial_t w_k \to \partial_t w$ and $D^2 w_k \to D^2 w$ strongly in $L^p_{\ell oc}(\mathbb{R}^{d+1}_{0,T})$ for all $p\in[1,\infty)$, as $k\to\infty$. For notational simplicity, denote the operator $(\partial_t + \mathcal{L} - r)$ by $\tilde{\mathcal{L}}$. We have that

$$(\tilde{\mathcal{L}} w * \zeta_k)(t_0,x_0) = \int_{\mathbb{R}^{d+1}_{0,T}} \Big(w_t(t,x) + \sum_{i,j=1}^d a_{ij}(x) w_{x_i x_j}(t,x) + \sum_{i=1}^d b_i(x) w_{x_i}(t,x) - r w(t,x)\Big)\zeta_k(t_0-t, x_0-x)\,\mathrm{d}t\mathrm{d}x$$

and

$$(\tilde{\mathcal{L}} w_k)(t_0,x_0) = (\partial_t + \mathcal{L}(x_0) - r)\Big(\int_{\mathbb{R}^{d+1}_{0,T}} w(t,x)\zeta_k(t_0-t,x_0-x)\,\mathrm{d}t\mathrm{d}x\Big).$$

Thanks to the properties on the derivatives of a convolution we have

$$(\tilde{\mathcal{L}} w_k)(t_0,x_0) = \int_{\mathbb{R}^{d+1}_{0,T}} \Big(w_t(t,x) + \sum_{i,j=1}^d a_{ij}(x_0) w_{x_i x_j}(t,x) + \sum_{i=1}^d b_i(x_0) w_{x_i}(t,x) - r w(t,x)\Big)\zeta_k(t_0-t, x_0-x)\,\mathrm{d}t\mathrm{d}x.$$

Thus, we have

$$
\begin{aligned}
&|\,(\tilde{\mathcal{L}}w * \zeta_k)(t_0, x_0) - (\tilde{\mathcal{L}}w_k)(t_0, x_0)| \\
&\quad = \Big| \int_{\mathbb{R}^{d+1}_{0,T}} \Big(\sum_{i,j=1}^{d} (a_{ij}(x) - a_{ij}(x_0)) w_{x_i x_j}(t, x) \\
&\qquad + \sum_{i=1}^{d} (b_i(x) - b_i(x_0)) w_{x_i}(t, x) \Big) \zeta_k(t_0 - t, x_0 - x)\, \mathrm{d}t \mathrm{d}x \Big|.
\end{aligned}
$$

Since first and second order derivatives of w belong to $L^p_{\ell oc}(\mathbb{R}^{d+1}_{0,T})$, by Hölder's inequality and the fact that a and b are continuous we have that for all fixed m

$$
\lim_{k\to\infty} \sup_{(t,x)\in\overline{\mathcal{O}}_m} |(\tilde{\mathcal{L}}w * \zeta_k)(t, x) - (\tilde{\mathcal{L}}w_k)(t, x)| = 0. \tag{3.127}
$$

For $\eta > 0$, we set

$$
\mathcal{C}^\eta_w = \{(t, x) \in \mathbb{R}^{d+1}_{0,T} : w(t, x) > g(t, x) + \eta\};
$$

it still holds that

$$
\lim_{k\to\infty} \sup_{(t,x)\in\overline{\mathcal{O}}_m\cap\mathcal{C}^\eta_w} |(\tilde{\mathcal{L}}w * \zeta_k)(t, x) - (\tilde{\mathcal{L}}w_k)(t, x)| = 0.
$$

Since w is a solution of Problem A, we have that $(\tilde{\mathcal{L}}w + h)(t, x) \geq 0$ for almost every $(t, x) \in \mathcal{C}^\eta_w$, therefore

$$
((\tilde{\mathcal{L}}w + h) * \zeta_k)(t, x) \geq 0 \tag{3.128}
$$

for all $(t, x) \in \mathcal{C}^\eta_w$, or in other terms $(\tilde{\mathcal{L}}w * \zeta_k)(t, x) \geq -(h * \zeta_k)(t, x)$ for all $(t, x) \in \mathcal{C}^\eta_w$.

For fixed $\eta > 0$ and $m \in \mathbb{N}$, we define

$$
N_{k,m} \coloneqq \inf_{(t,x)\in\overline{\mathcal{O}_m\cap\mathcal{C}^\eta_w}} |(\tilde{\mathcal{L}}w_k)(t, x) + h(t, x)|.
$$

By the convergence of (3.127) and thanks to (3.128) we have

$$\liminf_{k\to\infty} N_{k,m} \geq 0. \tag{3.129}$$

Pick an arbitrary pair $(n,\nu)\in\mathcal{A}_t$ and denote

$$\zeta_\eta := \inf\{s\geq 0 | w(t+s, X_s^{[n,\nu]}) \leq g(t+s, X_s^{[n,\nu]}) + \eta\} \wedge (T-t).$$

By an application of Itô's formula to $e^{-rs}w_k(t+s, X_s^{[n,\nu]})$ up to the stopping time $\zeta_\eta\wedge\rho_m$, we obtain

$$\begin{aligned} w_k(t,x) = \mathsf{E}_x\Bigg[& e^{-r(\zeta_\eta\wedge\rho_m)} w_k(t+\zeta_\eta\wedge\rho_m, X^{[n,\nu]}_{\zeta_\eta\wedge\rho_m}) \\ & - \int_0^{\zeta_\eta\wedge\rho_m} e^{-rs}(\partial_t w_k + \mathcal{L}w_k - rw_k)(t+s, X_s^{[n,\nu]})\mathrm{d}s \\ & - \int_0^{\zeta_\eta\wedge\rho_m} e^{-rs}\langle \nabla w_k(t+s, X_{s-}^{[n,\nu]}), n_s\rangle\, \mathrm{d}\nu_s^c \\ & - \sum_{s\leq\zeta_\eta\wedge\rho_m} e^{-rs}\int_0^{\Delta\nu_s}\langle\nabla w_k(t+s, X_{s-}^{[n,\nu]}+\lambda n_s), n_s\rangle\,\mathrm{d}\lambda\Bigg]. \end{aligned}$$

Letting $k\to\infty$, dominated convergence (up to possibly selecting a subsequence) and reverse Fatou's lemma (justified by (3.129)) allow us to pass the limit under expectation. Then, exploiting the uniform convergence of $(w_k, \nabla w_k)$ to $(w, \nabla w)$ on $\overline{\mathcal{O}_m\cap\mathcal{C}_w^\eta}$, (3.129), the definition of $\zeta_\eta\wedge\rho_m$ and the fact that $|\nabla w|_d\leq f$ we have

$$\begin{aligned} w(t,x) \leq\ \eta + \mathsf{E}_x\Bigg[& e^{-r\zeta_\eta} g(t+\zeta_\eta, X^{[n,\nu]}_{\zeta_\eta})\mathbb{1}_{\{\zeta_\eta\leq\rho_m\}} + e^{-r\rho_m} w(t+\rho_m, X^{[n,\nu]}_{\rho_m})\mathbb{1}_{\{\zeta_\eta>\rho_m\}} \\ & + \int_0^{\zeta_\eta\wedge\rho_m} e^{-rs} h(t+s, X_s^{[n,\nu]})\mathrm{d}s \\ & + \int_0^{\zeta_\eta\wedge\rho_m} e^{-rs} f(t+s, X_s^{[n,\nu]}) \circ \mathrm{d}\nu_s\Bigg]. \end{aligned}$$

Finally, letting $m\to\infty$, the same arguments that lead to (3.124) give $w(t,x)\leq \eta + \mathcal{J}_{t,x}(n,\nu,\zeta_\eta)$. Hence, $w(t,x)\leq\eta+v(t,x)$ and letting $\eta\to 0$ we conclude. □

Remark 3.34: It is worth noticing that the proof above can be repeated verbatim if we replace $\mathcal{A}_t$ with $\mathcal{A}_t^\circ$ everywhere. Thus we conclude that

$$v(t,x) = \inf_{(n,\nu)\in\mathcal{A}_t^\circ} \sup_{\tau\in\mathcal{T}_t} \mathcal{J}_{t,x}(n,\nu,\tau) = \sup_{\tau\in\mathcal{T}_t} \inf_{(n,\nu)\in\mathcal{A}_t^\circ} \mathcal{J}_{t,x}(n,\nu,\tau).$$

That is, the game with absolutely continuous controls admits the same value as the game with singular controls. It is however expected, but it will not be proven here, that an optimal control cannot be found in $\mathcal{A}_t^\circ$ whereas it should be possible to find one in $\mathcal{A}_t$ in some cases. ■

Remark 3.35 (η-optimal control strategies)**:** Theorem 3.31 proves the uniform convergence of a subsequence u^{ε_i} to the value function of the game u (see (3.125)). This means that for any $\eta > 0$, there exists a k such that for all $i \geq k$

$$|u^{\varepsilon_i}(t,x) - u(t,x)| < \eta.$$

From the proof of the theorem, we have the explicit η-optimal strategy for the controller in the game with value u^{ε_i}, i.e., the control pair $(n,\nu) = (n^{\varepsilon_i}, \nu^{\varepsilon_i})$ defined as in (3.99) but with $u^{\epsilon,\delta}$ therein replaced by u^{ε_i}. ■

Chapter 4

Degenerate Cases

This chapter presents extensions of our analysis to two degenerate cases of ZSG. In Section 4.1, we study the case where the controller is allowed to use controls in selected directions of the state-space. In Section 4.2, we study the case where the underlying process has a degenerate diffusion.

4.1 A Restriction on the Controller

This chapter presents a similar game to the one studied in Chapter 3 but in which the controller has a constraint on the control directions of the underlying process. Under a new set of assumptions, we are able to prove through an approximation procedure that the game admits a value and we provide an optimal strategy for the stopper. The reason why the theory of Chapter 3 cannot be applied directly can be found in Remark 4.5 after the presentation of the problem.

4.1.1 The Problem

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathsf{P})$ be a stochastic basis on which an adapted d'-dimensional Brownian motion $(W_t)_t$ is defined. Let T be the terminal time, and $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^{d'}$ be measurable functions, with $d \leq d'$. Denote

$$\mathcal{T}_t := \{\tau | \tau \text{ is a stopping time}, \tau \leq T - t\}.$$

The difference with the game in Chapter 3 is that now the control affects only some coordinates of the underlying process which we assume for simplicity of exposition to be the first $\hbar$ coordinates ($\hbar < d$) without loss of generality. We define by $\mathcal{A}_t^\hbar$ the class of admissible controls as

$$\mathcal{A}_t^\hbar := \left\{ (n,\nu) \middle| \begin{array}{l} (n_s)_{s\in[0,\infty)} \text{ is progressively measurable, } \mathbb{R}^\hbar \text{ valued,} \\ \text{with } |n_s|_\hbar = 1,\ \mathsf{P}\text{-a.s. } \forall s \in [0,\infty); \\ (\nu_s)_{s\in[0,\infty)} \text{ is } \mathbb{F}\text{-adapted, real valued, non-decreasing and} \\ \text{right-continuous with } \nu_{0-} = 0,\ \mathsf{P}\text{-a.s., and } \mathsf{E}[|\nu_{T-t}|^2] < \infty \end{array} \right\}.$$

Moreover, we indicate by $\nabla^\hbar f$ and $\nabla^{d-\hbar} f$ the gradient of a function f truncated to the first $\hbar$ coordinates and last $d-\hbar$ coordinates, respectively.

Consider the following d-dimensional controlled stochastic differential where we will use the notation $X_s^{[n,\nu]}$ to underline that the process is controlled by $(n,\nu) \in \mathcal{A}_t^\hbar$:

$$\begin{aligned} \mathrm{d}X_{i,t}^{[n,\nu]} &= b_i(X_s^{[n,\nu]})\,\mathrm{d}s + \sum_{j=1}^{d'} \sigma_{ij}(X_s^{[n,\nu]})\mathrm{d}W_s^j + n_{i,s}\,\mathrm{d}\nu_s, \qquad \text{for } 1 \le i \le \hbar, \\ \mathrm{d}X_{i,t}^{[n,\nu]} &= b_i(X_s^{[n,\nu]})\,\mathrm{d}s + \sum_{j=1}^{d'} \sigma_{ij}(X_s^{[n,\nu]})\mathrm{d}W_s^j, \qquad \text{for } \hbar+1 \le i \le d. \end{aligned}$$

Sometimes we will use $X_t^{x,[n,\nu]}$ to keep track of the fact that the process is starting in x at time 0.

Similar to Chapter 3, the controller pays a reward to the stopper when the game ends and the two players are optimising the expectation of the reward in which the controller now is minimising over the class $\mathcal{A}_t^\hbar$. Here, we recall explicitly the payoff from Chapter 3 where now the action cost f does not depend on the spatial coordinates:

$$\begin{aligned} \mathcal{J}_{t,x}(n,\nu,\tau) = \mathsf{E}_x\Big[& e^{-r\tau} g(t+\tau, X_\tau^{[n,\nu]}) + \int_0^\tau e^{-rs} h(t+s, X_s^{[n,\nu]})\,\mathrm{d}s \\ & + \int_{[0,\tau]} e^{-rs} f(t+s) \circ \mathrm{d}\nu_s \Big]. \end{aligned} \tag{4.1}$$

Since the function f depends only on time, then the integral defined in (4.1) is now consistent with the Riemann-Stieltjes integral, i.e.,

$$\int_{[0,\tau]} e^{-rs} f(t+s) \circ \mathrm{d}\nu_s = \int_0^\tau e^{-rs} f(t+s)\,\mathrm{d}\nu_s^c + \sum_{0\le s\le\tau} e^{-rs} f(t+s)\Delta\nu_s \tag{4.2}$$
$$= \int_0^\tau e^{-rs} f(t+s)\,\mathrm{d}\nu_s.$$

We define

$$\underline{v}(t,x) \coloneqq \sup_{\tau\in\mathcal{T}_t} \inf_{(n,\nu)\in\mathcal{A}_t^\hbar} \mathcal{J}_{t,x}(n,\nu,\tau) \text{ and } \overline{v}(t,x) \coloneqq \inf_{(n,\nu)\in\mathcal{A}_t^\hbar} \sup_{\tau\in\mathcal{T}_t} \mathcal{J}_{t,x}(n,\nu,\tau), \tag{4.3}$$

so that $\underline{v}(t,x) \le \overline{v}(t,x)$. If the equality holds then we say that the game admits a value:

$$v(t,x) \coloneqq \underline{v}(t,x) = \overline{v}(t,x). \tag{4.4}$$

We recall that $\mathcal{L}$ denotes the infinitesimal generator of the uncontrolled process $X^{[e_1,0]}$ (where e_1 is the unit vector in $\mathbb{R}^\hbar$ with 1 in the first entry) and it reads

$$(\mathcal{L}\varphi)(x) = \frac{1}{2}\mathrm{tr}\left(a(x)D^2\varphi(x)\right) + \langle b(x), \nabla\varphi(x)\rangle,$$

with $a(x) \coloneqq (\sigma\sigma^\top)(x)$.

Next we give assumptions under which we obtain our main result (Theorem 4.3).

Assumption 4.1 (Controlled SDE) *The functions b and σ are Lipschitz with constant D_1 and continuously differentiable on $\mathbb{R}^d$ and σ is such that*

$$\sigma^i(x) = \sigma^i(x_i), \qquad \textit{for } i = 1,\dots d,$$

where $\sigma_i = (\sigma_{i1},\dots\sigma_{id'})$. Recalling $a = \sigma\sigma^\top$, for any bounded set $B\subset\mathbb{R}^d$ there is $\theta_B > 0$ such that $a(\cdot)$ is locally elliptic

$$\langle\zeta, a(x)\zeta\rangle \ge \theta_B|\zeta|_d^2 \qquad \textit{for any } \zeta\in\mathbb{R}^d \textit{ and all } x\in\overline{B}.$$

Notice that the Lipschitz continuity of b and σ implies that there exists D_2

such that

$$|b(x)|_d + |\sigma(x)|_{d\times d'} \leq D_2(1 + |x|_d), \qquad \text{for all } x \in \mathbb{R}^d. \tag{4.5}$$

Assumption 4.2 (Functions f, g, h) *For the functions $f, g, h : \mathbb{R}^{d+1}_{0,T} \to [0, \infty)$ the following hold:*

(i) $g \in C^{1,2,\alpha}_{\ell oc}(\mathbb{R}^{d+1}_{0,T})$ and $h \in C^{0,1,\alpha}_{\ell oc}(\mathbb{R}^{d+1}_{0,T})$ for some $\alpha \in (0, 1)$;

(ii) f is non-increasing, positive and f^2 is differentiable;

(iii) there is $K_0 \in (0, \infty)$ such that for all $0 \leq s < t \leq T$ and all $x \in \mathbb{R}^{d+1}_{0,T}$

$$h(t, x) - h(s, x) \leq K_0(t - s) \quad \text{and} \quad g(t, x) - g(s, x) \leq K_0(t - s);$$

(iv) there is $K_1 \in (0, \infty)$ such that

$$|g(t, x) - g(t, y)| + |h(t, x) - h(t, y)| \leq K_1|x - y|_d, \quad \text{for } x, y \in \mathbb{R}^d; \tag{4.6}$$

(v) f and g are such that

$$|\nabla^{\hbar} g(t, x)|_{\hbar} \leq f(t), \quad \text{for all } (t, x) \in \mathbb{R}^{d+1}_{0,T}; \tag{4.7}$$

(vi) there is $K_2 \in (0, \infty)$ such that

$$(h + \partial_t g + \mathcal{L}g - rg)(t, x) \geq -K_2, \quad \text{for all } (t, x) \in \mathbb{R}^{d+1}_{0,T}.$$

Notice that (4.6) implies that there exists $K_3 \in (0, \infty)$ such that

$$0 \leq |g(t, x)| + |h(t, x)| \leq K_3(1 + |x|_d), \quad \text{for } (t, x) \in \mathbb{R}^{d+1}_{0,T}. \tag{4.8}$$

A comment on these conditions can be found in Chapter 3 below Assumption 3.5.

We can now state the main result of the chapter.

Theorem 4.3 *The game described above admits a value v (i.e., (4.4) holds). Moreover, for any given $(t, x) \in \mathbb{R}^{d+1}_{0,T}$ and any admissible control $(n, \nu) \in \mathcal{A}^{\hbar}_t$, the* *stopping time* $\tau_* = \tau_*(t, x; n, \nu) \in \mathcal{T}_t$ *defined under* P_x *as*

$$\tau_* := \inf \left\{ s \geq 0 \,\middle|\, v(t+s, X_s^{[n,\nu]}) = g(t+s, X_s^{[n,\nu]}) \right\} \wedge (T - t) \tag{4.9}$$

is optimal for the stopper.

Remark 4.4: In this framework, we are not able to connect the value function with a variational inequality as we did in Chapter 3. ■

4.1.2 Approximated Problem

The theory developed in Chapter 3 does not apply directly to this new type of game. Results proved by probabilistic arguments apply without changes to this class of games, but the results proved by analytic arguments fail and we need an approximation procedure in order to use them.

We explain in the next remark the main reason why some proofs stop to hold.

Remark 4.5: The theory in Chapter 3 does not apply directly because some of the results therein (such as, Propositions 3.17 and 3.22, and Lemma 3.29) cease to hold. Under the setting of Section 4.1, the heuristic argument and the result obtained later through the approximated procedure lead to the following condition: the value function of the game v should satisfy $|\nabla^\hbar v|_\hbar \leq f$. It means that the penalisation term introduced in the corresponding PDE is $\psi_\varepsilon(\mathcal{H}(\nabla v) - f^2)$ with $\mathcal{H}(p) = |\bar{p}|_\hbar^2$ and $\bar{p}$ is the vector of the first $\hbar$ coordinates of p. Repeating the proof of Proposition 3.17, we get from the corresponding (3.53) the following

$$-2\langle \nabla w^n, \nabla(\mathcal{H}(\nabla u^n) - f_m^2)\rangle \leq -2\lambda |\nabla^\hbar u^n|_\hbar^2 + 2|\nabla u|_d |\nabla f_m^2|_d + \tilde{R}_n.$$

This inequality is not sufficient to conclude the proof of the analogue of Proposition 3.17. Indeed, the first term on the right-hand side above does not bound the terms as in Proposition 3.17. In a similar way Proposition 3.22 does not hold. The other result which fails is Lemma 3.29, indeed (3.111) in the proof of Lemma 3.29 does not lead to a bound on $|D^2 w^n|_{d\times d}^2$ but it leads only to a bound on a strict subset of the partial derivatives included in $D^2 w^n$ and thus the result does not hold. ■

4.1.3 Approximation Procedure

We introduce a sequence of approximated games to which the results from Chapter 3 apply. These games are indexed by a parameter γ and we obtain a family of value functions associated to $\gamma \in (0,1)$. The existence of the value of the original game is found through a convergent subsequence of this family.

Fix a $\gamma \in (0,1)$. For a given pair $(n,\nu) \in \mathcal{A}_t^d$ we consider the following (controlled) stochastic differential equation:

$$\begin{aligned} \mathrm{d}X_{i,t}^{[n,\nu],\gamma} &= b_i(X_s^{[n,\nu],\gamma})\,\mathrm{d}s + \sum_{j=1}^{d'} \sigma_{ij}(X_{i,s}^{[n,\nu],\gamma})\mathrm{d}W_s^j + n_{i,s}\,\mathrm{d}\nu_s, \text{ for } 1 \le i \le \hbar, \\ \mathrm{d}X_{i,t}^{[n,\nu],\gamma} &= b_i(X_s^{[n,\nu],\gamma})\,\mathrm{d}s + \sum_{j=1}^{d'} \sigma_{ij}(X_{i,s}^{[n,\nu],\gamma})\mathrm{d}W_s^j + \gamma n_{i,s}\,\mathrm{d}\nu_s, \text{ for } \hbar+1 \le i \le d. \end{aligned} \tag{4.10}$$

The term γ is a weight applied to the last $d-\hbar$ coordinates.

Let $p = (\bar{p}, \tilde{p}) \in \mathbb{R}^\hbar \times \mathbb{R}^{d-\hbar}$. For $\gamma \in (0,1]$, we introduce a function $\mathcal{H}^\gamma(p) : \mathbb{R}^d \to \mathbb{R}$ defined as $\mathcal{H}^\gamma(p) := |\bar{p}|_\hbar^2 + \gamma|\tilde{p}|_{d-\hbar}^2$ (it is clear that $\mathcal{H}^1$ is equal to $\mathcal{H}$ used in Chapter 3). The gradient is $\nabla\mathcal{H}^\gamma(p) = 2\bar{p} + 2\gamma\tilde{p} = (2p_1, \ldots, 2p_\hbar, 2\gamma p_{\hbar+1}, \ldots, 2\gamma p_d)$ and the Hessian matrix $D^2\mathcal{H}^\gamma(p)$ is a diagonal matrix with the first $\hbar$ entries equal to 2 and the last $d-\hbar$ entries equal to 2γ.

We introduce an approximation of f as

$$f^\gamma(t) := \sqrt{f^2(t) + \gamma K_1^2} \quad \text{for } t \in [0,T], \tag{4.11}$$

where K_1 comes from (4.6). By construction $f^\gamma \to f$ uniformly on $[0,T]$ as $\gamma \to 0$. Notice that (4.7) and (4.11) imply

$$\begin{aligned} \mathcal{H}^\gamma(\nabla g(t,x)) &= |\nabla^\hbar g(t,x)|_\hbar^2 + \gamma|\nabla^{d-\hbar} g(t,x)|_{d-\hbar}^2 \\ &\le f^2(t) + \gamma K_1^2 \\ &= (f^\gamma(t))^2, \end{aligned}$$

for all $(t,x) \in \mathbb{R}_{0,T}^{d+1}$.

We consider a new payoff $\mathcal{J}_{t,x}^\gamma$ similar to the payoff defined in (4.1) with $X_t^{[n,\nu]}$

and f therein replaced by $X_t^{[n,\nu],\gamma}$ and f^γ, respectively, i.e.,

$$\mathcal{J}^{\gamma}_{t,x}(n,\nu,\tau) = \mathsf{E}_x\bigg[e^{-r\tau} g(t+\tau, X_\tau^{[n,\nu],\gamma}) + \int_0^\tau e^{-rs} h(t+s, X_s^{[n,\nu],\gamma})\,\mathrm{d}s \\ + \int_0^\tau e^{-rs} f^\gamma(t+s)\,\mathrm{d}\nu_s \bigg]. \tag{4.12}$$

We say that the game (4.12) admits a value if

$$u^\gamma(t,x) = \sup_{\tau\in\mathcal{T}_t} \inf_{(n,\nu)\in\mathcal{A}_t^d} \mathcal{J}^\gamma_{t,x}(n,\nu,\tau) = \inf_{(n,\nu)\in\mathcal{A}_t^d} \sup_{\tau\in\mathcal{T}_t} \mathcal{J}^\gamma_{t,x}(n,\nu,\tau). \tag{4.13}$$

The game described above satisfies Assumptions 3.4 and 3.5. The dynamics of $X^{[n,\nu],\gamma}$ has a weight γ in the last $d-\hbar$ coordinates which does not appear in the dynamics of (3.1). The γ parameter in the dynamics allows to recover the theory of [11]. Indeed, it leads to a different penalisation term $\psi_\epsilon(\mathcal{H}^\gamma(\nabla u) - (f_m^\gamma)^2)$ in (3.18) and a different Hamiltonian (3.19) that now reads

$$H_m^{\varepsilon,\gamma}(t,x,y) \coloneqq \sup_{p\in\mathbb{R}^d} \big\{ \langle \bar{y},\bar{p}\rangle + \gamma\langle \tilde{y},\tilde{p}\rangle - \psi_\varepsilon(\mathcal{H}^\gamma(p) - (f_m^\gamma(t,x))^2)\big\},$$

where $y=(\bar{y},\tilde{y})$ and $p=(\bar{p},\tilde{p})$ belong to $\mathbb{R}^\hbar\times\mathbb{R}^{d-\hbar}$. We notice that the first-order condition for $H_m^{\varepsilon,\gamma}$ is the same as the one for H_m^ε used in the proof of Proposition 3.9. Indeed we have

$$\begin{aligned} y_i &= \psi'_\varepsilon(\mathcal{H}^\gamma(p) - (f_m^\gamma(t,x))^2)\mathcal{H}_i(p) = \psi'_\varepsilon(\mathcal{H}^\gamma(p) - (f_m^\gamma(t,x))^2)2p_i \quad \text{for } i\le\hbar, \\ \gamma y_i &= \psi'_\varepsilon(\mathcal{H}^\gamma(p) - (f_m^\gamma(t,x))^2)\mathcal{H}_i(p) = \psi'_\varepsilon(\mathcal{H}^\gamma(p) - (f_m^\gamma(t,x))^2)2\gamma p_i \quad \text{for } i>\hbar, \end{aligned}$$

which can be expressed as $y = \psi'_\varepsilon(\mathcal{H}(p) - f_m^2(t,x))2p$ using vectorial notation. We now explain in detail the small changes that are required to adapt the results from Chapter 3 to the case with dynamics (4.10).

Propositions 3.9 and 3.11 hold with $\mathcal{H}^\gamma$ in place of $\mathcal{H}$. Even if $H_m^{\varepsilon,\gamma}$ is replaced by H_m^ε, they lead to the same first-order condition (as described above). Similarly, Propositions 3.25 and 3.26 hold by the same argument with

$$H^{\varepsilon,\gamma}(t,x,y) \coloneqq \sup_{p\in\mathbb{R}^d} \big\{ \langle \bar{y},\bar{p}\rangle + \gamma\langle \tilde{y},\tilde{p}\rangle - \psi_\varepsilon(\mathcal{H}^\gamma(p) - (f^\gamma(t,x))^2)\big\},$$

in place of H^ε and the extension of (3.93) to $H^{\varepsilon,\gamma}$, i.e., we have

$$\begin{aligned} H^{\varepsilon,\gamma}(t,x,y) &\geq \frac{\varepsilon}{2}(|\bar{y}|_\hbar^2 + \gamma|\tilde{y}|_{d-\hbar}^2) - \psi_\varepsilon\big(\mathcal{H}^\gamma(\tfrac{\varepsilon}{2}y) - (f^\gamma(t,x))^2\big) \\ &\geq \frac{\varepsilon}{2}(|\bar{y}|_\hbar^2 + \gamma|\tilde{y}|_{d-\hbar}^2) - \psi_\varepsilon\big(\mathcal{H}^\gamma(\tfrac{\varepsilon}{2}y)\big) \\ &= \frac{\varepsilon}{2}(|\bar{y}|_\hbar^2 + \gamma|\tilde{y}|_{d-\hbar}^2) - \psi_\varepsilon\big(\tfrac{\varepsilon^2}{4}(|\bar{y}|_\hbar^2 + \gamma|\tilde{y}|_{d-\hbar}^2)\big) \\ &\geq \frac{\varepsilon\gamma}{4}|y|_d^2. \end{aligned}$$

Thus the corresponding equation of (3.102) holds with a γ^{-1} multiplicative factor on the right-hand side of (3.102).

Proposition 3.17 holds with an adjustments in (3.53). That equation leads now to

$$-2\langle \nabla w^n, \nabla(\mathcal{H}^\gamma(\nabla u^n) - (f_m^\gamma)^2)\rangle \leq 2\lambda\mathcal{H}^\gamma(\nabla u^n) + 2|\nabla u|_d|\nabla(f_m^\gamma)^2|_d + \tilde{R}_n.$$

Since $\mathcal{H}^\gamma(\nabla u^n) \geq \gamma|\nabla u^n|_d^2$ then the equation corresponding to (3.54) is

$$0 \leq (C_1 - \lambda\gamma)|\nabla u|_d^2 + C_2 + \lambda r M_1 + R_n + \tilde{R}_n.$$

The proof continues with the same arguments and the γ factor is maintained until the end. It follows that the final constant M_3 depends on γ. The proof of Proposition 3.22 requires the same changes. In particular, (3.84) becomes

$$\begin{aligned} \xi\langle \nabla w^n, \nabla\big(\mathcal{H}^\gamma(\nabla u^n) - (f^\gamma)^2\big)\rangle &\geq \lambda\mathcal{H}^\gamma(\nabla u) - |\nabla u|_d^3|\nabla\xi|_d - \xi\hat{R}_n, \\ &\geq \lambda\gamma|\nabla u|_d^2 - |\nabla u|_d^3|\nabla\xi|_d - \xi\hat{R}_n, \end{aligned}$$

where $2|\nabla u|_d|\nabla(f^\gamma)^2|_d$ in (3.84) disappears because f^γ does not depend on the spatial variables and we can conclude the proof in the same way.

Lemma 3.29 still holds, but in this case we have that (3.111) holds in the following way:

$$\frac{1}{2}\big\langle D^2\mathcal{H}^\gamma(\nabla w^n)\nabla w_{x_i}^n, \nabla w_{x_j}^n\big\rangle \geq \gamma\big\langle \nabla w_{x_i}^n, \nabla w_{x_j}^n\big\rangle.$$

This inequality forces us to change other inequalities in order to obtain γ as a multiplicative factor. In particular, (3.112) becomes

$$\begin{aligned}&\left\langle a\nabla\xi,\nabla(\mathcal{H}^{\gamma}(\nabla w^n)-(f^{\gamma})^2)\right\rangle\\&\qquad\geq-\xi\frac{\theta\gamma}{4}|D^2w^n|^2_{d\times d}-\frac{16}{\theta\gamma}\bar{a}^2_m d^4C_0|\nabla w^n|^2_d\end{aligned}$$

and (3.119) becomes

$$\sum_{k=1}^{d}w^n_{x_k}\mathcal{L}_{x_k}w^n\leq\tfrac{\theta\gamma}{8}|D^2w^n|^2_{d\times d}+\frac{C_1}{\gamma}(N_1+1)^2.$$

The proof continues as in the proof of Proposition 3.17 and it leads to the equation corresponding to (3.115):

$$\left(1-\sqrt{\lambda\gamma^{-1}}A_{m+1}\right)\xi\psi_{\varepsilon}(\zeta_n)\leq C_6\sqrt{1+\lambda^{-1}+\kappa_{\delta,m}R_n}.$$

Choosing $\lambda=\frac{\gamma}{4A^2_{m+1}}$ we can conclude in the same way.

Finally, Theorem 3.33 holds with the same arguments using the same idea used for Propositions 3.9 and 3.11, i.e., the use of an optimal control that satisfies the first-order condition of the Hamiltonian $H^{\varepsilon,\gamma}$.

The remaining results of Chapter 3 do not need any changes.

Theorem 4.6 *The game described above admits a value u^{γ} (i.e., (4.13) holds). Moreover, for any given $(t,x)\in\mathbb{R}^{d+1}_{0,T}$ and any admissible control $(n,\nu)\in\mathcal{A}^d_t$, the stopping time $\tau^{\gamma}_*=\tau^{\gamma}_*(t,x;n,\nu)\in\mathcal{T}_t$ defined under P_x as*

$$\tau^{\gamma}_*:=\inf\left\{s\geq0|u^{\gamma}(t+s,X^{[n,\nu]}_s)=g(t+s,X^{[n,\nu]}_s)\right\}\wedge(T-t),$$

is optimal for the stopper.

Following the theory of Chapter 3, we have that the function u^{γ} from Theorem 4.6 has the following properties.

Lemma 4.7 *For any $\gamma\in(0,1)$, the function u^{γ} belongs to $(C^{0,1,\beta}_{\ell oc}\cap W^{1,2,p}_{\ell oc})(\mathbb{R}^{d+1}_{0,T})$ and*

$$0\leq u^{\gamma}(t,x)\leq K_4(1+|x|_d),$$

where $K_4 = K_4(T, D_2, K_3)$ is independent of γ and comes from Lemma 3.13 proved using the stronger assumption (4.8).

The next lemma proves that the controller has no advantage to use a control ν with an expectation greater than a suitable constant K_6. This result is crucial to prove the main theorem of this section (Theorem 4.12), but it has also an interest in its own right.

Lemma 4.8 *There exists a constant $K_6 = K_6(x; T, f(T), K_4, D_1)$ independent of γ such that*

$$u^\gamma(t,x) = \inf_{(n,\nu)\in\mathcal{A}^{d,opt}_{t,x}} \sup_{\tau\in\mathcal{T}_t} \mathcal{J}^\gamma_{t,x}(n,\nu,\tau) = \sup_{\tau\in\mathcal{T}_t} \inf_{(n,\nu)\in\mathcal{A}^{d,opt}_{t,x}} \mathcal{J}^\gamma_{t,x}(n,\nu,\tau),$$

where $\mathcal{A}^{d,opt}_{t,x} := \big\{(n,\nu)\in\mathcal{A}^d_t \big| \mathsf{E}_x[\nu_{T-t}] \le K_6\big\}$.

Proof. Let $(e_1, 0)\in\mathcal{A}^d_t$ be the null control and denote $X = X^{[e_1,0]}$. Following the idea of Lemma 3.13, we have

$$\begin{aligned}\sup_{\tau\in\mathcal{T}_t}\inf_{(n,\nu)\in\mathcal{A}^d_t} \mathcal{J}^\gamma_{t,x}(n,\nu,\tau) &\le \sup_{\tau\in\mathcal{T}_t}\mathcal{J}^\gamma_{t,x}(e_1,0,\tau)\\ &= \sup_{\tau\in\mathcal{T}_t}\mathsf{E}_x\left[e^{-r\tau}g(t+\tau,X_\tau)+\int_0^\tau e^{-rs}h(t+s,X_s)\,\mathrm{d}s\right]\\ &\le K_3(1+T)\mathsf{E}_x\left[\sup_{s\in[0,T]}(1+e^{-rs}|X_s|_d)\right] \qquad (4.14)\\ &\le K_4(1+|x|_d),\end{aligned}$$

where the second inequality is using the linear growth of g and h (see (4.8)), the third inequality is by standard estimates for SDEs with coefficients with linear growth ([48, Cor. 2.5.10]) and the constant $K_4 > 0$ depends only on T, D_2 and K_3 from (4.5) and (4.8), respectively. Therefore, the controller tries to minimise over the class

$$\mathcal{A}^{d,opt}_{t,x} := \left\{(n,\nu)\in\mathcal{A}^d_t \,\middle|\, \sup_{\tau\in\mathcal{T}_t}\mathcal{J}^\gamma_{t,x}(n,\nu,\tau) \le K_4(1+|x|_d)\right\}.$$

For $(n, \nu) \in \mathcal{A}^{d,opt}_{t,x}$ we have

$$
\begin{aligned}
\mathsf{E}_x\big[|\nu_{T-t}|\big] &= \mathsf{E}_x\left[\int_{[0,T-t]} \mathrm{d}\nu_s\right] \\
&\leq \mathsf{E}_x\left[\Big(\min_{s\in[0,T-t]} f^\gamma(t+s)\Big)^{-1}\int_{[0,T-t]} f^\gamma(t+s)\,\mathrm{d}\nu_s\right] \\
&\leq \frac{e^{rT}}{f^\gamma(T)}\mathsf{E}_x\bigg[e^{-r(T-t)}g(T, X_T^{[n,\nu],\gamma}) + \int_0^{T-t} e^{-rs}h(t+s, X_s^{[n,\nu],\gamma})\,\mathrm{d}s \\
&\qquad\qquad + \int_{[0,T-t]} e^{-rs} f^\gamma(t+s)\,\mathrm{d}\nu_s\bigg] \\
&\leq \frac{e^{rT}}{f(T)}\mathcal{J}^\gamma_{t,x}(n,\nu,T-t)
\end{aligned}
\tag{4.15}
$$

where the first inequality is using $\frac{f^\gamma(t+s)}{f^\gamma(t+s)}$ inside the integral and taking the minimum of the denominator; the second inequality is justified because g, h are non-negative and f^γ is non-increasing in time, we also multiplied by $e^{r(T-t)}e^{-r(T-t)}$ and $e^{rs}e^{-rs}$ the function g and inside the two integrals, respectively, and finally we collected $e^{-r(T-t)}$ and e^{rs} in front of the expectation taking the maximum on time of them, i.e., e^{rT}; the third inequality is because f^γ is decreasing in γ and thus $f^\gamma(T) \geq f(T)$.

Combining (4.15) with (4.14) and $(n, \nu) \in \mathcal{A}^{d,opt}_{t,x}$, we have

$$
\mathsf{E}_x\big[|\nu_{T-t}|\big] \leq \frac{e^{rT}K_4}{f(T)}(1+|x|_d) =: K_6,
$$

where $K_6 = K_6(x; T, f(T), K_4)$ is independent of γ. □

Remark 4.9: The result in Lemma 4.8 can be extended to $\underline{v}(t, x)$ and we obtain that

$$
\underline{v}(t,x) = \sup_{\tau\in\mathcal{T}_t}\inf_{(n,\nu)\in\mathcal{A}^{\hbar,opt}_{t,x}} \mathcal{J}_{t,x}(n,\nu,\tau) = \inf_{(n,\nu)\in\mathcal{A}^{\hbar,opt}_{t,x}}\sup_{\tau\in\mathcal{T}_t} \mathcal{J}_{t,x}(n,\nu,\tau),
$$

with $\mathcal{A}^{\hbar,opt}_{t,x} = \{(n,\nu) \in \mathcal{A}^{\hbar}_t | \mathsf{E}_x[\nu_{T-t}] \leq K_6\}$ and K_6 as in Lemma 4.8.

The proof follows the steps of the proof of Lemma 4.8. Indeed, (4.14) and (4.15) hold with: $f(t+s)$ and $\mathcal{J}_{t,x}(n,\nu,\tau)$ in place of $f^\gamma(t+s)$ and $\mathcal{J}^\gamma_{t,x}(n,\nu,\tau)$, respectively, and we can repeat the same steps therein which lead to the result. ■

Before we state the next result we recall the definition of local time at 0 of a process X from [57, Sec. IV.7, p. 212].

Definition 4.10 *The local time at* 0 *of a semimartingale* X *is defined as*

$$L_t^0(X) := A_t^0 - \sum_{0<s\le t} |X_s| - |X_{s-}| - \operatorname{sign}(X_{s-})\Delta X_s,$$

where A_t^0 *is the increasing process that satisfies*

$$|X_t| = |X_0| + \int_{0+}^t \operatorname{sign}(X_{s-})\,\mathrm{d}X_s + A_t^0.$$

Theorem 4.11 *Let* $t \in [0,T]$*. For all* $\tau \le T-t$ *stopping time and* $(n,\nu) \in \mathcal{A}_t^d$*, there exists* $(\bar{n},\bar{\nu}) \in \mathcal{A}_t^\hbar$ *such that*

$$\mathsf{E}_x\Big[|X_\tau^{[n,\nu],\gamma} - X_\tau^{[\bar{n},\bar{\nu}],0}|_d\Big] \le \gamma K_7 \mathsf{E}_x[\nu_{T-t}],$$

where $K_7 = K_7(D_1,d,T)$ *with* D_1 *from Assumption 4.1 is a constant independent of* γ*.*

Proof. For each pair $(n,\nu) \in \mathcal{A}_t^d$, with $n = (n^\hbar, n^{d-\hbar}) \in \mathbb{R}^\hbar \times \mathbb{R}^{d-\hbar}$, we can define a pair $(\bar{n},\bar{\nu}) \in \mathcal{A}_t^\hbar$ as

$$\begin{aligned} \bar{n}_{i,s} &:= \begin{cases} \frac{n_{i,s}}{|n_s^\hbar|_\hbar}, & |n_s^\hbar|_\hbar \ne 0, \\ \bar{e}_1, & |n_s^\hbar|_\hbar = 0, \end{cases} \quad \text{for } i = 1,\dots \hbar; \\ \bar{\nu}_s &:= \int_0^s |n_r^\hbar|_\hbar \,\mathrm{d}\nu_r; \end{aligned} \tag{4.16}$$

where $\bar{e}_1$ is the $\hbar$-dimensional unit vector with value 1 in the first coordinate. We have that $\int_0^s \bar{n}_{i,r}\mathrm{d}\bar{\nu}_r = \int_0^s n_{i,r}\mathrm{d}\nu_r$ for all $s \in [0,T-t]$. The vector process $(\bar{n}_s)_{s\in[0,\infty)}$ is progressively measurable. Indeed, take a measurable subset U of the $\hbar$-dimensional unit ball. We need to prove that for all $s \in [0,T-t]$, the set

$$U^{-1} := \big\{(r,\omega) \in [0,s]\times\Omega \,\big|\, \bar{n}_r(\omega) \in U\big\}$$

is a measurable set. Let $Z := \{(r,\omega) \in [0,s]\times\Omega : |n_r^\hbar(\omega)|_\hbar \ne 0\}$, we have that $U^{-1} = (U^{-1}\cap Z)\cup(U^{-1}\cap Z^c)$. On the set $U^{-1}\cap Z$, we have that $\bar{n}_s$ is the ratio of two progressively measurable processes with a denominator always different from 0 and thus the set is measurable. The other set $U^{-1} \cap Z^c$ is equal to $\emptyset$ if $\bar{e}_1 \notin U$ and equal to Z^c if $\bar{e}_1 \in U$; in both cases we have that $U^{-1} \cap Z^c$ is measurable because $\emptyset$ and Z^c are. Indeed, the latter is measurable because it is the pre-image of a closed set under a measurable function, in particular, it is the composition of $|\cdot|$ (which is a continuous function and thus also a measurable function) with n which is progressively measurable by assumption. We conclude that the set U is measurable as it is a finite union of measurable sets.

We consider two processes, $X^{[n,\nu],\gamma}$ and, $X^{[\bar{n},\bar{\nu}],0}$, whose dynamics follow for the coordinates $i = 1, \dots \hbar$

$$\begin{aligned} \mathrm{d}X^{[n,\nu],\gamma}_{i,s} &= b_i(X^{[n,\nu],\gamma}_s)\mathrm{d}s + \sum_{j=1}^{d'} \sigma_{ij}(X^{[n,\nu],\gamma}_{i,s})\mathrm{d}W^j_s + n_{i,s}\mathrm{d}\nu_s, \\ \mathrm{d}X^{[\bar{n},\bar{\nu}],0}_{i,s} &= b_i(X^{[\bar{n},\bar{\nu}],0}_s)\mathrm{d}s + \sum_{j=1}^{d'} \sigma_{ij}(X^{[\bar{n},\bar{\nu}],0}_{i,s})\mathrm{d}W^j_s + \bar{n}_{i,s}\mathrm{d}\bar{\nu}_s, \end{aligned} \tag{4.17}$$

and for $i = \hbar + 1, \dots d$

$$\begin{aligned} \mathrm{d}X^{[n,\nu],\gamma}_{i,s} &= b_i(X^{[n,\nu],\gamma}_s)\mathrm{d}s + \sum_{j=1}^{d'} \sigma_{ij}(X^{[n,\nu],\gamma}_{i,s})\mathrm{d}W^j_s + \gamma n_{i,s}\mathrm{d}\nu_s, \\ \mathrm{d}X^{[\bar{n},\bar{\nu}],0}_{i,s} &= b_i(X^{[\bar{n},\bar{\nu}],0}_s)\mathrm{d}s + \sum_{j=1}^{d'} \sigma_{ij}(X^{[\bar{n},\bar{\nu}],0}_{i,s})\mathrm{d}W^j_s. \end{aligned} \tag{4.18}$$

Let $\tau \in \mathcal{T}_t$ and denote $X^\gamma = X^{[n,\nu],\gamma}$ and $X^0 = X^{[\bar{n},\bar{\nu}],0}$ for simplicity. Define the exit time from the ball of radius R as

$$\tau_R := \inf\left\{s \geq 0 \middle| |X^\gamma_s|_d \vee |X^0_s|_d \geq R\right\}.$$

We denote the two processes $X^\gamma_{\cdot\wedge\tau\wedge\tau_R}$ and $X^0_{\cdot\wedge\tau\wedge\tau_R}$ by $X^{\gamma,R}_\cdot$ and $X^{0,R}_\cdot$, respectively; in a similar way, we denote the difference process, stopped at the time $\tau \wedge \tau_R$ as $J^{\gamma,R}_\cdot := X^{\gamma,R}_\cdot - X^{0,R}_\cdot$. The process $J^{\gamma,R}$ is a semimartingale. For any $s \in [0, T - t]$, by Meyer-Itô Formula for semimartingales (see [57, Thm. IV.70]) applied to the

i-th coordinate of $J^{\gamma,R}$, $J_i^{\gamma,R}$, we have

$$|J_{i,s}^{\gamma,R}| = \int_{0+}^{s\wedge\tau\wedge\tau_R} \operatorname{sign}(J_{i,\lambda-}^{\gamma,R})\,\mathrm{d}(J_{i,\lambda}^{\gamma,R}) + L^0_{s\wedge\tau\wedge\tau_R}(J_i^{\gamma,R}) + \sum_{0<\lambda\leq s\wedge\tau\wedge\tau_R} |J_{i,\lambda}^{\gamma,R}| - |J_{i,\lambda-}^{\gamma,R}| - \operatorname{sign}(J_{i,\lambda-}^{\gamma,R})\Delta(J_{i,\lambda}^{\gamma,R}) \tag{4.19}$$

for $i = 1,\dots d$, where $\operatorname{sign}(y) = -1$ for $y < 0$, $\operatorname{sign}(y) = 1$ for $y > 0$ and 0 otherwise, $L^0_{s\wedge\tau\wedge\tau_R}(J_i^{\gamma,R})$ is the local time of the process $J_i^{\gamma,R}$ at 0 in the time interval $[0, s\wedge\tau\wedge\tau_R]$, and $\Delta(J_{i,\lambda}^{\gamma,R}) := J_{i,\lambda}^{\gamma,R} - J_{i,\lambda-}^{\gamma,R}$. Notice that $J_{i,\lambda}^{\gamma,R} = J_{i,\lambda-}^{\gamma,R}$ for $i = 1,\dots \hbar$ because we have that $\int_0^s \bar{n}_{i,r}\mathrm{d}\bar{\nu}_r = \int_0^s n_{i,r}\mathrm{d}\nu_r$. Thus, using the SDEs defined above we have

$$|J_{i,s}^{\gamma,R}| = \int_0^{s\wedge\tau\wedge\tau_R} \operatorname{sign}(J_{i,\lambda}^{\gamma,R})(b^i(X_\lambda^{\gamma,R}) - b^i(X_\lambda^{0,R}))\,\mathrm{d}\lambda + \int_0^{s\wedge\tau\wedge\tau_R} \operatorname{sign}(J_{i,\lambda}^{\gamma,R})(\sigma^i(X_{i,\lambda}^{\gamma,R}) - \sigma^i(X_{i,\lambda}^{0,R}))\,\mathrm{d}W_\lambda + L^0_{s\wedge\tau\wedge\tau_R}(J_i^{\gamma,R})$$

for $i = 1,\dots \hbar$.

Taking expectation in the equation above we get

$$\begin{aligned}\mathsf{E}_x\big[|J_{i,s}^{\gamma,R}|\big] &= \mathsf{E}_x\left[\int_0^{s\wedge\tau\wedge\tau_R} \operatorname{sign}(J_{i,\lambda}^{\gamma,R})(b^i(X_\lambda^{\gamma,R}) - b^i(X_\lambda^{R0}))\,\mathrm{d}\lambda + L^0_{s\wedge\tau\wedge\tau_R}(J_i^{\gamma,R})\right] \\ &\leq \mathsf{E}_x\left[\int_0^s |b^i(X_\lambda^{\gamma,R}) - b^i(X_\lambda^{0,R})|\,\mathrm{d}\lambda + L^0_s(J_i^{\gamma,R})\right] \\ &\leq \mathsf{E}_x\left[D_1\int_0^s |J_\lambda^{\gamma,R}|_d\,\mathrm{d}\lambda + L^0_s(J_i^{\gamma,R})\right],\end{aligned} \tag{4.20}$$

where D_1 comes from Assumption 4.1. In order to estimate the local time, we follow the idea of [20, Lem. 5.1] which we can apply because $J_i^{\gamma,R}$ is a continuous semimartingale:

$$\begin{aligned}&\mathsf{E}_x\big[L^0_s(J_i^{\gamma,R})\big] \\ &\leq 4\epsilon - 2\mathsf{E}_x\left[\int_0^s \Big(\mathbb{1}_{\{J_{i,\lambda}^{\gamma,R}\in[0,\epsilon)\}} + \mathbb{1}_{\{J_{i,\lambda}^{\gamma,R}\geq\epsilon\}}e^{1-\frac{J_{i,\lambda}^{\gamma,R}}{\epsilon}}\Big)(b^i(X_\lambda^{\gamma,R}) - b^i(X_\lambda^{0,R}))\,\mathrm{d}\lambda\right] \\ &\quad + \frac{1}{\epsilon}\mathsf{E}_x\left[\int_0^s \mathbb{1}_{\{J_{i,\lambda}^{\gamma,R}>\epsilon\}}e^{1-\frac{J_{i,\lambda}^{\gamma,R}}{\epsilon}}(\sigma^i(X_{i,\lambda}^{\gamma,R}) - \sigma^i(X_{i,\lambda}^{0,R}))^2\,\mathrm{d}\lambda\right]\end{aligned} \tag{4.21}$$

for any $\epsilon > 0$. Denote by I_ϵ the last integral on the right-hand side above, we estimate it as follows: pick $\zeta \in (\frac{1}{2}, 1)$, we have

$$\begin{aligned}
I_\epsilon =& \frac{1}{\epsilon}\mathsf{E}_x\Big[\int_0^s \mathbb{1}_{\{J_{i,\lambda}^{\gamma,R}\in(\epsilon,\epsilon^\zeta)\}} e^{1-\frac{J_{i,\lambda}^{\gamma,R}}{\epsilon}}(\sigma^i(X_{i,\lambda}^{\gamma,R}) - \sigma^i(X_{i,\lambda}^{0,R}))^2\,\mathrm{d}\lambda\Big] \\
&+ \frac{1}{\epsilon}\mathsf{E}_x\Big[\int_0^s \mathbb{1}_{\{J_{i,\lambda}^{\gamma,R}\geq\epsilon^\zeta\}} e^{1-\frac{J_{i,\lambda}^{\gamma,R}}{\epsilon}}(\sigma^i(X_{i,\lambda}^{\gamma,R}) - \sigma^i(X_{i,\lambda}^{0,R}))^2\,\mathrm{d}\lambda\Big] \\
\leq& \frac{1}{\epsilon}\mathsf{E}_x\Big[D_1^2\int_0^s \mathbb{1}_{\{J_{i,\lambda}^{\gamma,R}\in(\epsilon,\epsilon^\zeta)\}} |J_{i,\lambda}^{\gamma,R}|^2\,\mathrm{d}\lambda\Big] \\
&+ \frac{1}{\epsilon}\mathsf{E}_x\Big[e^{1-\epsilon^{\zeta-1}}\int_0^s \mathbb{1}_{\{J_{i,\lambda}^{\gamma,R}\geq\epsilon^\zeta\}}(\sigma^i(X_{i,\lambda}^{\gamma,R}) - \sigma^i(X_{i,\lambda}^{0,R}))^2\,\mathrm{d}\lambda\Big] \\
\leq& D_1^2\epsilon^{2\zeta-1}T + \kappa_R^2\epsilon^{-1}e^{1-\epsilon^{\zeta-1}}\mathsf{E}_x\Big[\int_0^s |J_{i,\lambda}^{\gamma,R}|_d\,\mathrm{d}\lambda\Big]
\end{aligned}$$

where the first inequality for the first integral is by the Lipschitz property of σ with D_1 from Assumption 4.1 and $e^{1-\frac{J_{i,\lambda}^{\gamma,R}}{\epsilon}} \leq 1$ on the event $\mathbb{1}_{\{J_{i,\lambda}^{\gamma,R}\in(\epsilon,\epsilon^\zeta)\}}$, and for the second integral is by $e^{1-\frac{J_{i,\lambda}^{\gamma,R}}{\epsilon}} \leq e^{1-\epsilon^{\zeta-1}}$ on the event $\mathbb{1}_{\{J_{i,\lambda}^{\gamma,R}\geq\epsilon^\zeta\}}$. The second inequality for the first integral is by $|J_{i,\lambda}^{\gamma,R}| \leq \epsilon^\zeta$ on the event $\mathbb{1}_{\{J_{i,\lambda}^{\gamma,R}\in(\epsilon,\epsilon^\zeta)\}}$ and the computation of the integral, and for the second integral we use that σ^i is $\frac{1}{2}$-Hölder with constant κ_R on the random time interval $[0, \tau_R]$.

Thus we obtain that

$$\mathsf{E}_x\big[L_s^0(J_i^{\gamma,R})\big] \leq 4\epsilon + \Big(4D_1 + \frac{\kappa_R^2}{\epsilon}e^{1-\epsilon^{\zeta-1}}\Big)\mathsf{E}_x\Big[\int_0^s |J_\lambda^{\gamma,R}|_d\,\mathrm{d}\lambda\Big] + D_1^2\epsilon^{2\zeta-1}T, \qquad (4.22)$$

where $4D_1$ comes from the second term of the right-hand side of (4.21) and we have also bound $|J_{i,\lambda}^{\gamma,R}| \leq |J_\lambda^{\gamma,R}|_d$. Combining (4.20) and (4.22), we obtain the following estimate:

$$\mathsf{E}_x\big[|J_{i,s}^{\gamma,R}|\big] \leq 4\epsilon + \Big(5D_1 + \frac{\kappa_R^2}{\epsilon}e^{1-\epsilon^{\zeta-1}}\Big)\mathsf{E}_x\Big[\int_0^s |J_\lambda^{\gamma,R}|_d\,\mathrm{d}\lambda\Big] + D_1^2\epsilon^{2\zeta-1}T, \qquad (4.23)$$

for $i = 1, \ldots \hbar$. The remaining coordinates are estimated as follows. Let $i =$

$\hbar+1,\dots d$, we have from (4.19)

$$\begin{aligned}|J_{i,s}^{\gamma,R}| = &\int_0^s \operatorname{sign}(J_{i,\lambda}^{\gamma,R})(b^i(X_\lambda^{\gamma,R}) - b^i(X_\lambda^{0,R}))\,\mathrm{d}\lambda \\ &+ \int_0^s \operatorname{sign}(J_{i,\lambda}^{\gamma,R})(\sigma^i(X_{i,\lambda}^{\gamma,R}) - \sigma^i(X_{i,\lambda}^{0,R}))\,\mathrm{d}W_\lambda \\ &+ \gamma\int_{0+}^s \operatorname{sign}(J_{i,\lambda-}^{\gamma,R}) n_{i,\lambda-}\,\mathrm{d}\nu_\lambda + L_s^0(J_{i,\lambda}^{\gamma,R}) \\ &+ \sum_{0<\lambda\le s} |J_{i,\lambda}^{\gamma,R}| - |J_{i,\lambda-}^{\gamma,R}| - \operatorname{sign}(J_{i,\lambda-}^{\gamma,R})\Delta(J_{i,\lambda}^{\gamma,R}).\end{aligned} \tag{4.24}$$

We notice that

$$|J_{i,\lambda}^{\gamma,R}| = |J_{i,\lambda-}^{\gamma,R} + \gamma n_{i,\lambda}\Delta\nu_\lambda| \le |J_{i,\lambda-}^{\gamma,R}| + \gamma\Delta\nu_\lambda$$

$$\gamma\int_0^s \operatorname{sign}(J_{i,\lambda}^{\gamma,R}) n_{i,\lambda}\,\mathrm{d}\nu_\lambda^c = \gamma\int_{0+}^s \operatorname{sign}(J_{i,\lambda-}^{\gamma,R}) n_{i,\lambda-}\,\mathrm{d}\nu_\lambda - \sum_{0<\lambda\le s} \operatorname{sign}(J_{i,\lambda-}^{\gamma,R})\Delta(J_{i,\lambda}^{\gamma,R}),$$

where ν^c denotes the continuous part of ν. It is clear that

$$\gamma\int_0^s \operatorname{sign}(J_{i,\lambda}^{\gamma,R}) n_{i,\lambda}\,\mathrm{d}\nu_\lambda^c + \gamma\sum_{0<\lambda\le s}\Delta\nu_\lambda \le \nu_s.$$

Thus, we get from (4.24) the inequality:

$$\begin{aligned}|J_{i,s}^{\gamma,R}| \le &\int_0^s \operatorname{sign}(J_{i,\lambda}^{\gamma,R})(b^i(X_\lambda^{\gamma,R}) - b^i(X_\lambda^{0,R}))\,\mathrm{d}\lambda \\ &+ \int_0^s \operatorname{sign}(J_{i,\lambda}^{\gamma,R})(\sigma^i(X_{i,\lambda}^{\gamma,R}) - \sigma^i(X_{i,\lambda}^{0,R}))\,\mathrm{d}W_\lambda \\ &+ \gamma\nu_s + L_s^0(J_{i,\lambda}^{\gamma,R}).\end{aligned}$$

Since we have a non-continuous semimartingale, then [20, Lemma 5.1] does not apply directly. We have some extra terms, i.e, the second expectation and the last term on the right-hand side below, similar to (4.24), in the proof of [20, Lemma

5.1]. We give the full result in detail in C.1 in Appendix.

$$
\begin{aligned}
&\mathsf{E}_x\big[\,L^0_s(J_i^{\gamma,R})\big]\\
&\le 4\epsilon - 2\mathsf{E}_x\bigg[\int_0^s(\mathbb{1}_{\{J_{i,\lambda}^{\gamma,R}\in[0,\epsilon)\}} + \mathbb{1}_{\{J_{i,\lambda}^{\gamma,R}\ge\epsilon\}}e^{1-\frac{J_{i,\lambda}^{\gamma,R}}{\epsilon}})(b^i(X_\lambda^{\gamma,R}) - b^i(X_\lambda^{0,R}))\,\mathrm{d}\lambda\bigg]\\
&\quad - 2\mathsf{E}_x\bigg[\int_0^s(\mathbb{1}_{\{J_{i,\lambda}^{\gamma,R}\in[0,\epsilon)\}} + \mathbb{1}_{\{J_{i,\lambda}^{\gamma,R}\ge\epsilon\}}e^{1-\frac{J_{i,\lambda}^{\gamma,R}}{\epsilon}})\gamma n_{i,\lambda}\,\mathrm{d}\nu_\lambda^c\bigg]\\
&\quad + \frac{1}{\epsilon}\mathsf{E}_x\bigg[\int_0^s \mathbb{1}_{\{J_{i,\lambda}^{\gamma,R}>\epsilon\}}e^{1-\frac{J_{i,\lambda}^{\gamma,R}}{\epsilon}}(\sigma^i(X_{i,\lambda}^{\gamma,R}) - \sigma^i(X_{i,\lambda}^{0,R}))^2\,\mathrm{d}\lambda\bigg]\\
&\quad + 2\mathsf{E}_x[\gamma\sum_{0<\lambda\le s}\Delta\nu_\lambda].
\end{aligned}
$$

Repeating the same idea of [20] and the above estimates for the first $\hbar$ coordinates, we get

$$
\begin{aligned}
\mathsf{E}_x\big[|J_{i,s}^{\gamma,R}|\big] \le 4\epsilon + \Big(5D_1 + \frac{\kappa_R^2}{\epsilon}e^{1-\epsilon^{\zeta-1}}\Big)\mathsf{E}_x\bigg[\int_0^s |J_\lambda^{\gamma,R}|_d\,\mathrm{d}\lambda\bigg]\\
+ D_1^2\epsilon^{2\zeta-1}T + 7\gamma\mathsf{E}_x\big[\nu_s\big].
\end{aligned}
\tag{4.25}
$$

Combining (4.23) and (4.25) for $i = 1,\dots d$ we have

$$
\begin{aligned}
\mathsf{E}_x\big[|J_s^{\gamma,R}|_d\big] &\le \sum_{i=1}^d \mathsf{E}_x\big[|J_{i,s}^{\gamma,R}|\big]\\
&\le 4d\epsilon + d\Big(5D_1 + \frac{\kappa_R^2}{\epsilon}e^{1-\epsilon^{\zeta-1}}\Big)\mathsf{E}_x\bigg[\int_0^s |J_\lambda^{\gamma,R}|_d\,\mathrm{d}\lambda\bigg]\\
&\quad + dD_1^2\epsilon^{2\zeta-1}T + 7d\gamma\mathsf{E}_x\big[\nu_s\big].
\end{aligned}
$$

Sending $\epsilon\downarrow 0$, we get

$$
\mathsf{E}_x\big[|J_s^{\gamma,R}|_d\big] \le 5dD_1\mathsf{E}_x\bigg[\int_0^s |J_\lambda^{\gamma,R}|_d\,\mathrm{d}\lambda\bigg] + 7d\gamma\mathsf{E}_x\big[\nu_s\big].
$$

We can now apply Gronwall's lemma to get

$$
\mathsf{E}_x\bigg[|J_s^{\gamma,R}|_d\bigg] \le \gamma K_7\mathsf{E}_x\big[\nu_{T-t}\big], \quad \text{for any } s\in[0,T-t],
$$

with $K_7 = K_7(D_1, d, T)$. Passing to the limit for $R\to\infty$ and by Fatou's Lemma, we get

$$\mathsf{E}_x\left[|J_s^\gamma|_d\right] \le \liminf_{R\to\infty} \mathsf{E}_x\left[|J_s^{\gamma,R}|_d\right] \le \gamma K_7 \mathsf{E}_x\left[\nu_{T-t}\right], \quad \text{for any } s \in [0, T-t].$$

Recalling that $J_\cdot^\gamma = X_{\cdot\wedge\tau}^\gamma - X_{\cdot\wedge\tau}^0 = X_{\cdot\wedge\tau}^{[n,\nu],\gamma} - X_{\cdot\wedge\tau}^{[\bar{n},\bar{\nu}],0}$ and picking $s = T - t$, we have that

$$\begin{aligned}\mathsf{E}_x\left[|X_\tau^{[n,\nu],\gamma} - X_\tau^{[\bar{n},\bar{\nu}],0}|_d\right] &= \mathsf{E}_x\left[|X_{\tau\wedge(T-t)}^{[n,\nu],\gamma} - X_{\tau\wedge(T-t)}^{[\bar{n},\bar{\nu}],0}|_d\right] \\ &= \mathsf{E}_x\left[|J_{T-t}^\gamma|_d\right] \\ &\le \gamma K_7 \mathsf{E}_x\left[\nu_{T-t}\right],\end{aligned}$$

with K_7 independent of ν and γ. □

4.1.4 The Value of the Game

Theorem 4.12 *The game in (4.3) admits a value v. Moreover, for any compact $\mathcal{K} \subset \mathbb{R}_{0,T}^{d+1}$, there exists a $C_\mathcal{K}$ such that*

$$\sup_{(t,x)\in\mathcal{K}} |u^\gamma(t,x) - v(t,x)| \le C_\mathcal{K}\gamma^{\frac{1}{2}},$$

where u^γ is the function described in Theorem 4.6.

Proof. Let u^γ be the value of the game described in Theorem 4.6. We introduce $\underline{u} := \liminf_{\gamma\to 0} u^\gamma$ and $\overline{u} := \limsup_{\gamma\to 0} u^\gamma$. If we prove that

$$\overline{u}(t,x) \le \underline{v}(t,x) \quad \text{and} \quad \underline{u}(t,x) \ge \overline{v}(t,x),$$

for all $(t,x) \in \mathbb{R}_{0,T}^{d+1}$, we get that $\underline{u}(t,x) = \overline{u}(t,x) = v(t,x) = \underline{v}(t,x) = \overline{v}(t,x)$ and Theorem 4.12 holds.

Fix $(t,x) \in \mathbb{R}_{0,T}^{d+1}$. We first prove that $\underline{u} \ge \overline{v}$. Let $(n^\gamma, \nu^\gamma) \in \mathcal{A}_t^d$ be an η-optimal control for $u^\gamma(t,x)$ and consider the associated $(\bar{n}^\gamma, \bar{\nu}^\gamma) \in \mathcal{A}_t^\hbar$ constructed as in (4.16). Consider the processes $X^{[n^\gamma,\nu^\gamma],\gamma}$ and $X^{[\bar{n}^\gamma,\bar{\nu}^\gamma],0}$ from (4.17) and (4.18) with (n,ν) and $(\bar{n},\bar{\nu})$ therein replaced by (n^γ,ν^γ) and $(\bar{n}^\gamma,\bar{\nu}^\gamma)$, respectively. For notational simplicity, denote the two processes $X^{[n^\gamma,\nu^\gamma],\gamma}$ and $X^{[\bar{n}^\gamma,\bar{\nu}^\gamma],0}$ by X^γ and

X^0, respectively. Let $\tau \in \mathcal{T}_t$ be an η-optimal stopping time for $\overline{v}(t,x)$. We have

$$\begin{aligned} u^\gamma(t,x) - \overline{v}(t,x) &\geq \mathcal{J}^\gamma_{t,x}(n^\gamma, \nu^\gamma, \tau) - \mathcal{J}_{t,x}(\bar{n}^\gamma, \bar{\nu}^\gamma, \tau) - 2\eta \\ &= \mathsf{E}_x\Bigg[e^{-r\tau}(g(t+\tau, X^\gamma_\tau) - g(t+\tau, X^0_\tau)) \\ &\qquad + \int_0^\tau e^{-rs}(h(t+s, X^\gamma_s) - h(t+s, X^0_s))\,\mathrm{d}s \\ &\qquad + \int_{[0,\tau]} e^{-rs} f^\gamma(t+s)\,\mathrm{d}\nu^\gamma_s - \int_{[0,\tau]} e^{-rs} f(t+s)\,\mathrm{d}\bar{\nu}^\gamma_s \Bigg] - 2\eta \\ &\geq -K_1 \mathsf{E}_x\Big[|X^\gamma_\tau - X^0_\tau|_d\Big] - \mathsf{E}_x\Big[\int_0^{T-t} |X^\gamma_s - X^0_s|_d\,\mathrm{d}s\Big] \\ &\geq -K_1 \mathsf{E}_x\Big[|X^\gamma_\tau - X^0_\tau|_d\Big] - K_1 T \sup_{s\in[0,T-t]} \mathsf{E}_x\Big[|X^\gamma_s - X^0_s|_d\Big] - 2\eta \end{aligned}$$

where K_1 comes from (4.6). The first inequality is by the choice of (n^γ, ν^γ) and τ. The second inequality is by the definition of $\bar{\nu}$, i.e., $\frac{\mathrm{d}\bar{\nu}_s(\omega)}{\mathrm{d}\nu_s(\omega)} = |n^\hbar_s(\omega)|_\hbar \leq 1$ for all $\omega \in \Omega$ and that $f^\gamma \geq f$ by (4.11). The last inequality is by Fubini's theorem and taking supremum inside the integral. Using Theorem 4.11 combined with $\mathsf{E}_x[|\nu^\gamma_{T-t}] \leq K_6$ from Lemma 4.8 we have that

$$u^\gamma(t,x) - \overline{v}(t,x) \geq -K_1(1+T)\gamma K_7 K_6 - 2\eta,$$

and passing to the limit inferior as $\gamma \downarrow 0$ we get

$$\underline{u}(t,x) - \overline{v}(t,x) \geq -2\eta.$$

By the arbitrariness of η, we obtain $\underline{u}(t,x) \geq \overline{v}(t,x)$.

We prove now that $\overline{u}(t,x) \leq \underline{v}(t,x)$. Let $\tau_\gamma \in \mathcal{T}_t$ be an η-optimal stopping time for u^γ and let $(n,\nu) \in \mathcal{A}^\hbar_t$ be an η-optimal control for $\underline{v}(t,x)$. Notice that since $(n,\nu) \in \mathcal{A}^\hbar_t \subset \mathcal{A}^d_t$, we have $X^{[n,\nu],\gamma}_s(\omega) = X^{[n,\nu],0}_s(\omega)$ for all $s \in [0, T-t]$, almost every $\omega \in \Omega$. Thus

$$\begin{aligned} u^\gamma(t,x) - \underline{v}(t,x) &\le \mathcal{J}^\gamma_{t,x}(n,\nu,\tau_\gamma) - \mathcal{J}_{t,x}(n,\nu,\tau_\gamma) + 2\eta \\ &= \mathsf{E}_x\Big[\int_{[0,\tau_\gamma]} e^{-rs}(f^\gamma(t+s) - f(t+s))\,\mathrm{d}\nu_s\Big] + 2\eta \\ &\le \sqrt{\gamma}K_1\mathsf{E}_x[|\nu_{T-t}|] + 2\eta \\ &\le \sqrt{\gamma}K_1K_6 + 2\eta. \end{aligned}$$

The first inequality is by the choice of τ_γ and (n,ν); the equality is because the two processes are indistinguishable; the second inequality is by the definition of f^γ and thus $f^\gamma \le f + \sqrt{\gamma}K_1$ (see (4.11)); the third inequality is by Remark 4.9. Passing to the limit superior as $\gamma \downarrow 0$ and by the arbitrariness of η we get $\overline{u}(t,x) \le \underline{v}(t,x)$ and we have $\lim_{\gamma\to 0} u^\gamma(t,x) = \underline{v}(t,x) = \overline{v}(t,x) = v(t,x)$.

Using that $\gamma \in (0,1)$, we obtain

$$|u^\gamma(t,x) - v(t,x)| \le K_1K_6(1 + K_7(1+T))\gamma^{\frac{1}{2}}$$

for all $(t,x) \in \mathbb{R}^{d+1}_{0,T}$, with K_1 from (4.6), $K_6 = K_6(x)$ from Lemma 4.8 and K_7 as in Theorem 4.11. Since $K_6 = K_6(x)$ is dependent continuously on x, let $\mathcal{K} \subset \mathbb{R}^{d+1}_{0,T}$ be a compact, we have that

$$\sup_{(t,x)\in\mathcal{K}} |u^\gamma(t,x) - v(t,x)| \le C_{\mathcal{K}}\gamma^{\frac{1}{2}}$$

with $C_{\mathcal{K}} := K_1(1 + K_7(1+T))\max_{(t,x)\in\mathcal{K}} K_6(x)$. □

We show the optimality of the stopping time defined in (4.9). We follow an approach that can be found for example in [13, Thm. 4.12].

Theorem 4.13 *For any given $(t,x) \in \mathbb{R}^{d+1}_{0,T}$ and any admissible control $(n,\nu) \in \mathcal{A}^\hbar_t$, the stopping time $\tau_* = \tau_*(t,x;n,\nu) \in \mathcal{T}_t$ defined under P_x as*

$$\tau_* := \inf\big\{s \ge 0 \,\big|\, v(t+s, X^{[n,\nu]}_s) = g(t+s, X^{[n,\nu]}_s)\big\} \wedge (T-t) \tag{4.26}$$

is optimal for the stopper.

Proof. Let $(t,x) \in \mathbb{R}^{d+1}_{0,T}$ and $(n,\nu) \in \mathcal{A}^{\hbar}_t$. Define

$$\tau^{\gamma}_* := \inf\left\{s \geq 0 \,\middle|\, u^{\gamma}(t+s, X^{[n,\nu]}_s) \leq g(t+s, X^{[n,\nu]}_s)\right\} \wedge (T-t).$$

We first prove that

$$\liminf_{\gamma \to 0} \tau^{\gamma}_* \geq \tau_*. \tag{4.27}$$

Fix $\omega \in \Omega$. If $\tau_*(\omega) = 0$, then (4.27) holds. If $\tau_*(\omega) > 0$, we pick $\delta > 0$ such that $\delta < \tau_*(\omega)$. By the same arguments that lead to the optimality of τ_* in the proof of Theorem 3.33, we have that $\tau_0 = \tau_*$ for almost every $\omega \in \Omega$ and $\Delta\nu_{\tau_*} = 0$, i.e., $X^{[n,\nu]}_{\tau_0 -} = X^{[n,\nu]}_{\tau_* -} = X^{[n,\nu]}_{\tau_*}$, where

$$\tau_0 = \inf\{s \geq 0 | v(t+s, X^{[n,\nu]}_{s-}) - g(t+s, X^{[n,\nu]}_{s-}) = 0\}.$$

We have that the sequence $(\tau_\eta)_{\eta \in (0,1)}$ with

$$\tau_\eta := \inf\{s \geq 0 | v(t+s, X^{[n,\nu]}_s) - g(t+s, X^{[n,\nu]}_s) \leq \eta\}$$

admits a strictly increasing subsequence $(\tau_{\eta_m})_{m \in \mathbb{N}}$ for almost every $\omega \in \Omega$, i.e., there exists an $\Omega_0 \subset \Omega$ with $\mathsf{P}(\Omega_0) = 1$ where $\tau_{\eta_m}(\omega) \uparrow \tau_0(\omega)$ for all ω as $m \to \infty$. Moreover,

$$\lim_{m \to \infty} X^{[n,\nu]}_{\tau_{\eta_m}} = X^{[n,\nu]}_{\tau_0 -}$$

holds for all $\omega \in \Omega_0$. It means that there exists a constant $C_\omega > 0$ such that

$$\inf_{0 \leq s \leq \delta} \left(v(t+s, X^{[n,\nu]}_s(\omega)) - g(t+s, X^{[n,\nu]}_s(\omega))\right) \geq C_\omega,$$

otherwise we would have that for all $\eta_m > 0$ from the sequence above,

$$\inf_{0 \leq s \leq \delta} \left(v(t+s, X^{[n,\nu]}_s(\omega)) - g(t+s, X^{[n,\nu]}_s(\omega))\right) < \eta_m,$$

and the inequality would imply that $\tau_{\eta_m} \leq \delta$ for all $m \in \mathbb{N}$. Since $\tau_{\eta_m} \uparrow \tau_* > \delta$, we reach a contradiction.

Moreover $(n,\nu) \in \mathcal{A}^{\hbar}_t$ is fixed, therefore we have that the second moment of the controlled process is finite and thus there exists a compact $\mathcal{K}_\omega \subset \mathbb{R}^{d+1}_{0,T}$ such that the trajectory of $X^{n,\nu}$ lies in it

$$\{(t+s, X^{[n,\nu]}_s) : s \in [0,\delta]\} \subset \mathcal{K}_\omega.$$

By Theorem 4.12, we have that there exists a γ' such that for all $0 < \gamma \leq \gamma'$:

$$\sup_{(t,x)\in\mathcal{K}_\omega} \big|u^\gamma(t,x) - v(t,x)\big| < \frac{C_\omega}{2},$$

and it follows

$$\inf_{0\leq s\leq\delta} \big|u^\gamma(t+s, X^{[n,\nu]}_s(\omega)) - g(t+s, X^{[n,\nu]}_s(\omega))\big| \geq \frac{C_\omega}{2}.$$

It means that $\tau^\gamma_* > \delta$ for $0 < \gamma \leq \gamma'$, and

$$\liminf_{\gamma\to 0} \tau^\gamma_* \geq \delta.$$

Sending $\delta \uparrow \tau_*$ we have that (4.27) holds and we obtain

$$\lim_{\gamma\to 0} \tau^\gamma_* \wedge \tau_* = \tau_*, \tag{4.28}$$

with τ_* as in (4.26).

Fix $\eta > 0$ and consider the stopping time

$$\tau^\gamma_\eta \coloneqq \inf\{s \geq 0 | u^\gamma(t+s, X^{[n,\nu]}_s) \leq g(t+s, X^{[n,\nu]}_s) + \eta\}.$$

Since $(n,\nu) \in \mathcal{A}^\hbar_t$, then (n,ν) is admissible for the class $\mathcal{A}^d_t$ (precisely, $n \mapsto \hat{n} \coloneqq (n, \mathbf{0}) \in \mathbb{R}^\hbar \times \mathbb{R}^{d-\hbar}$, we have that $(\hat{n},\nu) \in \mathcal{A}^d_t$). By the arguments in proof of Theorem 3.33 about the optimality of τ_* therein, we have that $\tau^\gamma_\eta \uparrow \tau^\gamma_*$ as $\eta \downarrow 0$. By the first line of (3.124) with g and τ_η therein replaced by u^γ and $\tau^\gamma_\eta \wedge \tau_*$, respectively, we have

$$\begin{aligned} u^\gamma(t,x) \leq \eta + \mathsf{E}_x\bigg[& e^{-r(\tau^\gamma_\eta\wedge\tau_*)} u^\gamma(t+\tau^\gamma_\eta\wedge\tau_*, X^{[n,\nu]}_{\tau^\gamma_\eta\wedge\tau_*}) + \int_0^{\tau^\gamma_\eta\wedge\tau_*} e^{-rs} h(t+s, X^{[n,\nu]}_s)\,\mathrm{d}s \\ & + \int_{[0,\tau^\gamma_\eta\wedge\tau_*]} e^{-rs} f^\gamma(t+s)\,\mathrm{d}\nu_s \bigg]. \end{aligned}$$

Sending $\eta \downarrow 0$, by dominated convergence theorem we obtain

$$u^\gamma(t,x) \le \mathsf{E}_x\bigg[e^{-r(\tau_*^\gamma \wedge \tau_*)} u^\gamma(t+\tau_*^\gamma \wedge \tau_*, X^{[n,\nu]}_{\tau_*^\gamma \wedge \tau_*}) + \int_0^{\tau_*^\gamma \wedge \tau_*} e^{-rs} h(t+s, X_s^{[n,\nu]})\,\mathrm{d}s \\ + \int_{[0,\tau_*^\gamma \wedge \tau_*]} e^{-rs} f^\gamma(t+s)\,\mathrm{d}\nu_s \bigg].$$

Finally, sending $\gamma \downarrow 0$ we have that

$$v(t,x) \le \lim_{\gamma \to 0} \mathsf{E}_x\bigg[e^{-r(\tau_*^\gamma \wedge \tau_*)} u^\gamma(t+\tau_*^\gamma \wedge \tau_*, X^{[n,\nu]}_{\tau_*^\gamma \wedge \tau_*}) + \int_0^{\tau_*^\gamma \wedge \tau_*} e^{-rs} h(t+s, X_s^{[n,\nu]})\,\mathrm{d}s \\ + \int_{[0,\tau_*^\gamma \wedge \tau_*]} e^{-rs} f^\gamma(t+s)\,\mathrm{d}\nu_s \bigg].$$

Using that $(n,\nu) \in \mathcal{A}_t^\hbar$, f is bounded, the functions g and h have linear growth, dominated convergence theorem applies and we obtain by (4.28) that

$$v(t,x) \le \mathsf{E}_x\bigg[e^{-r\tau_*} g(t+\tau_*, X^{[n,\nu]}_{\tau_*}) + \int_0^{\tau_*} e^{-rs} h(t+s, X_s^{[n,\nu]})\,\mathrm{d}s \\ + \int_{[0,\tau_*]} e^{-rs} f(t+s)\,\mathrm{d}\nu_s \bigg].$$

Thus, the strategy τ_* is optimal for the stopper. □

The next remark shows that the value function v of the game does not change if we impose the controller to use only absolute continuous controls.

Remark 4.14: Remark 3.34 says that

$$u^\gamma(t,x) = \inf_{(n,\nu)\in\mathcal{A}_t^{d,\circ}} \sup_{\tau \in \mathcal{T}_t} \mathcal{J}^\gamma_{t,x}(n,\nu,\tau) = \sup_{\tau \in \mathcal{T}_t} \inf_{(n,\nu)\in\mathcal{A}_t^{d,\circ}} \mathcal{J}^\gamma_{t,x}(n,\nu,\tau),$$

where $\mathcal{A}_t^{d,\circ} := \{(n,\nu) \in \mathcal{A}_t^d | \nu \text{ is absolutely continuous}\}$. We introduce the lower and upper value functions of the games with absolutely continuous admissible controls:

$$\underline{v}^\circ(t,x) = \sup_{\tau \in \mathcal{T}_t} \inf_{(n,\nu)\in\mathcal{A}_t^{\hbar,\circ}} \mathcal{J}_{t,x}(n,\nu,\tau) \quad \text{and} \quad \overline{v}^\circ(t,x) = \inf_{(n,\nu)\in\mathcal{A}_t^{\hbar,\circ}} \sup_{\tau \in \mathcal{T}_t} \mathcal{J}_{t,x}(n,\nu,\tau).$$

Repeating the proof of Theorem 4.12 replacing $\underline{v}$ and $\overline{v}$ therein with $\underline{v}^\circ$ and $\overline{v}^\circ$, respectively, we get that $\lim_{\gamma\to 0} u^\gamma = v^0(t,x)$ and

$$v^0(t,x) = \inf_{(n,\nu)\in\mathcal{A}_t^{\hbar,\circ}} \sup_{\tau\in\mathcal{T}_t} \mathcal{J}_{t,x}(n,\nu,\tau) = \sup_{\tau\in\mathcal{T}_t} \inf_{(n,\nu)\in\mathcal{A}_t^{\hbar,\circ}} \mathcal{J}_{t,x}(n,\nu,\tau).$$

Since the limit is unique, then we have that $v^0 = v$ with v as in Theorem 4.12. ■

4.2 Degenerate Processes

In this section, we relax the local ellipticity condition on the diffusion coefficient of the SDEs which we assumed in Chapter 3. We prove that stochastic games played with this new type of processes admit a value under a new set of assumptions. We use an approximation procedure to perturb the original process with a parametrised non-degenerate noise. The approximated games satisfy assumptions of Chapter 3 and we can apply results therein. The family of value functions admits a a convergent subsequence that converges to the the value of the original game. Moreover, we provide an optimal strategy for the stopper.

4.2.1 The Problem

We consider a zero-sum game as in Chapter 3 where there are still two players, a stopper (maximiser) and a controller (minimiser). In this new type of game we allow the diffusion coefficient (see σ in (4.29)) to be degenerate.

We repeat briefly the model without giving the full details which can be found in Chapter 3. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathsf{P})$ be a stochastic basis on which an adapted d'-dimensional Brownian motion $(W_t)_t$ is defined. Let T be the terminal time, and $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^{d'}$ be measurable functions. Denote

$$\mathcal{T}_t \coloneqq \{\tau | \tau \text{ is a stopping time}, \tau \le T - t\}.$$

We define by $\mathcal{A}_t$ the class of admissible controls as

$$\mathcal{A}_t := \left\{ (n,\nu) \left| \begin{array}{l} (n_s)_{s\in[0,\infty)} \text{ is progressively measurable, } \mathbb{R}^d \text{ valued,} \\ \text{with } |n_s|_d = 1,\ \mathsf{P}\text{-a.s. } \forall s \in [0,\infty); \\ (\nu_s)_{s\in[0,\infty)} \text{ is } \mathbb{F}\text{-adapted, real valued, non-decreasing and} \\ \text{right-continuous with } \nu_{0-} = 0,\ \mathsf{P}\text{-a.s., and } \mathsf{E}[|\nu_{T-t}|^2] < \infty \end{array} \right. \right\}.$$

Consider the following d-dimensional controlled stochastic differential equation. We will use the notation $X_s^{[n,\nu]}$ to underline that the process is controlled by (n,ν):

$$\begin{cases} \mathrm{d}X_s^{[n,\nu]} = b(X_s^{[n,\nu]})\mathrm{d}s + \sigma(X_s^{[n,\nu]})\mathrm{d}W_s + n_s \mathrm{d}\nu_s, & 0 \le s \le T, \\ X_{0-}^{n,\xi} = x. \end{cases} \tag{4.29}$$

Since $\sigma \in \mathbb{R}^{d\times d'}$, the resulting matrix $a(x) := (\sigma\sigma^\top)(x) \in \mathbb{R}^{d\times d}$ is a positive-semidefinite matrix which means that the matrix $a(x)$ can be singular, i.e., there can be points $x \in \mathbb{R}^d$ and $\zeta \in \mathbb{R}^d$ such that

$$\langle \zeta, a(x)\zeta \rangle = 0,$$

or, alternatively, that $\det(a(x)) = 0$ for some $x \in \mathbb{R}^d$.

Remark 4.15: In Chapters 3 and 4, we required that $d \le d'$ because we assumed the diffusion coefficient to be non-degenerate. In the case where $d > d'$, the diffusion coefficient is always degenerate because $\det(a(x)) = 0$ for all $x \in \mathbb{R}^d$. ■

The players are playing the same zero-sum game introduced in (3.2) and we show it below for completeness

$$\mathcal{J}_{t,x}(n,\nu,\tau) = \mathsf{E}_x\Bigg[e^{-r\tau} g(t+\tau, X_\tau^{[n,\nu]}) + \int_0^\tau e^{-rs} h(t+s, X_s^{[n,\nu]})\,\mathrm{d}s + \int_{[0,\tau]} e^{-rs} f(t+s) \circ \mathrm{d}\nu_s \Bigg]. \tag{4.30}$$

In this setting, the function f is independent on the state space, and the integral defined in (4.30) is now consistent with the Riemann-Stieltjes integral, i.e.,

$$\int_{[0,\tau]} e^{-rs} f(t+s) \circ \mathrm{d}\nu_s = \int_0^\tau e^{-rs} f(t+s)\, \mathrm{d}\nu_s,$$

as in (4.2).

We define

$$\underline{v}(t,x) := \sup_{\tau\in\mathcal{T}_t} \inf_{(n,\nu)\in\mathcal{A}_t} \mathcal{J}_{t,x}(n,\nu,\tau), \text{ and } \overline{v}(t,x) := \inf_{(n,\nu)\in\mathcal{A}_t} \sup_{\tau\in\mathcal{T}_t} \mathcal{J}_{t,x}(n,\nu,\tau), \quad (4.31)$$

so that $\underline{v}(t,x) \le \overline{v}(t,x)$. If the equality holds then we say that the game admits a value:

$$v(t,x) := \underline{v}(t,x) = \overline{v}(t,x). \quad (4.32)$$

The local ellipticity condition in Assumption 3.4 does not hold anymore. Indeed, we allow the diffusion coefficient to be degenerate and thus (3.10) fails. This implies that some of the proofs in Chapter 3 are not valid anymore, in particular, Propositions 3.17 and 3.22, Lemma 3.29 and the argument in Section B.6.

We recall that $\mathcal{L}$ denotes the infinitesimal generator of the uncontrolled process $X^{[e_1,0]}$ (where e_1 is the unit vector with 1 in the first entry) and it reads

$$(\mathcal{L}\varphi)(x) = \frac{1}{2}\mathrm{tr}\left(a(x)D^2\varphi(x)\right) + \langle b(x), \nabla\varphi(x)\rangle. \quad (4.33)$$

Next we give assumptions under which we obtain our main result (Theorem 4.18).

Assumption 4.16 (Controlled SDE) *The functions b and σ are continuously differentiable on $\mathbb{R}^d$ and Lipschitz with constant D_1, i.e.,*

$$|b(x)-b(y)|_d + |\sigma(x)-\sigma(y)|_{d\times d'} \le D_1|x-y|_d, \qquad \textit{for all } x,y\in\mathbb{R}^d. \quad (4.34)$$

The matrix σ is such that there exists $D_2>0$ such that

$$|\sigma(x)|_{d\times d'} \le D_2(1+|x|_d)^{\frac{1}{2}}, \qquad \textit{for all } x\in\mathbb{R}^d. \quad (4.35)$$

Thanks to (4.34), we can assume that there exists D_2 such that

$$|b(x)|_d \leq D_2(1 + |x|_d), \quad \text{for all } x \in \mathbb{R}^d,$$

and without loss of generality we can assume that D_2 is the same of (4.35).

Assumption 4.17 (Functions f, g, h) *For the functions $f, g, h : \mathbb{R}^{d+1}_{0,T} \to [0, \infty)$ the following hold:*

(i) $g \in C^{1,2,\alpha}_{\ell oc}(\mathbb{R}^{d+1}_{0,T})$ *and* $h \in C^{0,1,\alpha}_{\ell oc}(\mathbb{R}^{d+1}_{0,T})$ *for some* $\alpha \in (0, 1)$*;*

(ii) f *is non-increasing, positive and* f^2 *is differentiable;*

(iii) there is $K_0 \in (0, \infty)$ *such that for all* $0 \leq s < t \leq T$ *and all* $x \in \mathbb{R}^{d+1}_{0,T}$

$$h(t, x) - h(s, x) \leq K_0(t - s) \quad \textit{and} \quad g(t, x) - g(s, x) \leq K_0(t - s);$$

(iv) there is $K_1 \in (0, \infty)$ *such that*

$$0 \leq |h(t, x)| \leq K_1(1 + |x|^2_d), \quad \textit{for } (t, x) \in \mathbb{R}^{d+1}_{0,T};$$

(v) there is $K_2 \in (0, \infty)$ *and* $\beta \in (0, 1)$ *such that*

$$|h(t, x) - h(t, y)| \leq K_2(1 + |x|_d + |y|_d)^\beta |x - y|_d, \quad \textit{for } (t, x) \in \mathbb{R}^{d+1}_{0,T}; \quad (4.36)$$

(vi) f *and* g *are such that*

$$|\nabla g(t, x)|_d \leq f(t), \quad \textit{for all } (t, x) \in \mathbb{R}^{d+1}_{0,T}; \quad (4.37)$$

(vii) there is $K_3 \in (0, \infty)$ *such that*

$$(h + \partial_t g + \mathcal{L}g - rg)(t, x) \geq -K_3, \quad \textit{for all } (t, x) \in \mathbb{R}^{d+1}_{0,T}.$$

Notice that (ii) and (4.37) in Assumption 4.17 imply that

$$|\nabla g(t, x)| \leq f(0) \quad \text{for all } (t, x) \in \mathbb{R}^{d+1}_{0,T}. \quad (4.38)$$

A comment on these conditions can be found in Chapter 3 below Assumption 3.5.

We can now state the main result of the chapter.

Theorem 4.18 *The game described above admits a value v (i.e., (4.32) holds). Moreover, for any given $(t,x) \in \mathbb{R}^{d+1}_{0,T}$ and any admissible control $(n,\nu) \in \mathcal{A}_t$, the stopping time $\tau_* = \tau_*(t,x;n,\nu) \in \mathcal{T}_t$ defined under P_x as*

$$\tau_* := \inf\left\{s \geq 0 \,\middle|\, v(t+s, X_s^{[n,\nu]}) = g(t+s, X_s^{[n,\nu]})\right\} \wedge (T-t)$$

is optimal for the stopper.

4.2.2 Approximated Problem

The theory developed in Chapter 3 does not apply directly to this new type of game. Results proved by probabilistic arguments apply without changes to this class of games, but the results proved by analytic arguments fail.

4.2.3 Approximation Procedure

We introduce an approximating sequence of games to which the results from Chapter 3 apply. These games are indexed by a parameter θ and we obtain a family of value functions of these games. The existence of the value of the original game is found through a convergent subsequence in this family.

Fix a $\theta \in (0,1)$. For a given pair $(n,\nu) \in \mathcal{A}_t$ we consider the (controlled) stochastic differential equation:

$$\mathrm{d}X_s^{[n,\nu],\theta} = b(X_s^{[n,\nu],\theta})\,\mathrm{d}s + n_s\,\mathrm{d}\nu_s + \sigma(X_s^{[n,\nu],\theta})\mathrm{d}W_s + \theta I_d \mathrm{d}\tilde{W}_s$$

where I_d is the d-dimensional identity matrix and $\tilde{W}$ is a d-dimensional Brownian motion independent from the original Brownian motion W. Let $\tau \in \mathcal{T}_t$, consider the same payoff defined in (4.30) but with $X^{[n,\nu],\theta}$ in the place of $X^{[n,\nu]}$ as underlying process, i.e.,

$$\mathcal{J}^{\theta}_{t,x}(n,\nu,\tau) = \mathsf{E}_x\bigg[e^{-r\tau} g(t+\tau, X_\tau^{[n,\nu],\theta}) + \int_0^\tau e^{-rs} h(t+s, X_s^{[n,\nu],\theta})\,\mathrm{d}s \qquad (4.39)$$
$$+ \int_{[0,\tau]} e^{-rs} f(t+s)\,\mathrm{d}\nu_s\bigg].$$

We say that the game (4.30) admits a value if

$$u^\theta(t,x) = \sup_{\tau\in\mathcal{T}_t} \inf_{(n,\nu)\in\mathcal{A}_t} \mathcal{J}^\theta_{t,x}(n,\nu,\tau) = \inf_{(n,\nu)\in\mathcal{A}_t} \sup_{\tau\in\mathcal{T}_t} \mathcal{J}^\theta_{t,x}(n,\nu,\tau). \tag{4.40}$$

Since we add a noise $\theta I_d \mathrm{d}\tilde{W}_s$ dependent on θ, then the differential operator $\mathcal{L}$ defined in (4.33) changes in $\mathcal{L}\theta$. Indeed the matrix $a(x)$ becomes $a^\theta(x) = a(x)+\theta^2 I_d$ which is non-degenerate because

$$\begin{aligned}\langle \zeta, a^\theta(x)\zeta\rangle &= \langle \zeta, (a(x) + \theta^2 I_d)\zeta\rangle \\ &= \langle \zeta, a(x)\zeta\rangle + \theta^2|\zeta|_d^2 \\ &\geq \theta^2|\zeta|_d^2\end{aligned}$$

for all $x, \zeta \in \mathbb{R}^d$. This game satisfies the Assumptions 3.4 and 3.5 and we are allowed to apply Theorem 3.6 with a replaced by a^θ. For simplicity we collect the results from Chapter 3 that we need in the next two theorems.

Theorem 4.19 *The game described above admits a value function u^θ (i.e., (4.40) holds). Moreover, for any given $(t,x) \in \mathbb{R}^{d+1}_{0,T}$ and any admissible control $(n,\nu) \in \mathcal{A}_t$, the stopping time $\tau^\theta_* = \tau^\theta_*(t,x;n,\nu) \in \mathcal{T}_t$ defined under P_x as*

$$\tau^\theta_* := \inf\left\{ s \geq 0 \,\middle|\, u^\theta(t+s, X^{[n,\nu],\theta}_s) = g(t+s, X^{[n,\nu],\theta}_s) \right\} \wedge (T-t) \tag{4.41}$$

is optimal for the stopper.

If we assume that the function h is uniformly Lipschitz in x, then the theorem still holds even if we drop (4.35) in Assumption 4.16.

Lemma 4.20 *For any $\theta \in (0,1)$, the function u^θ belongs to $(C^{0,1,\gamma}_{\ell oc} \cap W^{1,2,p}_{\ell oc})(\mathbb{R}^{d+1}_{0,T})$ for $\gamma \in (0,1)$ and $p \in (1,\infty)$. Moreover*

$$\begin{aligned}&0 \leq |u^\theta(t,x)| \leq K_4(1+|x|_d^2), \\ &|\nabla u^\theta(t,x)|_d \leq f(t),\end{aligned}$$

where $K_4 = K_4(T, D_1, K_1)$ with D_1 and K_1 from Assumptions 4.16 and 4.17, respectively, and it is independent of θ.

The Assumptions 4.16 and 4.17 are slightly different from Assumptions 4.1

and 4.2 in Section 4.1. Using Lemma 3.13 in order to repeat and complete the proof of Lemma 4.8 for u^θ, it is straightforward to show that:

Lemma 4.21 *There exists a $K_6 = K_6(x; T, f(T), K_4, D_1)$ independent of θ such that*

$$u^\theta(t,x) = \inf_{(n,\nu)\in\mathcal{A}^{opt}_{t,x}} \sup_{\tau\in\mathcal{T}_t} \mathcal{J}^\theta_{t,x}(n,\nu,\tau) = \sup_{\tau\in\mathcal{T}_t} \inf_{(n,\nu)\in\mathcal{A}^{opt}_{t,x}} \mathcal{J}^\theta_{t,x}(n,\nu,\tau)$$

where $\mathcal{A}^{opt}_{t,x} := \left\{(n,\nu) \in \mathcal{A}_t \middle| \mathsf{E}_x[\nu_{T-t}] \le K_6\right\}$ and

$$K_6 := \frac{e^{rT} K_4}{f(T)}(1 + |x|^2_d)$$

independent of θ and $K_4 = K_4(T, D_2, K_1)$ from Lemma 4.20.

Remark 4.22: The result in Lemma 4.21 can be extended to $\underline{v}(t,x)$ and we obtain that

$$\underline{v}(t,x) = \sup_{\tau\in\mathcal{T}_t} \inf_{(n,\nu)\in\mathcal{A}^{opt}_{t,x}} \mathcal{J}_{t,x}(n,\nu,\tau) = \inf_{(n,\nu)\in\mathcal{A}^{opt}_{t,x}} \sup_{\tau\in\mathcal{T}_t} \mathcal{J}_{t,x}(n,\nu,\tau),$$

with $\mathcal{A}^{opt}_{t,x} = \{(n,\nu) \in \mathcal{A}_t | \mathsf{E}_x[\nu_{T-t}] \le K_6\}$ and K_6 as in Lemma 4.21.

The proof follows the steps of the proofs of Lemmas 4.8 and 4.21. ■

We state a classical inequality which will be used in the next theorem: for all $p \in [1, \infty)$, the following inequality holds

$$\Big(\sum_{i=1}^d x_i\Big)^p \le d^{p-1}\Big(\sum_{i=1}^d |x_i|^p\Big) \tag{4.42}$$

for all $x \in \mathbb{R}^d$.

We show in the next theorem the rate of convergence of the process with the noise to the process without the noise.

Theorem 4.23 *Fix* $(t,x) \in \mathbb{R}^{d+1}_{0,T}$ *and* $\beta \in (0,1)$*. For any* $(n,\nu) \in \mathcal{A}_t$*, we have*

$$\mathsf{E}_x\Big[\sup_{s\in[0,T-t]} |X^{[n,\nu],\theta}_s - X^{[n,\nu],0}_s|_d^{\frac{1}{1-\beta}}\Big] \leq \theta K_7, \tag{4.43}$$

where $K_7 = K_7(D_1, d, T, \beta)$ *with* D_1 *from Assumption 4.16. Moreover, it holds*

$$\mathsf{E}_x\Bigg[\sup_{s\in[0,T-t]} |X^{[n,\nu],\theta}_s|\Bigg] \leq K_8 \mathsf{E}_x[\nu_{T-t}], \tag{4.44}$$

where $K_8 = K_8(d, T, D_2)$ *independent of* θ*.*

Proof. For simplicity denote $X^{[n,\nu],\theta}$, $X^{[n,\nu],0}$ and the difference process $X^{[n,\nu],\theta} - X^{[n,\nu],0}$ by X^θ, X^0 and J^θ, respectively. Taking the supremum over time and the expectation we get

$$\begin{aligned}\mathsf{E}_x\Bigg[\sup_{\lambda\in[0,s]} |J^\theta_\lambda|_d^{\frac{1}{1-\beta}}\Bigg] \leq 3^{\frac{\beta}{1-\beta}} \mathsf{E}_x\Bigg[\sup_{\lambda\in[0,s]}\Bigg(\Bigg|\int_0^\lambda \big(b(X^\theta_r) - b(X^0_r)\big)\,\mathrm{d}r\Bigg|_d^{\frac{1}{1-\beta}} + |\theta I_d \tilde{W}_\lambda|_d^{\frac{1}{1-\beta}} \\ + \Bigg|\int_0^\lambda \big(\sigma(X^\theta_r) - \sigma(X^0_r)\big)\,\mathrm{d}W_r\Bigg|_d^{\frac{1}{1-\beta}}\Bigg)\Bigg]\end{aligned} \tag{4.45}$$

where the inequality follows by the definition of J^θ and (4.42). The first term on the right-hand side of (4.45) is bounded from above using the Hölder's inequality and the Lipschitz property of b (see (4.34))

$$\begin{aligned}\mathsf{E}_x\Bigg[\sup_{\lambda\in[0,s]}\Bigg|\int_0^\lambda \big(b(X^\theta_r) - b(X^0_r)\big)\,\mathrm{d}r\Bigg|_d^{\frac{1}{1-\beta}}\Bigg] &\leq \mathsf{E}_x\Bigg[\sup_{\lambda\in[0,s]} \lambda^{\frac{1}{\beta}} \int_0^\lambda |b(X^\theta_r) - b(X^0_r)|_d^{\frac{1}{1-\beta}}\,\mathrm{d}r\Bigg] \\ &\leq \mathsf{E}_x\Bigg[T^{\frac{1}{\beta}} D_1^{\frac{1}{1-\beta}} \int_0^s \sup_{r\in[0,\lambda]} |J^\theta_r|_d^{\frac{1}{1-\beta}}\,\mathrm{d}r\Bigg];\end{aligned}$$

the second term of the right-hand side of (4.45) is bounded from above using the Doob's maximal inequality applied to $\tilde{W}$

$$\mathsf{E}_x\Bigg[\sup_{\lambda\in[0,s]} |\theta I_d \tilde{W}_\lambda|_d^{\frac{1}{1-\beta}}\Bigg] \leq \mathsf{E}_x\Bigg[d\Big(\tfrac{1}{\beta}\theta|Z_T|\Big)^{\frac{1}{1-\beta}}\Bigg] \leq \kappa_1 \theta^{\frac{1}{1-\beta}},$$

where Z_T is an independent normal random variable distributed as $\mathcal{N}(0,T)$ and $\kappa_1 = d\mathsf{E}_x[(\frac{1}{\beta}|Z_T|)^{\frac{1}{1-\beta}}]$; the last term on the right-hand side of (4.45) is bounded from above using [48, Cor. 2.5.11]

$$\begin{aligned}\mathsf{E}_x\Big[\sup_{\lambda\in[0,s]}\Big|\int_0^\lambda\big(\sigma(X_r^\theta)-\sigma(X_r^0)\big)\,\mathrm{d}W_r\Big|_d^{\frac{1}{1-\beta}}\Big]&\\ \le 2^{\frac{1}{2-2\beta}+2}\frac{\beta}{1-\beta}s^{\frac{1}{2-2\beta}-1}\mathsf{E}_x\Big[\int_0^s|\sigma(X_r^\theta)-\sigma(X_r^0)|_{d\times d'}^{\frac{1}{1-\beta}}\,\mathrm{d}r\Big]&\\ \le \kappa_2\mathsf{E}_x\Big[D_1^{\frac{1}{1-\beta}}\int_0^s\sup_{r\in[0,\lambda]}|J_r^\theta|_d^{\frac{1}{1-\beta}}\,\mathrm{d}r\Big],&\end{aligned}$$

where $\kappa_2=\kappa_2(T,\beta)=2^{\frac{5-4\beta}{2-2\beta}}\frac{\beta}{1-\beta}T^{\frac{2\beta-1}{2-2\beta}}$ and D_1 comes from (4.34). Using the three inequalities above in (4.45) we get

$$\mathsf{E}_x\Big[\sup_{\lambda\in[0,s]}|J_\lambda^\theta|_d^{\frac{1}{1-\beta}}\Big]\le 3^{\frac{\beta}{1-\beta}}\mathsf{E}_x\Big[D_1^{\frac{1}{1-\beta}}(1+\kappa_2)\int_0^s\sup_{r\in[0,\lambda]}|J_r^\theta|_d^{\frac{1}{1-\beta}}\,\mathrm{d}\lambda+\kappa_1\theta^{\frac{1}{1-\beta}}\Big].$$

By Gronwall's lemma we get for all $s\in[0,T-t]$

$$\mathsf{E}_x\Big[\sup_{\lambda\in[0,s]}|X_\lambda^\theta-X_\lambda^0|_d\Big]=\mathsf{E}_x\Big[\sup_{\lambda\in[0,s]}|J_\lambda^\theta|_d\Big]\le\theta^{\frac{1}{1-\beta}}K_7$$

with $K_7=K_7(D_1,d,\beta,T)$ independent of $(n,\nu)\in\mathcal{A}_t$.

Now we prove (4.44). Let $(n,\nu)\in\mathcal{A}_t$ such that $\mathsf{E}[\nu_{T-t}]\le K_6$ with K_6 as in Lemma 4.21. We have

$$\begin{aligned}\mathsf{E}_x\Big[\sup_{\lambda\in[0,s]}|X_\lambda^\theta|_d\Big]\le\mathsf{E}_x\Big[\sup_{\lambda\in[0,s]}\Big(\Big|\int_0^\lambda b(X_r^\theta)\,\mathrm{d}r\Big|_d+\Big|\int_0^\lambda\sigma(X_r^\theta)\,\mathrm{d}W_r\Big|_d&\\ +\Big|\int_0^\lambda\theta I_d\,\mathrm{d}\tilde{W}_r\Big|_d+\Big|\int_0^\lambda n_r\mathrm{d}\nu_r\Big|_d\Big)\Big].&\end{aligned}\tag{4.46}$$

We now estimate all the terms above separately. The first term on the right-hand side of (4.46):

$$\begin{aligned}\mathsf{E}_x\Big[\sup_{\lambda\in[0,s]}\Big|\int_0^\lambda b(X_r^\theta)\,\mathrm{d}r\Big|_d\Big]&\le\mathsf{E}_x\Big[D_2\int_0^s(1+\sup_{r\in[0,\lambda]}|X_r^\theta|_d)\,\mathrm{d}\lambda\Big]\\ &\le TD_2+\mathsf{E}_x\Big[D_2\int_0^s\sup_{r\in[0,\lambda]}|X_r^\theta|_d\,\mathrm{d}\lambda\Big];\end{aligned}$$

the second term on the right-hand side of (4.46):

$$\begin{aligned}\mathsf{E}_x\left[\sup_{\lambda\in[0,s]}\left|\int_0^\lambda \sigma(X_r^\theta)\,\mathrm{d}W_r\right|_d\right] &\le \mathsf{E}_x\left[\sup_{\lambda\in[0,s]}\left|\int_0^\lambda \sigma(X_r^\theta)\,\mathrm{d}r\right|_d^2\right]^{\frac{1}{2}}\\
&\le \kappa_3 \mathsf{E}_x\left[\int_0^s |(\sigma\sigma^\top)(X_r^\theta)|_{d\times d}\,\mathrm{d}\lambda\right]^{\frac{1}{2}}\\
&\le \kappa_3\left(1+\mathsf{E}_x\left[\int_0^s |(\sigma\sigma^\top)(X_r^\theta)|_{d\times d}\,\mathrm{d}\lambda\right]\right)\\
&\le \kappa_3\left(1+\mathsf{E}_x\left[D_2\int_0^s (1+\sup_{r\in[0,\lambda]}|X_r^\theta|_d)\,\mathrm{d}\lambda\right]\right)\\
&\le \kappa_3\left(1+TD_2+\mathsf{E}_x\left[D_2\int_0^s \sup_{r\in[0,\lambda]}|X_r^\theta|_d\,\mathrm{d}\lambda\right]\right),\end{aligned}$$

where the first inequality follows by Hölder's inequality, the second inequality by [48, Cor. 2.5.11] with $\kappa_3 = \kappa_3(T)$ and the last inequality follows by Assumption 4.16; the third term on the right-hand side of (4.46):

$$\mathsf{E}_x\left[\sup_{\lambda\in[0,s]}\left|\int_0^s \theta I_d\,\mathrm{d}\tilde{W}_\lambda\right|_d\right] \le d\theta\sqrt{8T} \le d\sqrt{8T} \tag{4.47}$$

using [48, Cor. 2.5.11]; the last term on the right-hand side of (4.46):

$$\mathsf{E}_x\left[\sup_{\lambda\in[0,s]}\left|\int_0^s n_r \mathrm{d}\nu_r\right|_d\right] \le \mathsf{E}_x[\nu_s]$$

by Lemma 4.21. Combining the four inequalities we have

$$\begin{aligned}\mathsf{E}_x\left[\sup_{\lambda\in[0,s]}|X_\lambda^\theta|_d\right] \le\, & TD_2(1+\kappa_3)+D_2(1+\kappa_3)\int_0^s \mathsf{E}_x\left[\sup_{r\in[0,\lambda]}|X_r^\theta|_d\right]\mathrm{d}\lambda\\
&+\kappa_3+d\sqrt{8T}+\mathsf{E}_x[\nu_s].\end{aligned}$$

By Gronwall's lemma, we get

$$\mathsf{E}_x\left[\sup_{s\in[0,T-t]}|X_s^\theta|_d\right] \le K_8\mathsf{E}_x[\nu_{T-t}],$$

with $K_8 = K_8(d, T, D_2)$. Analogously, we get the same estimates for X^0 without

(4.47), therefore we obtain:

$$\mathsf{E}_x\left[\sup_{s\in[0,T-t]}|X_s^0|_d\right] \le K_8\mathsf{E}_x[\nu_{T-t}].$$

with the same K_8. □

4.2.4 The Value of the Game

Theorem 4.24 *The game in (4.31) admits a value v. Moreover, for any compact $\mathcal{K} \subset \mathbb{R}^{d+1}_{0,T}$, there exists a $C_\mathcal{K} > 0$ such that*

$$\sup_{(t,x)\in\mathcal{K}} |u^\theta(t,x) - v(t,x)| \le C_\mathcal{K}\theta^{1-\beta} \tag{4.48}$$

for β as in (4.36) and u^θ is the value function described in Theorem 4.19.

Proof. Let u^θ be the value of the game described in Section 4.2.3. We introduce $\underline{u} := \liminf_{\theta\to 0} u^\theta$ and $\overline{u} := \limsup_{\theta\to 0} u^\theta$. If we prove that

$$\overline{u}(t,x) \le \underline{v}(t,x) \quad \text{and} \quad \underline{u}(t,x) \ge \overline{v}(t,x)$$

for all $(t,x) \in \mathbb{R}^{d+1}_{0,T}$, we get that $\underline{u}(t,x) = \overline{u}(t,x) = v(t,x) = \underline{v}(t,x) = \overline{v}(t,x)$ and Theorem 4.24 holds.

We prove first that $\underline{u}(t,x) \ge \overline{v}(t,x)$. For $(n,\nu) \in \mathcal{A}_t$ we consider the two processes whose dynamics read

$$\begin{aligned}
\mathrm{d}X_s^{[n,\nu],\theta} &= b(X_s^{[n,\nu],\theta})\mathrm{d}s + \sigma(X_s^{[n,\nu],\theta})\mathrm{d}W_s + \theta I_d \mathrm{d}\tilde{W}_s + n_s\mathrm{d}\nu_s,\\
\mathrm{d}X_s^{[n,\nu],0} &= b(X_s^{[n,\nu],0})\mathrm{d}s + \sigma(X_s^{[n,\nu],0})\mathrm{d}W_s + n_s\mathrm{d}\nu_s,
\end{aligned}$$

where we recall that $\tilde{W}$ is a d-dimensional Brownian motion independent from the original Brownian motion W and I_d is the d-dimensional identity matrix.

Let (n^θ,ν^θ) be an η-optimal control for $u^\theta(t,x)$ and let $\tau \in \mathcal{T}_t$ be an η-optimal stopping time for $\overline{v}(t,x)$. Let $X^\theta = X^{[n^\theta,\nu^\theta],\theta}$ and $X^0 = X^{[n^\theta,\nu^\theta],0}$ for notational simplicity. We have

$$
\begin{aligned}
&u^\theta(t,x) - \overline{v}(t,x) \\
&\quad \geq \mathcal{J}^\theta_{t,x}(n^\theta,\nu^\theta,\tau) - \mathcal{J}_{t,x}(n^\theta,\nu^\theta,\tau) - 2\eta \\
&\quad = \mathsf{E}_x\bigg[e^{-r\tau}(g(t+\tau,X^\theta_\tau) - g(t+\tau,X^0_\tau)) \\
&\qquad\qquad + \int_0^\tau e^{-rs}(h(t+s,X^\theta_s) - h(t+s,X^0_s))\,\mathrm{d}s\bigg] - 2\eta \qquad (4.49) \\
&\quad \geq -f(0)\mathsf{E}_x\Big[|X^\theta_\tau - X^0_\tau|_d\Big] \\
&\qquad - K_2\mathsf{E}_x\bigg[\int_0^{T-t}(1+|X^\theta_s|_d + |X^0_s|_d)^\beta |X^\theta_s - X^0_s|_d\,\mathrm{d}s\bigg] - 2\eta \\
&\quad \geq -f(0)\mathsf{E}_x\Big[\sup_{s\in[0,T-t]}|X^\theta_s - X^0_s|_d\Big] \\
&\qquad - K_2(T-t)\sup_{s\in[0,T-t]}\mathsf{E}_x\Big[(1+|X^\theta_s|_d + |X^0_s|_d)^\beta |X^\theta_s - X^0_s|_d\Big] - 2\eta,
\end{aligned}
$$

where $f(0)$ and K_2 come from (4.38) and (4.36), respectively. The second inequality follows by the Lipschitz property of g and h, and we extend the integral up to time $T-t$. The third inequality is by Fubini's theorem and taking supremum inside the time integral.

The first term in the last line of (4.49) is bounded from below by Hölder inequality:

$$
\begin{aligned}
&\sup_{s\in[0,T-t]}\mathsf{E}_x\Big[(1+|X^\theta_s|_d + |X^0_s|_d)^\beta |X^\theta_s - X^0_s|_d\Big] \\
&\qquad\qquad \leq \sup_{s\in[0,T-t]}\mathsf{E}_x\Big[(1+|X^\theta_s|_d + |X^0_s|_d)\Big]^\beta \mathsf{E}_x\Big[|X^\theta_s - X^0_s|_d^{\frac{1}{1-\beta}}\Big]^{1-\beta}.
\end{aligned}
$$

Applying (4.43) and (4.44) in (4.49), we get

$$
u^\theta(t,x) - \overline{v}(t,x) \geq -\Big(f(0) + K_2(T-t)(K_6K_8)^\beta\Big)(\theta K_7)^{1-\beta} - 2\eta.
$$

It means that when we pass to the limit inferior as $\theta \downarrow 0$ we have $\underline{u}(t,x) - \overline{v}(t,x) \geq -2\eta$ and by the arbitrariness of η we get $\underline{u}(t,x) \geq \overline{v}(t,x)$.

Let $\tau^\theta \in \mathcal{T}_t$ be an η-optimal stopping time for u^θ and let $(n,\nu) \in \mathcal{A}_t$ be an

η-optimal control for $\underline{v}(t,x)$. The estimates above still hold thanks also to Remark 4.22, thus we have

$$u^\theta(t,x) - \underline{v}(t,x) \le \Big(f(0) + K_2(T-t)(K_6K_8)^\beta\Big)(\theta K_7)^{1-\beta} + 2\eta.$$

Passing to the limit superior as $\theta \downarrow 0$ and then $\eta \downarrow 0$, we have $\overline{u}(t,x) \le \underline{v}(t,x)$ and the theorem is proved. It means that (4.48) holds with

$$C_{\mathcal{K}} := \max_{(t,x)\in\mathcal{K}} \big(f(0) + K_2T(K_6K_8)^\beta\big)K_7^{1-\beta},$$

where K_6 comes from Lemma 4.21 and it depends continuously on the initial state $x \in \mathcal{K}$. □

If we impose stronger continuity condition on h we can drop the growth condition on σ. This is done in the next simple corollary.

Corollary 4.25 *If the function h is Lipschitz in space uniformly in time, i.e.,*

$$|h(t,x) - h(t,x)| \le K_2|x-y|_d \quad \textit{for all } x,y \in \mathbb{R}^d,$$

with K_2 independent of t, then Theorem 4.24 holds without condition (4.35).

Proof. Since the function h is Lipschitz, then the bound in (4.49) becomes

$$u^\theta(t,x) - \overline{v}(t,x) \ge -(f(0) + K_2T)\mathsf{E}_x\left[\sup_{s\in[0,T-t]} |X_s^\theta - X_s^0|_d\right] - 2\eta.$$

Recalling that (4.43) is obtained using only the Lipschitz property of σ, we can repeat all the remaining steps in the proof of Theorem 4.24 after (4.49). □

We show the optimality of the stopping time defined in (4.41). We follow an approach that can be found for example in [13, Thm. 4.12].

Theorem 4.26 *For any given $(t,x) \in \mathbb{R}^{d+1}_{0,T}$ and any admissible control $(n,\nu) \in \mathcal{A}_t$, the stopping time $\tau_* = \tau_*(t,x;n,\nu) \in \mathcal{T}_t$ defined under P_x as*

$$\tau_* := \inf\left\{s \ge 0 \,\middle|\, v(t+s, X_s^{[n,\nu]}) = g(t+s, X_s^{[n,\nu]})\right\} \wedge (T-t) \tag{4.50}$$

is optimal for the stopper.

If we assume that the function h is uniformly Lipschitz in x, then the statement holds even if we drop (4.35) in Assumption 4.16.

The proof is similar to Theorem 4.13.

Proof. Let $(t,x) \in \mathbb{R}^{d+1}_{0,T}$, $(n,\nu) \in \mathcal{A}_t$ and for simplicity denote $X_s^{[n,\nu],\theta}$ and $X_s^{[n,\nu],0}$ by X_s^θ and X_s^0, respectively. Define

$$\tau_*^\theta := \inf\left\{s \geq 0 \,\middle|\, u^\theta(t+s, X_s^\theta) \leq g(t+s, X_s^\theta)\right\} \wedge (T-t).$$

By the convergence in (4.43), we have that there exists a strictly decreasing subsequence $(\theta_k)_{k\in N} \subset (0,1)$ such that $X^{\theta_k} \to X^0$ almost surely as $k \to \infty$, i.e., there exists an $\Omega_0 \subseteq \Omega$ with $\mathsf{P}(\Omega_0) = 1$ such that for all $\omega \in \Omega_0$ we have $X^{\theta_k}(\omega) \to X^0(\omega)$ uniformly as $k \to \infty$. Thus, for all $\Gamma > 0$, there exists $k' = k'(\omega)$ such that for all $k \geq k'$ we have

$$|X_s^{\theta_k}(\omega) - X_s^0(\omega)|_d \leq \Gamma.$$

We prove now that

$$\liminf_{k\to\infty} \tau_*^{\theta_k} \geq \tau_*,$$

with τ_* as in (4.50). Fix $\omega \in \Omega_0$. If $\tau_*(\omega) = 0$, then (4.27) holds. If $\tau_*(\omega) > 0$, we pick $\delta > 0$ such that $\delta < \tau_*(\omega)$. By the same arguments of the proof of the optimality of τ_* in the proof of Theorem 3.33, we have for almost every $\omega \in \Omega_0$ that $\tau_0 = \tau_*$ and $\Delta\nu_{\tau_*} = 0$, i.e., $X^0_{\tau_0-} = X^0_{\tau_*-} = X^0_{\tau_*}$, where

$$\tau_0 = \inf\{s \geq 0 | v(t+s, X^0_{s-}) - g(t+s, X^0_{s-}) = 0\}.$$

We have that the sequence $(\tau_\eta)_{\eta\in(0,1)}$ with

$$\tau_\eta := \inf\{s \geq 0 | v(t+s, X_s^0) - g(t+s, X_s^0) \leq \eta\}$$

admits a strictly increasing subsequence $(\tau_{\eta_m})_{m\in\mathbb{N}}$, with $\eta_m \downarrow 0$ as $m \to \infty$, for almost every $\omega \in \Omega_0$, i.e., there exists an $\Omega_0' \subset \Omega_0$ with $\mathsf{P}(\Omega_0') = 1$ where

$\tau_{\eta_m}(\omega) \uparrow \tau_0(\omega)$ for all $\omega \in \Omega_0'$ as $m \to \infty$ and

$$\lim_{m\to\infty} X^0_{\tau_{\eta_m}} = X^0_{\tau_0 -}.$$

This limit means that there exists a constant $C_\omega > 0$ such that

$$\inf_{0\le s\le\delta} \big(v(t+s, X^0_s(\omega)) - g(t+s, X^0_s(\omega))\big) \ge C_\omega,$$

otherwise we would have that for all $\eta_m > 0$,

$$\inf_{0\le s\le\delta} \big(v(t+s, X^0_s(\omega)) - g(t+s, X^0_s(\omega))\big) < \eta_m,$$

and the inequality would imply $\tau_{\eta_m} \le \delta$ for all $m \in \mathbb{N}$. Since $\tau_{\eta_m} \uparrow \tau_* > \delta$, we reach a contradiction.

Moreover, since $(n,\nu) \in \mathcal{A}_t$ is fixed, we have that the second moment of the controlled process is finite and thus there exists a compact $\mathcal{K}_\omega \subset \mathbb{R}^{d+1}_{0,T}$ such that the trajectory of $X^0(\omega)$ lies in it

$$\{(t+s, X^0_s(\omega)) : s \in [0,\delta]\} \subset \mathcal{K}_\omega.$$

By Theorem 4.24, we have that there exists a k'' such that for all $k \ge k''$, we have

$$\sup_{(t,x)\in\mathcal{K}_\omega} |u^{\theta_k}(t,x) - v(t,x)| < \frac{C_\omega}{4}.$$

Thus we have that for $k \ge k' \vee k''$

$$\begin{aligned}
|u^{\theta_k}(t+s, X^{\theta_k}_s(\omega)) - v(t+s, X^0_s(\omega))| &\le |u^{\theta_k}(t+s, X^{\theta_k}_s(\omega)) - u^{\theta_k}(t+s, X^0_s(\omega))| \\
&\quad + |u^{\theta_k}(t+s, X^0_s(\omega)) - v(t+s, X^0_s(\omega))| \\
&\le f(0)|X^{\theta_k}_s(\omega) - X^0_s(\omega)| + \frac{C_\omega}{4} \\
&\le f(0)\Gamma + \frac{C_\omega}{4},
\end{aligned}$$

where we used $|\nabla u^\theta(t,x)|_d \le f(t) \le f(0)$ for all θ (see Lemma 4.20). Taking Γ

such that $f(0)\Gamma < \frac{C_\omega}{4}$, it follows

$$\inf_{0\le s\le \delta} |u^{\theta_k}(t+s, X^{\theta_k}_s(\omega)) - g(t+s, X^0_s(\omega)| \ge \frac{C_\omega}{2}.$$

This means that $\tau^{\theta_k}_* > \delta$ for $k \ge k' \vee k''$ and

$$\liminf_{k\to\infty} \tau^{\theta_k}_* \ge \delta.$$

Sending $\delta \uparrow \tau_*$ we have that (4.27) holds and we obtain

$$\lim_{k\to\infty} \tau^{\theta_k}_* \wedge \tau_* = \tau_*, \tag{4.51}$$

with τ_* as in (4.26). We can repeat the same arguments of the proof of Theorem 4.13. For $\eta > 0$ we consider

$$\tau^{\theta_k}_\eta := \inf\{s \ge 0 | u^{\theta_k}(t+s, X^{\theta_k}_s) \le g(t+s, X^{\theta_k}_s) + \eta\},$$

where u^{τ_k} is the value of the game (4.40); we have that $\tau^{\theta_k}_\eta \uparrow \tau^{\theta_k}_*$ as $\eta \downarrow 0$. By the first line of (3.124) with g and τ_η therein replaced by u^{θ_k} and $\tau^{\theta_k}_\eta \wedge \tau_*$, respectively, we have

$$u^{\theta_k}(t,x) \le \eta + \mathsf{E}_x\Bigg[e^{-r(\tau^{\theta_k}_\eta \wedge \tau_*)} u^{\theta_k}(t+\tau^{\theta_k}_\eta \wedge \tau_*, X^{\theta_k}_{\tau^{\theta_k}_\eta \wedge \tau_*}) + \int_0^{\tau^{\theta_k}_\eta \wedge \tau_*} e^{-rs} h(t+s, X^{\theta_k}_s)\, \mathrm{d}s$$
$$+ \int_0^{\tau^{\theta_k}_\eta \wedge \tau_*} e^{-rs} f(t+s)\, \mathrm{d}\nu_s\Bigg].$$

Sending $\eta \downarrow 0$, by dominated convergence theorem we obtain

$$u^{\theta_k}(t,x) \le \mathsf{E}_x\Bigg[e^{-r(\tau^{\theta_k}_* \wedge \tau_*)} u^{\theta_k}(t+\tau^{\theta_k}_* \wedge \tau_*, X^{\theta_k}_{\tau^{\theta_k}_* \wedge \tau_*}) + \int_0^{\tau^{\theta_k}_* \wedge \tau_*} e^{-rs} h(t+s, X^{\theta_k}_s)\, \mathrm{d}s$$
$$+ \int_0^{\tau^{\theta_k}_* \wedge \tau_*} e^{-rs} f(t+s)\, \mathrm{d}\nu_s\Bigg].$$

Finally, sending $k \to \infty$ we have that

$$v(t,x) \le \lim_{k\to\infty} \mathsf{E}_x\Big[e^{-r(\tau_*^{\theta_k}\wedge\tau_*)}u^{\theta_k}(t+\tau_*^{\theta_k}\wedge\tau_*, X^{\theta_k}_{\tau_*^{\theta_k}\wedge\tau_*})+\int_0^{\tau_*^{\theta_k}\wedge\tau_*} e^{-rs}h(t+s, X^{\theta_k}_s)\,\mathrm{d}s \\ + \int_0^{\tau_*^{\theta_k}\wedge\tau_*} e^{-rs}f(t+s)\,\mathrm{d}\nu_s\Big]$$

Using that $(n,\nu)\in\mathcal{A}_t$, X^{θ_k} converges uniformly to X^0 almost surely, f is bounded, the functions g and h have linear growth, we have that dominated convergence theorem applies and we obtain by (4.51) that

$$v(t,x) \le \mathsf{E}_x\Big[e^{-r\tau_*}g(t+\tau_*, X^0_{\tau_*})+\int_0^{\tau_*} e^{-rs}h(t+s, X^0_s)\,\mathrm{d}s+\int_0^{\tau_*} e^{-rs}f(t+s)\,\mathrm{d}\nu_s\Big].$$

Thus, the strategy τ_* is optimal for the stopper.

The last statement of Theorem 4.26 holds because Corollary 4.25 allows the use of Theorem 4.24, and we have that the proof above can be repeated under the assumptions that h is uniformly Lipschitz in x and σ does not satisfy (4.35) in Assumption 4.16. □

The next remark shows that the value function v of the game does not change if we impose the controller to use only absolutely continuous controls.

Remark 4.27: Remark 3.34 says that

$$u^\theta(t,x) = \inf_{(n,\nu)\in\mathcal{A}_t^\circ}\sup_{\tau\in\mathcal{T}_t}\mathcal{J}^\theta_{t,x}(n,\nu,\tau) = \sup_{\tau\in\mathcal{T}_t}\inf_{(n,\nu)\in\mathcal{A}_t^\circ}\mathcal{J}^\theta_{t,x}(n,\nu,\tau),$$

where $\mathcal{A}_t^\circ := \{(n,\nu)\in\mathcal{A}_t|\nu \text{ is absolutely continuous}\}$. We introduce the lower and upper value functions of the games with absolutely continuous admissible controls:

$$\underline{v}^\circ(t,x) := \sup_{\tau\in\mathcal{T}_t}\inf_{(n,\nu)\in\mathcal{A}_t^\circ}\mathcal{J}_{t,x}(n,\nu,\tau) \quad \text{and} \quad \overline{v}^\circ(t,x) := \inf_{(n,\nu)\in\mathcal{A}_t^\circ}\sup_{\tau\in\mathcal{T}_t}\mathcal{J}_{t,x}(n,\nu,\tau).$$

Repeating the proof of Theorem 4.24 replacing $\underline{v}$ and $\overline{v}$ therein with $\underline{v}^\circ$ and $\overline{v}^\circ$, respectively, we get that $\lim_{\theta\to 0} u^\theta(t,x) = v^0(t,x)$ and

$$v^0(t,x) = \inf_{(n,\nu)\in\mathcal{A}_t^\circ} \sup_{\tau\in\mathcal{T}_t} \mathcal{J}^\theta_{t,x}(n,\nu,\tau) = \sup_{\tau\in\mathcal{T}_t} \inf_{(n,\nu)\in\mathcal{A}_t^\circ} \mathcal{J}^\theta_{t,x}(n,\nu,\tau).$$

Since the limit is unique, then we have that $v^0 = v$ with v as in Theorem 4.24. ■

Appendix

In this appendix, we store results that we prefer to keep separate from the previous chapters because they are more auxiliary tools for us than results on ZSGs. Each section stores results from a correspondent chapter.

A Chapter 1

In this section we give the proof of the compact embedding between parabolic Sobolev spaces and parabolic Hölder spaces (see (1.6)).

A.1 Parabolic Sobolev Embedding

Let $d \in \mathbb{N}$ and $a : \mathbb{R} \to \mathbb{R}^d$ be uniformly elliptic with θ as coefficient of ellipticity, i.e., there exists a θ such that

$$\theta^{-1}|\zeta|_d^2 \geq \langle \zeta, a(t)\zeta \rangle \geq \theta|\zeta|_d^2 \tag{A.1}$$

for all $(t, \zeta) \in \mathbb{R}^{1+d}$. We also introduce three objects related to a $Q \subset \mathbb{R}^d$:

- the diameter of Q as $\mathrm{diam}(Q) := \sup_{x,y\in Q} |x-y|_Q$ where $|\cdot|_Q$ is the euclidean norm on Q;
- the size of Q as $|Q| := \displaystyle\int_Q \mathrm{d}z$;
- the mean value of a function u inside Q as $u_Q := \dfrac{1}{|Q|}\displaystyle\int_Q u(z)\mathrm{d}z$.

We give now a Theorem without proof which can be found in the correspondent book [47].

Theorem A.1 ([47, Thm. 10.2.5]) *Let $p \in [0,\infty)$ and let $Q \subset \mathbb{R}^d$ be a convex bounded domain and $u \in W^{1,p}(Q)$. Then*

$$\int_Q \int_Q |u(y) - u(x)|^p \, \mathrm{d}x\mathrm{d}y \leq 2^{d+1} \operatorname{diam}(Q)^p |Q| \int_Q |\nabla u(x)|_d^p \, \mathrm{d}x.$$

The next Lemma can be found with its proof in [47, Lemma 4.2.1], we give below the statement and its proof in detail.

Lemma A.2 ([47, Lemma 4.2.1]) *Let $p \in [1,\infty)$, $\rho \in (0,\infty)$, $u \in C^\infty(\mathbb{R}^{d+1})$, $f^i \in C^\infty(\mathbb{R}^{d+1})$ for $i = 1, \dots d$ and $g \in C^\infty(\mathbb{R}^{d+1})$. Assume that the function u satisfies*

$$u_t(t,x) + \sum_{i,j=1}^d a_{ij}(t) u_{x_i x_j}(t,x) = \sum_{i=1}^d f^i_{x_i}(t,x) + g(t,x),$$

where $a(t) = (a_{ij}(t))_{ij} \in \mathbb{R}^{d\times d}$ is symmetric for all t and elliptic with θ as coefficient of ellipticity (see (A.1)). Then

$$\int_{Q_\rho} |u(t,x) - u_{Q_\rho}|^p \, \mathrm{d}t\mathrm{d}x \leq N_1 \rho^p \int_{Q_\rho} \sum_{i=1}^d |u_{x_i}(t,x)|^p + |f^i(t,x)|^p + \rho^p |g(t,x)|^p \, \mathrm{d}t\mathrm{d}x, \tag{A.2}$$

where Q_ρ is any cylinder of length ρ^2 with basis $B_\rho \subset \mathbb{R}^d$, i.e., $Q_\rho := (0,T) \times B_\rho$ and $N_1 = N_1(d,\theta,p)$.

Proof. We first prove that the theorem holds with $\rho = 1$. Then we prove it for all ρ. Let $c \in \mathbb{R}$. Using the inequality (4.42), we have that:

$$\begin{aligned} \int_{Q_1} |u(t,x) - u_{Q_1}|^p \, \mathrm{d}t\mathrm{d}x &\leq 2^{p-1} \Big(\int_{Q_1} |u(t,x) - c|^p + |c - u_{Q_1}|^p \mathrm{d}t\mathrm{d}x \Big) \\ &= 2^{p-1} \Big(\int_{Q_1} |u(t,x) - c|^p \mathrm{d}t\mathrm{d}x + |Q_1||c - u_{Q_1}|^p \Big) =: A_1. \end{aligned} \tag{A.3}$$

Using the definition of u_{Q_1} and Hölder's inequality we get

$$\begin{aligned} A_1 &= 2^{p-1}\Big(\int_{Q_1}|u(t,x)-c|^p\mathrm{d}t\mathrm{d}x + |Q_1|\Big|\frac{1}{|Q_1|}\int_{Q_1} c-u(t,x)\mathrm{d}t\mathrm{d}x\Big|^p\Big) \\ &\le 2^{p-1}\Big(\int_{Q_1}|u(t,x)-c|^p\mathrm{d}t\mathrm{d}x + |Q_1|\frac{1}{|Q_1|^p}|Q_1|^{p-1}\int_{Q_1}|c-u(t,x)|^p\mathrm{d}t\mathrm{d}x\Big) \\ &= 2^p\int_{Q_1}|u(t,x)-c|^p\mathrm{d}t\mathrm{d}x. \end{aligned} \tag{A.4}$$

Consider a $\xi \in C_c^\infty(B_1)$ with unit integral. We define a function $\bar{u}$ of time $t \in [0,1]$ as

$$\bar{u}(t) := \int_{B_1}\xi(x)u(t,x)\mathrm{d}x,$$

it can be viewed as the mean value of u under the measure whose density is ξ.

We compute the L^p-norm of the difference between u and $\bar{u}$ for each fixed t and using the Hölder's inequality we get:

$$\begin{aligned} \int_{B_1}|u(t,y)-\bar{u}(t)|^p\mathrm{d}y &= \int_{B_1}\Big|u(t,y)-\int_{B_1}\xi(x)u(t,x)\mathrm{d}x\Big|^p\mathrm{d}y \\ &= \int_{B_1}\Big|\int_{B_1}(u(t,y)-u(t,x))\xi(x)\mathrm{d}x\Big|^p\mathrm{d}y \\ &\le |B_1|\Big(\int_{B_1}|\xi(x)|^{\frac{p}{p-1}}\mathrm{d}x\Big)^{p-1}\int_{B_1}\int_{B_1}|u(t,y)-u(t,x)|^p\,\mathrm{d}x\mathrm{d}y \\ &\le |B_1|\Big(\int_{B_1}|\xi(x)|^{\frac{p}{p-1}}\mathrm{d}x\Big)^{p-1}2^{d+1}2^p|B_1|\int_{B_1}|\nabla u(t,y)|_d^p\,\mathrm{d}y \\ &= C_1\int_{B_1}|\nabla u(t,y)|_d^p\mathrm{d}y \end{aligned} \tag{A.5}$$

where for $p=1$ we consider the supremum of ξ instead of the $L^{\frac{p}{p-1}}$-norm. In the last inequality, we used Theorem A.1 and $C_1 = C_1(\xi,d,p)$. From (A.3) and (A.4) we have

$$
\begin{aligned}
&2^p \int_{Q_1} |u(t,x) - c|^p \,\mathrm{d}t\mathrm{d}x \\
&\quad \le 2^{2p-1}\Big(\int_{Q_1} |u(t,x) - \bar{u}(t)|^p \mathrm{d}t\mathrm{d}x + |B_1| \int_0^1 |\bar{u}(t) - c|^p \mathrm{d}t \Big) \\
&\quad \le 2^{2p-1}\Big(C_1 \int_{Q_1} |\nabla u(t,x)|_d^p \,\mathrm{d}t\mathrm{d}x + |B_1| \int_0^1 |\bar{u}(t) - c|^p \,\mathrm{d}t \Big).
\end{aligned}
$$

where we used (A.5) in the second inequality. We concentrate now on the last term above, we define $c \in \mathbb{R}$ as $c := \int_0^1 \bar{u}(s)\mathrm{d}s$ then

$$
\begin{aligned}
\int_0^1 |\bar{u}(t) - c|^p \,\mathrm{d}t &= \int_0^1 \Big| \bar{u}(t) - \int_0^1 \bar{u}(s)\,\mathrm{d}s \Big|^p \,\mathrm{d}t \\
&= \int_0^1 \Big| \int_0^1 \bar{u}(t) - \bar{u}(s)\,\mathrm{d}s \Big|^p \,\mathrm{d}t \\
&\le \int_0^1 \int_0^1 |\bar{u}(t) - \bar{u}(s)|^p \,\mathrm{d}s\mathrm{d}t.
\end{aligned}
$$

Using Theorem A.1, where $\mathrm{diam}([0,1]) = 1$ and $d = 1$, we get

$$
\begin{aligned}
\int_0^1 |\bar{u}(t) - c|^p \mathrm{d}t &\le 4 \int_0^1 |\bar{u}_t(s)|^p \,\mathrm{d}s \qquad \text{(A.6)} \\
&= 4 \int_0^1 \Big| \int_{B_1} \xi(x) u_t(s,x)\mathrm{d}x \Big|^p \,\mathrm{d}s
\end{aligned}
$$

where the equality holds because we can pass the derivative under the integral.

By assumption we can write

$$
u_t(s,x) = -\sum_{i,j=1}^d a_{ij}(s) u_{x_i x_j}(s,x) + \sum_{i=1}^d f^i_{x_i}(s,x) + g(s,x),
$$

thus we can substitute the equation above in (A.6) and obtain

$$\int_0^1 |\bar{u}(t) - c|^p \mathrm{d}t$$
$$\leq 4 \int_0^1 \left| \int_{B_1} \xi(x) \Big(- \sum_{i,j=1}^d a_{ij}(s) u_{x_i x_j}(s,x) + \sum_{i=1}^d f^i_{x_i}(s,x) + g(s,x) \Big) \, \mathrm{d}x \right|^p \mathrm{d}s =: A_2.$$

Integrating by parts A_2 and using that ξ has compact support on B_1, we get

$$A_2 = 4 \int_0^1 \left| \int_{B_1} \sum_{i,j=1}^d a_{ij}(s) u_{x_i}(s,x) \xi_{x_j}(x) - \sum_{i=1}^d f^i(s,x) \xi_{x_i} + \xi(x) g(s,x) \mathrm{d}x \right|^p \mathrm{d}s.$$

Using (4.42) and Hölder's inequality with p and $q := \frac{p}{p-1}$ we have

$$\begin{aligned} A_2 \leq C_2 \Big(& \sum_{i,j=1}^d \Big(\int_{Q_1} |a_{ij}(s) \xi_{x_j}(x)|^q \, \mathrm{d}s \mathrm{d}x \Big)^{p-1} \Big(\int_{Q_1} |u_{x_i}(s,x)|^p \, \mathrm{d}s \mathrm{d}x \Big) \\ & + \sum_{i=1}^d \Big(\int_{Q_1} |\xi_{x_i}(x)|^q \, \mathrm{d}s \mathrm{d}x \Big)^{p-1} \Big(\int_{Q_1} |f^i(s,x)|^p \, \mathrm{d}s \mathrm{d}x \Big) \\ & + \Big(\int_{Q_1} |\xi(x)|^q \, \mathrm{d}s \mathrm{d}x \Big)^{p-1} \int_{Q_1} |g(s,x)|^p \, \mathrm{d}s \mathrm{d}x \Big) \\ \leq C_3 & \Big(\int_{Q_1} \sum_{i=1}^d |u_{x_i}(s,x)|^p + \sum_{i=1}^d |f^i(s,x)|^p + |g(s,x)|^p \, \mathrm{d}x \mathrm{d}s \Big), \end{aligned}$$

where $C_2 := 4(d^2 + d + 1)^{p-1}$ and $C_3 = C_3(d,p,\theta)$ is C_2 multiplied by terms independent of u, f and g. The result holds with $N_1 := 2^{2p-1}(C_1 + |B_1| C_3)$.

We now prove the result for a generic $\rho \in (0,\infty)$. Consider the substitution $v(t,x) := u(\rho^2 t, \rho x)$, then

$$\begin{aligned} v_t(t,x) + \sum_{i,j=1}^d a_{ij}(\rho^2 t) v_{x_i x_j}(t,x) = & \rho^2 \Big(u_t(\rho^2 t, \rho x) + \sum_{i,j=1}^d a_{ij}(\rho^2 t) u_{x_i x_j}(\rho^2 t, \rho x) \Big) \\ = & \rho^2 \Big(\frac{1}{\rho} \sum_{i=1}^d f^i_{x_i}(\rho^2 t, \rho x) + g(\rho^2 t, \rho x) \Big) \\ = & \sum_{i=1}^d \rho f^i_{x_i}(\rho^2 t, \rho x) + \rho^2 g(\rho^2 t, \rho x), \end{aligned}$$

where $\tilde{u}_{Q_1}$ is the mean of $u(\rho^2 t, \rho x)$ inside Q_1. Considering the left-hand side of (A.2), we get

$$\int_{Q_\rho} |u(t,x) - u_{Q_\rho}|^p \,\mathrm{d}t\mathrm{d}x = \rho^{(d+2)p} \int_{Q_1} |u(\rho^2 t, \rho x) - \tilde{u}_{Q_1}|^p \,\mathrm{d}t\mathrm{d}x$$
$$= \rho^{(d+2)p} \int_{Q_1} |v(t,x) - v_{Q_1}|^p \,\mathrm{d}t\mathrm{d}x =: A_3$$

By (A.2) with $\rho = 1$ we get

$$\begin{aligned}
A_3 &\le \rho^{(d+2)p} N_1 \int_{Q_1} \sum_{i=1}^{d} |v_{x_i}(t,x)|^p + |\rho f^i(\rho^2 t, \rho x)|^p + |\rho^2 g(\rho^2 t, \rho x)|^p \,\mathrm{d}t\mathrm{d}x \\
&= \rho^{(d+2)p} N_1 \int_{Q_1} \sum_{i=1}^{d} \rho^p |u_{x_i}(\rho^2 t, \rho x)|^p + \rho^p |f^i(\rho^2 t, \rho x)|^p + \rho^{2p} |g(\rho^2 t, \rho x)|^p \,\mathrm{d}t\mathrm{d}x \\
&= \rho^{(d+2)p} \rho^p N_1 \int_{Q_1} \sum_{i=1}^{d} |u_{x_i}(\rho^2 t, \rho x)|^p + |f^i(\rho^2 t, \rho x)|^p + \rho^p |g(\rho^2 t, \rho x)|^p \,\mathrm{d}t\mathrm{d}x \\
&= N_1 \rho^p \int_{Q_\rho} \sum_{i=1}^{d} |u_{x_i}(t,x)|^p + |f^i(t,x)|^p + \rho^p |g(t,x)|^p \,\mathrm{d}t\mathrm{d}x,
\end{aligned}$$

and the Theorem holds. □

Similarly, we present [47, Lemma 4.2.2] with its straigthforward proof.

Lemma A.3 *Let $p \in [1, \infty)$. There exists a constant $N_2 = N_2(d, p)$ such that for any $\rho \in (0, \infty)$ and $u \in C^\infty(\mathbb{R}^{d+1})$ we have*

$$\begin{aligned}
&\sum_{i=1}^{d} \int_{Q_\rho} |u_{x_i}(t,x) - (u_{x_i})_{Q_\rho}|^p \,\mathrm{d}t\mathrm{d}x \\
&\qquad \le N_2 \rho^p \sum_{i,j=1}^{d} \int_{Q_\rho} |u_{x_i x_j}(t,x)|^p + |u_t(t,x)|^p \,\mathrm{d}t\mathrm{d}x.
\end{aligned} \tag{A.7}$$

Proof. We define the auxiliary elliptic operator $L := \sum_{i,j=1}^{d} \delta^{i,j} \frac{\partial}{\partial x_i \partial x_j}$ and $f :=$

$u_t + Lu$. If we differentiate f with respect to x_i we obtain

$$\begin{aligned} f_{x_i} =& \Big(u_t(s,x) + \sum_{j=1}^{d} u_{x_j,x_j} \Big)_{x_i} \\ =& (u_{x_i}(s,y))_t + L(u_{x_i}(s,x)). \end{aligned}$$

We can then apply Lemma A.2 with $g \equiv 0$ and then (A.7) follows. □

Thanks to these preliminary results we are able to solve [47, Exercise 10.1.14]. This is the Sobolev embedding for Parabolic Spaces.

Lemma A.4 ([47, Exercise 10.1.14]) *The space $W^{1,2,p}_{\ell oc}(\mathbb{R}^{d+1})$ is a subset of the space $C^{0,1,\alpha}_{\ell oc}(\mathbb{R}^{d+1})$ for $p > d+2$ with $\alpha = 1 - \frac{d+2}{p}$. Moreover the inclusion is compact for $\alpha < 1 - \frac{d+2}{p}$.*

Proof. Let $u \in W^{1,2,p}_{\ell oc}(\mathbb{R}^{d+1})$, we have $\partial_t u, \partial_{x_i} u \in L^p_{\ell oc}(\mathbb{R}^{d+1})$ for $i = 1, \dots d$, and it follows by standard argument that u is locally bounded and α-Hölder for $\alpha \leq 1 - \frac{d+1}{p}$ with respect to the $d+1$ euclidean distance, and thus it is also locally Hölder for $\alpha \leq 1 - \frac{d+2}{p}$ (see [25, Thm. 5.6.4]). This is sufficient to conclude that u is also α-Hölder continuous with respect to the parabolic distance. Indeed, consider a compact $\mathcal{K} \subset \mathbb{R}^{d+1}$, we have that for all $(t,x), (s,y) \in \mathcal{K}$

$$(|t-s|^{\frac{1}{2}} + |x-y|_d) \geq \begin{cases} (|t-s| + |x-y|_d^2)^{\frac{1}{2}}, & \text{if } |t-s| < 1, \\ (2\,\mathrm{diam}(\mathcal{K}))^{-1}(|t-s| + |x-y|_d^2)^{\frac{1}{2}}, & \text{if } |t-s| \geq 1. \end{cases}$$

For all $\alpha \in (0,1)$, we have on the left-hand side below the α-Hölder condition with respect to the parabolic $d+1$ norm:

$$\begin{aligned} &\sup_{\substack{(t,x),(s,y)\in\mathcal{K} \\ (t,x)\neq(s,y)}} \frac{|u(t,x) - u(s,y)}{(|t-s|^{\frac{1}{2}} + |x-y|_d)^{\alpha}} \\ &\quad\leq \sup_{\substack{(t,x),(s,y)\in\mathcal{K} \\ (t,x)\neq(s,y) \\ |t-s|<1}} \frac{|u(t,x) - u(s,y)}{(|t-s| + |x-y|_d^2)^{\frac{\alpha}{2}}} + \sup_{\substack{(t,x),(s,y)\in\mathcal{K} \\ |t-s|\geq 1}} \frac{2\,\mathrm{diam}(\mathcal{K})|u(t,x) - u(s,y)}{(|t-s| + |x-y|_d^2)^{\frac{\alpha}{2}}} \\ &\quad\leq (1 + 2\,\mathrm{diam}(\mathcal{K})) \sup_{\substack{(t,x),(s,y)\in\mathcal{K} \\ (t,x)\neq(s,y)}} \frac{|u(t,x) - u(s,y)}{(|t-s| + |x-y|_d^2)^{\frac{\alpha}{2}}}, \end{aligned}$$

which concludes the first part.

It remains to prove the α-Hölder continuity of the spatial derivatives u_{x_i} for $i = 1, \dots d$.

Define the following cylinder $Q_\rho(t,x) \coloneqq (t, t+\rho^2) \times B_\rho(x)$. First we prove that there exists M and $\alpha \in (0,1]$ such that for all $(t,x),(s,y) \in \mathbb{R}^{d+1}$ and $\rho > 0$

$$\sum_{i=1}^{d} A_i \leq M(|t-s|^{\frac{1}{2}} + |x-y|_d)^{2d+4+\alpha}, \qquad \text{with}$$
$$A_i \coloneqq \int_{Q_\rho(t,x)} \int_{Q_\rho(s,y)} |u_{x_i}(r_1,p_1) - u_{x_i}(r_2,p_2)| \,\mathrm{d}r_1 \mathrm{d}p_1 \mathrm{d}r_2 \mathrm{d}p_2 \quad \text{for } i = 1, \dots d,$$

whenever ρ is such that $4\rho \leq |t-s|^{\frac{1}{2}} + |x-y|_d$. For notational simplicity we denote $\bar{d} \coloneqq |t-s|^{\frac{1}{2}} + |x-y|_d$. Notice that the volume of $|Q_\rho| = \rho^{d+2}|B_1|$. Thus, we get

$$\begin{aligned} A_i &\leq \int_{Q_\rho(t,x)} \int_{Q_\rho(s,y)} |u_{x_i}(r_1,p_1) - c| + |u_{x_i}(r_2,p_2) - c| \,\mathrm{d}r_1 \mathrm{d}p_1 \mathrm{d}r_2 \mathrm{d}p_2 \\ &\leq |B_1|\rho^{d+2} \Big(\int_{Q_\rho(t,x)} |u_{x_i}(r_1,p_1) - c| \,\mathrm{d}r_1 \mathrm{d}p_1 + \int_{Q_\rho(s,y)} |u_{x_i}(r_2,p_2) - c| \,\mathrm{d}r_2 \mathrm{d}p_2 \Big), \end{aligned} \tag{A.8}$$

where $c \in \mathbb{R}$ is a constant that we chose later.

We can assume with no loss of generality that $t < s$ and we can estimate (A.8) integrating on a bigger domain. Using $Q_{2\bar{d}}(t,x)$, we have that $Q_\rho(t,x) \cup Q_\rho(s,y) \subset Q_{2\bar{d}}(t,x)$ because $s + \rho^2 = s - t + t + \rho^2 \leq \bar{d}^2 + t + \rho^2 \leq t + 2\bar{d}^2$ and a similar inequality holds for the spatial ball $B_\rho(y)$. We define c as

$$c \coloneqq \int_{Q_{2\bar{d}}(t,x)} u_{x_i}(r_1,p_1) \,\mathrm{d}r_1 \mathrm{d}p_1.$$

Applying Lemma A.3 to the right-hand side of (A.8) where now we are integrating on $Q_{2\bar{d}}(t,x)$, we obtain that

$$\sum_{i=1}^{d} A_i \leq 2N_2|B_1|\rho^{d+2}(2\bar{d}) \Big(\sum_{i,j=1}^{d} \int_{Q_{2\bar{d}}(t,x)} |u_{x_i x_j}(r_1,p_1)| + |u_t(r_1,p_1)| \,\mathrm{d}r_1 \mathrm{d}p_1 \Big).$$

The sum on the right-hand side above can be estimated using Hölder's inequality with p and $\frac{p}{p-1}$, thus we obtain

$$\begin{aligned}\sum_{i,j=1}^{d}\int_{Q_{2\bar{d}}(t,x)}&|u_{x_ix_j}(r_1,p_1)|+|u_t(r_1,p_1)|\,\mathrm{d}r_1\mathrm{d}p_1\\&\leq|Q_{2\bar{d}}|^{\frac{p-1}{p}}\Big(\sum_{i,j=1}^{d}\Big|\int_{Q_{2\bar{d}}(t,x)}|u_{x_ix_j}(r_1,p_1)|^p\,\mathrm{d}r_1\mathrm{d}p_1\Big|^{\frac{1}{p}}\\&\qquad+\Big|\int_{Q_{2\bar{d}}(t,x)}|u_t(r_1,p_1)|^p\,\mathrm{d}r_1\mathrm{d}p_1\Big|^{\frac{1}{p}}\Big)\\&\leq\big(|B_1|(2\bar{d})^{d+2}\big)^{\frac{p-1}{p}}\|u\|_{W^{1,2,p}(Q_{2\bar{d}}(t,x))}\\&=N_3\bar{d}^{(d+2)\frac{p-1}{p}}\|u\|_{W^{1,2,p}(Q_{2\bar{d}}(t,x))},\end{aligned}$$

where we denoted $N_3 \coloneqq (|B_1|2^{d+2})^{\frac{p-1}{p}}$. It means that

$$\begin{aligned}\sum_{i=1}^{d}A_i\leq&2N_2|B_1|\rho^{d+2}(2\bar{d})N_3\bar{d}^{(d+2)\frac{p-1}{p}}\|u\|_{W^{1,2,p}(Q_{2\bar{d}}(t,x))}\\=&N_4\|u\|_{W^{1,2,p}(Q_{2\bar{d}}(t,x))}\bar{d}^{2d+4+1-\frac{d+2}{p}},\end{aligned}\tag{A.9}$$

where $N_4 \coloneqq 4N_2|B_1|N_3$and $\alpha \coloneqq 1-\frac{d+2}{p}$ is the Hölder's coefficient.

Now we obtain the local α-Hölder continuity of the gradient with $\alpha = 1-\frac{d+2}{p}$. Consider a family of mollifiers $(\varphi_n)_{n\in\mathbb{N}}$ and the mollified sequence $(u_n)_{n\in\mathbb{N}} \subset C_c^\infty(\mathbb{R}^{d+1})$ with $u_n \coloneqq \varphi_n * u$. The family $(u_n)_{n\in\mathbb{N}}$ preserves the property (A.9). By properties of the family of mollifiers we have that $|u_n| \leq |u|$ and $D^m u_n = \varphi_n * D^m u$ for all multi-index m but only up to one time derivative and two spatial derivatives, because the function u is in $W^{1,2,p}_{\ell oc}(\mathbb{R}^{d+1})$. Moreover the mollified functions are smooth, this means that are α-Hölder continuous with some K_n depending on n. Consider two points (t,x) and (s,y) in $\mathbb{R}\times\mathbb{R}^d$, fix $\delta = \delta(\alpha) \in (0,\frac{1}{4}]$ such that $2\delta^\alpha \leq \frac{1}{2}$, define $\bar{d} \coloneqq |t-s|^{\frac{1}{2}} + |x-y|_d$ and $\rho > 0$ such that $\delta\bar{d}$ and $4\rho \leq \bar{d}$. For

$(r_1, p_1) \in Q_\rho(t, x)$ and $(r_2, p_2) \in Q_\rho(s, y)$ it holds

$$\begin{aligned}|u_n(t,x) - u_n(s,y)| \leq & |u_n(t,x) - u_n(r_1,p_1)| + |u_n(r_1,p_1) - u_n(r_2,p_2)| \\ & + |u_n(r_2,p_2) - u_n(s,y)| \\ \leq & K_n \left((|t - r_1|^{\frac{1}{2}} + |x - p_1|)^\alpha + (|s - r_2|^{\frac{1}{2}} + |y - p_2|)^\alpha \right) \\ & + |u(r_1,p_1) - u(r_2,p_2)| \\ \leq & 2K_n\rho^\alpha + |u(r_1,p_1) - u(r_2,p_2)|.\end{aligned}$$

By definition $\rho = \delta\bar{d}$ and we obtain

$$|u_n(t,x) - u_n(s,y)| \leq 2K_n(\delta\bar{d})^\alpha + |u(r_1,p_1) - u(r_2,p_2)|. \tag{A.10}$$

Integrating with respect to (r_1, p_1) and (r_2, p_2) in the two cylinders Q_ρ centred in (t, x) and (s, y), we get

$$\begin{aligned}(\delta\bar{d})^{2d+4}|B_1||u_n(t,x) - u_n(s,y)| \leq & 2|B_1|K_n(\delta\bar{d})^{\alpha+2d+4} \\ & + N_4\|u\|_{W^{1,2,p}(Q_{2\bar{d}}(t,x))}\bar{d}^{\alpha+2d+4},\end{aligned}$$

where the second term on the right-hand side of (A.10) has been estimated using (A.9). It means that

$$\begin{aligned}|u_n(t,x) - u_n(s,y)| \leq & 2K_n(\delta\bar{d})^\alpha + \tfrac{N_4}{|B_1|}\|u\|_{W^{1,2,p}(Q_{2\bar{d}}(t,x))}\bar{d}^\alpha\delta^{-2d-4} \\ = & (2K_n\delta^\alpha + \tfrac{N_4}{|B_1|}\|u\|_{W^{1,2,p}(Q_{2\bar{d}}(t,x))}\delta^{-2d-4})\bar{d}^\alpha.\end{aligned}$$

Recalling the definition of $\bar{d} = |t - s|^{\frac{1}{2}} + |x - y|_d$ and dividing by $\bar{d}^\alpha$ both sides of the inequality above we get

$$\frac{|u_n(t,x) - u_n(s,y)|}{(|t - s|^{\frac{1}{2}} + |x - y|_d)^\alpha} \leq 2K_n\delta^\alpha + \frac{N_4}{|B_1|}\|u\|_{W^{1,2,p}(Q_{2\bar{d}}(t,x))}\delta^{-2d-4}. \tag{A.11}$$

Using that $2\delta^\alpha \leq \frac{1}{2}$, we have from (A.11)

$$\frac{|u_n(t,x) - u_n(s,y)|}{(|t - s|^{\frac{1}{2}} + |x - y|_d)^\alpha} \leq \tfrac{1}{2}K_n + \frac{N_4}{|B_1|}\|u\|_{W^{1,2,p}(Q_{2\bar{d}}(t,x))}\delta^{-2d-4},$$

and taking the supremum on the right-hand side above

$$\sup_{\substack{(t,x),(s,y)\in Q_{2\bar{d}}(t,x)\\(t,x)\neq(s,y)}} \frac{|u_n(t,x)-u_n(s,y)|}{(|t-s|^{\frac{1}{2}}+|x-y|_d)^\alpha} \leq \frac{1}{2}K_n + \frac{N_4}{|B_1|}\|u\|_{W^{1,2,p}(Q_{2\bar{d}}(t,x))}\delta^{-2d-4}.$$

The left-hand side above is the α-Hölder constant of K_n, thus

$$K_n \leq \tfrac{1}{2}K_n + \tfrac{N_4}{|B_1|}\|u\|_{W^{1,2,p}(Q_{2\bar{d}}(t,x))}\delta^{-2d-4},$$

arranging the terms in the equation above, we have

$$K_n \leq 2\tfrac{N_4}{|B_1|}\|u\|_{W^{1,2,p}(Q_{2\bar{d}}(t,x))}\delta^{-2d-4},$$

which does not depend on n. We have obtain a bound on the α-Hölder constant uniformly in n and by uniform convergence on compacts we have that the limit function u is α-Hölder continuous with constant $N_5 := 2\frac{N_4}{|B_1|}\delta^{-2d-4}$.

It remains to prove that $|\nabla u|_d$ is locally bounded with a constant dependent on the $W^{1,2,p}$-norm of u in compact sets of $\mathbb{R}^{d+1}$. Fix $i \in (1,\dots d)$ and consider the partial derivative with respect to x_i of u computed at $(t,x)\in\mathbb{R}^{d+1}$, then

$$|u_{x_i}(t,x)| \leq |u_{x_i}(t,x)-u_{x_i}(s,y)| + |u_{x_i}(s,y)|.$$

We integrate with respect to (s,y) on $B_1(t,x)$ the unit ball centred in (t,x) and we obtain

$$|B_1^{d+1}||u_{x_i}(t,x)| \leq \int_{B_1(t,x)} |u_{x_i}(t,x)-u_{x_i}(s,y)| + |u_{x_i}(s,y)|\,\mathrm{d}s\mathrm{d}y,$$

where $|B_1^{d+1}|$ is the volume of a $d+1$-dimensional unit ball. Using the Hölder continuity property of u_{x_i} with constant $N_5\|u\|_{W^{1,2,p}(Q_{2\bar{d}}(t,x))}$ proved before on the first term on the right-hand side above and the Hölder inequality on the second term on the right-hand side above we get

$$
\begin{aligned}
|B_1^{d+1}||u_{x_i}(t,x)| \leq & \int_{B_1(t,x)} N_5 \|u\|_{W^{1,2,p}(Q_{2\bar{d}}(t,x))} (|t-s|^{\frac{1}{2}} + |x-y|_d)^\alpha \, \mathrm{d}s\mathrm{d}y \\
& + |B_1^{d+1}|^{\frac{p-1}{p}} \left(\int_{B_1(t,x)} |u_{x_i}(s,y)|^p \, \mathrm{d}s\mathrm{d}y \right)^{\frac{1}{p}} \\
\leq & N_5 C_1 \|u\|_{W^{1,2,p}(Q_{2\bar{d}}(t,x))} + |B_1^{d+1}|^{\frac{p-1}{p}} \|u_{x_i}\|_{L^p(B_1(t,x))}
\end{aligned}
$$

where C_1 is the integral in the first line above and it does not depend on u, and the norm in the last term above is finite because $u \in W^{1,2,p}_{\ell oc}(\mathbb{R}^{d+1})$. It follows that $|\nabla u|_d$ is locally bounded. □

Remark A.5: In order to be precise in Chapters 2 and 3, we should prove that there exists two constants $c_1 > 0$ and $c_2 > 0$ such that

$$
\begin{aligned}
c_1 \sup_{\substack{(t,x),(s,y)\in\mathcal{K} \\ (t,x)\neq(s,y)}} \frac{|v(t,x)-v(s,y)|}{(|t-s|^{\frac{1}{2}}+|x-y|)^\alpha} \leq & \sup_{\substack{(t,x),(s,y)\in\mathcal{K} \\ (t,x)\neq(s,y)}} \frac{|v(t,x)-v(s,y)|}{(|t-s|+|x-y|^2)^{\frac{\alpha}{2}}} \\
\leq & \, c_2 \sup_{\substack{(t,x),(s,y)\in\mathcal{K} \\ (t,x)\neq(s,y)}} \frac{|v(t,x)-v(s,y)|}{(|t-s|^{\frac{1}{2}}+|x-y|)^\alpha}.
\end{aligned}
\tag{A.12}
$$

Since we have that

$$
(|t-s|+|x-y|^2)^{\frac{1}{2}} \leq |t-s|^{\frac{1}{2}} + |x-y|,
$$

then the first inequality of (A.12) holds with $c_1 = 1$. The other inequality is straightforward using $c_2 = 2$. Indeed

$$
|t-s|^{\frac{1}{2}} + |x-y| \leq 2(|t-s|+|x-y|^2)^{\frac{1}{2}},
$$

using $x + y \leq 2\sqrt{x^2+y^2}$. ■

B Chapter 3

In this section, we give the auxiliary results for Chapter 3, i.e., the families of cut-off functions and penalty terms used in (3.18), the stability results for PDEs, the well-posedness of the penalised controlled process and the statement of the Shaefer's fixed point theorem.

B.1 Cut-off Functions

Here we give a construction of the function defined in (3.16) indexed by m. Let $\mu, \sigma : \mathbb{R} \to [0,1]$ be defined as

$$\mu(z) := \begin{cases} 0, & z \geq 1; \\ \exp\left(\frac{1}{z-1}\right), & z < 1; \end{cases} \quad \sigma(z) := \begin{cases} 0, & z \leq 0; \\ \exp\left(-\frac{1}{z}\right), & z > 0. \end{cases}$$

Notice that $\mu'(z) = -\mu(z)\frac{1}{(z-1)^2}$ and $\sigma'(z) = \sigma(z)\frac{1}{z^2}$. We can now use these two functions to define $\xi : \mathbb{R} \to [0,1]$ as

$$\xi(z) := \frac{\mu(z)}{\mu(z)+\sigma(z)} = \begin{cases} 1, & z \leq 0, \\ 0, & z \geq 1, \\ \exp\left(\frac{1}{z-1}\right) \Big/ \left[\exp\left(\frac{1}{z-1}\right) + \exp\left(-\frac{1}{z}\right)\right], & 0 < z < 1. \end{cases}$$

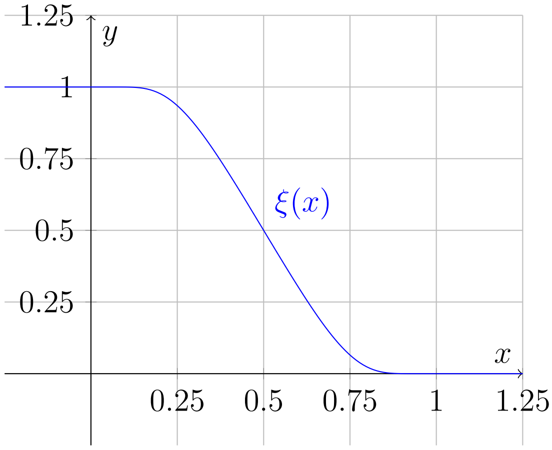

Figure 1: Graph of $\xi(x)$

We underline that $1-\xi(z)=\frac{\sigma(z)}{\mu(z)+\sigma(z)}$ for $z\in(0,1)$. Differentiate ξ with respect to z, we get

$$\begin{aligned}\xi'(z)=&\frac{\mu'(z)(\mu(z)+\sigma(z))-\mu(z)(\mu'(z)+\sigma'(z))}{(\mu(z)+\sigma(z))^2}\\=&\frac{\mu'(z)\sigma(z)-\mu(z)\sigma'(z)}{(\mu(z)+\sigma(z))^2}\\=&\frac{\mu(z)\sigma(z)}{(\mu(z)+\sigma(z))^2}\Big(\frac{1}{(z-1)^2}+\frac{1}{z^2}\Big)\\=&\,\xi(z)(1-\xi(z))\Big(\frac{1}{(z-1)^2}+\frac{1}{z^2}\Big),\end{aligned}$$

for $z\in(0,1)$, we have that $\xi'(z)=0$ for $z\in\mathbb{R}\setminus(0,1)$. We show the continuity of ξ' at $z=0$.

Notice that the function $\mu(z)+\sigma(z)=e^{-1}$ in $z=0$, thus

$$\begin{aligned}\lim_{z\downarrow 0}\xi(z)(1-\xi(z))\Big(\frac{1}{(z-1)^2}+\frac{1}{z^2}\Big)=&\lim_{z\downarrow 0}\frac{\sigma(z)}{\mu(z)+\sigma(z)}\Big(1+\frac{1}{z^2}\Big)\\=&\,e\lim_{z\downarrow 0}\sigma(z)\Big(\frac{1}{z^2}\Big)\\=&\,e\lim_{w\to\infty}e^{-w}w^2=0.\end{aligned}\tag{B.1}$$

The continuity at $z=1$ is the same because the function is symmetric respect to $\frac{1}{2}$.

We set $\xi_m(x)=\xi(|x|_d-m)$ for $x\in\mathbb{R}^d$. Then $\xi_m\in C_c^\infty(\mathbb{R}^d)$, $\xi_m:\mathbb{R}^d\to[0,1]$, $\xi_m=1$ on $\overline{B}_m$ and $\xi_m=0$ on $\mathbb{R}^d\setminus B_{m+1}$. It is clear that $\nabla\xi_m=\mathbf{0}$ on B_m and $\mathbb{R}^d\setminus B_{m+1}$. It can also be checked that for $x\in B_{m+1}\setminus B_m$

$$\partial_{x_k}\xi_m(x)=\frac{x_k}{|x|_d}\xi'(|x|_d-m),\quad\text{for } k=1,\dots d,$$

and therefore $|\nabla\xi_m(x)|_d^2=\big(\xi'(|x|_d-m)\big)^2$. Notice that $(\xi'(z))^2$ can be written as

$$\begin{aligned}\big(\xi'(z)\big)^2=&\,\xi^2(z)(1-\xi(z))^2\Big(\frac{1}{(z-1)^2}+\frac{1}{z^2}\Big)^2\\\le&\,\xi(z)\xi(z)(1-\xi(z))\Big(\frac{1}{(z-1)^2}+\frac{1}{z^2}\Big)^2.\end{aligned}$$

It remains to prove that $\zeta(z):=\xi(z)(1-\xi(z))(\frac{1}{(z-1)^2}+\frac{1}{z^2})^2$ has a maximum in

[0, 1].

Using the same computations as in (B.1), the function ζ is continuous at the boundary points 0 and 1. It is a product of continuous functions and it is continuous at the points where the two denominators vanishes. It is thus continuous on a compact and it admits a maximum. Then $|\nabla\xi_m(x)|_d^2 \leq C_0\xi_m(x)$ for all $x \in \mathbb{R}^d$, for a suitable $C_0 > 0$ independent of m.

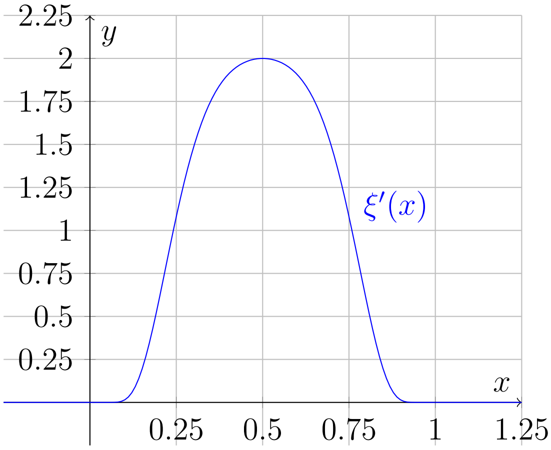

Figure 2: Graph of $\xi'(x)$

B.2 Penalty Functions

On the penalty functions ψ_ε introduced in Section 3.2.1 we have two possible candidates. The first one is a function defined by ξ from Section B.1:

$$\psi_1(x) := \begin{cases} 0 & x \leq 0, \\ 2(1 - \xi(\frac{x}{4})) & 0 \leq x \leq 2, \\ x - 1, & x \geq 2, \end{cases}$$

and we define $\psi_\varepsilon(x) := \psi_1(\frac{x}{\varepsilon})$.

The second candidate comes from [68, Page 363]

$$\tilde{\psi}_\varepsilon(x) := \begin{cases} 0, & x \le 0, \\ \frac{1}{6}\big(\frac{x}{\varepsilon}\big)^3, & 0 \le x \le \varepsilon, \\ -\frac{1}{6}\big(\frac{x}{\varepsilon} - 2\big)^3 + \big(\frac{x}{\varepsilon} - 2\big) + 1, & \varepsilon \le x \le 2\varepsilon, \\ \frac{x}{\varepsilon} - 1, & x \ge 2\varepsilon. \end{cases}$$

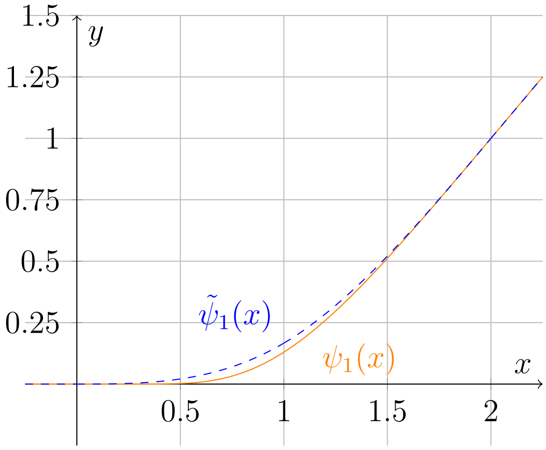

Figure 3: Graph of $\psi_1(x)$ and $\tilde{\psi}_1(x)$

B.3 Proof of Lemma 3.15

For existence and uniqueness of the solution to (3.34) we invoke [29, Thm. 3.3.7]. Indeed the smoothing of $u_m^{\varepsilon,\delta}$ and $(\,\cdot\,)^+$ guarantees that

$$h_m + \tfrac{1}{\delta}\chi_n(g_m - u_m^{\varepsilon,\delta}) - \psi_\varepsilon\big(|\nabla u^n|_d^2 - f_m^2 - \tfrac{1}{n}\big) \in C^{0,1,\alpha}(\overline{\mathcal{O}}_m). \tag{B.2}$$

Moreover, the compatibility condition

$$\lim_{s\uparrow T}(\partial_t g_m + \mathcal{L}g_m - rg_m)(s,x) \\ = \Big[-h_m - \tfrac{1}{\delta}\chi_n(g_m - u_m^{\varepsilon,\delta}) + \psi_\varepsilon\big(|\nabla u^n|_d^2 - f_m^2 - \tfrac{1}{n}\big)\Big](T,x),$$

for $x \in \partial B_m$, holds with both sides of the equation equal to zero. Indeed, given that $\xi_{m-1} \in C_c^\infty(B_m)$ we have $g_m = \partial_{x_i} g_m = \partial_{x_i x_j} g_m = 0$ on $[0,T]\times\partial B_m$. Moreover,

$g_m = 0$ on $[0,T] \times \partial B_m$ also implies $\partial_t g_m = 0$ on $[0,T] \times \partial B_m$. So the left-hand side of the equation is equal to zero. On the right-hand side, for $x \in \partial B_m$ we have $h_m(T,x) = g_m(T,x) = u_m^{\varepsilon,\delta}(T,x) = 0$ and

$$|\nabla u^n(T,x)|_d^2 \leq |\nabla u_m^{\varepsilon,\delta}(T,x)|_d^2 + \tfrac{1}{n} = |\nabla g_m(T,x)|_d^2 + \tfrac{1}{n} \leq f_m^2(T,x) + \tfrac{1}{n},$$

by uniform convergence of ∇u^n to $\nabla u_m^{\varepsilon,\delta}$ and (3.17). The compatibility condition follows upon recalling $\chi_n(0) = 0$ and $\psi_\varepsilon(z) = 0$ for $z \leq 0$.

The fact that $w^n \in C_{Loc}^{1,3,\alpha}(\mathcal{O}_m)$ is also consequence of (B.2) and standard interior estimates for PDEs [29, Thm. 3.5.11 and Cor. 3.5.1]. Instead, the convergence result $w^n \to u_m^{\varepsilon,\delta}$ in $C^{1,2,\beta}(\overline{\mathcal{O}}_m)$, as $n \to \infty$, for $\beta \in (0,\alpha)$, is a special case of Lemma B.1. □

Lemma B.1 *Let $F : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ be a Lipschitz continuous function. Fix $\phi, \varphi \in C^{0,1,\alpha}(\overline{\mathcal{O}}_m)$ and let u be a solution in $C^{1,2,\alpha}(\overline{\mathcal{O}}_m)$ of*

$$\begin{cases} \partial_t u + \mathcal{L}u - ru = -h_m + F(\phi, \nabla\varphi), & \text{on } \mathcal{O}_m, \\ u(t,x) = g_m(t,x), & (t,x) \in \partial_P \mathcal{O}_m. \end{cases} \tag{B.3}$$

Let $(\phi_n)_{n\in\mathbb{N}}, (\varphi_n)_{n\in\mathbb{N}} \subseteq C^{0,1,\alpha}(\overline{\mathcal{O}}_m)$ be such that $\phi_n \to \phi$ and $\varphi_n \to \varphi$ in $C^{0,1,\gamma}(\overline{\mathcal{O}}_m)$ as $n \to \infty$ for all $\gamma \in (0,\alpha)$. Let $(F_n)_{n\in\mathbb{N}}$ be equi-Lipschitz continuous functions $F_n : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ such that $F_n \to F$ in $C^0_{\ell oc}(\mathbb{R}^{1+d})$. Finally, denote by u_n a solution to (B.3) in $C^{1,2,\alpha}(\overline{\mathcal{O}}_m)$ with $F_n(\phi_n, \nabla\varphi_n)$ instead of $F(\phi, \nabla\varphi)$.

Then, up to possibly selecting a subsequence,

$$\lim_{n\to\infty} \|u_n - u\|_{C^{1,2,\gamma}(\overline{\mathcal{O}}_m)} = 0, \quad \text{for all } \gamma \in (0,\alpha). \tag{B.4}$$

If $(\varphi_n)_{n\in\mathbb{N}} \subseteq C^{0,2,\alpha}(\overline{\mathcal{O}}_m)$ and $(F_n)_{n\in\mathbb{N}} \subseteq C^{1,\alpha}(\mathbb{R}^{1+d})$, then $(u_n)_{n\in\mathbb{N}} \subseteq C_{Loc}^{1,3,\alpha}(\mathcal{O}_m) \cap C^{1,2,\alpha}(\overline{\mathcal{O}}_m)$.

Proof. Define $\hat{u}_n := u - u_n$. Then $\hat{u}_n$ solves

$$\begin{cases} \partial_t \hat{u}_n + \mathcal{L}\hat{u}_n - r\hat{u}_n = F(\phi, \nabla\varphi) - F_n(\phi_n, \nabla\varphi_n), & \text{on } \mathcal{O}_m, \\ \hat{u}_n(t,x) = 0, & (t,x) \in \partial_P \mathcal{O}_m. \end{cases}$$

By [29, Thm. 3.2.6] we have the estimate

$$\|\hat{u}_n\|_{C^{1,2,\gamma}(\overline{\mathcal{O}}_m)} \leq K\|F(\phi,\nabla\varphi) - F_n(\phi_n,\nabla\varphi_n)\|_{C^{0,0,\gamma}(\overline{\mathcal{O}}_m)}, \tag{B.5}$$

for a constant $K > 0$ independent of n. Notice that by equi-Lipschitz continuity, the sequence $(F_n)_{n\in\mathbb{N}}$ is compact in any $C^\beta(\overline{U})$ for $\beta \in (0,1)$ and bounded set $U \subset \mathbb{R}^{d+1}$. Thanks to the convergence of the functions ϕ_n, φ_n and F_n we have that, up to possibly selecting a subsequence, $F_n(\phi_n,\nabla\varphi_n) \to F(\phi,\nabla\varphi)$ in $C^\gamma(\overline{\mathcal{O}}_m)$ as $n \to \infty$ for all $\gamma \in (0,\alpha)$. Thus, (B.4) holds.

If we also assume that $(\varphi_n)_{n\in\mathbb{N}} \subseteq C^{0,2,\alpha}(\overline{\mathcal{O}}_m)$ and $(F_n)_{n\in\mathbb{N}} \subseteq C^{1,\alpha}(\mathbb{R}^{1+d})$, it turns out that $F_n(\phi_n,\varphi_n) \in C^{0,1,\alpha}(\overline{\mathcal{O}}_m)$ and since the coefficients of $\mathcal{L}$ are continuously differentiable then $u_n \in C^{1,3,\alpha}_{Loc}(\mathcal{O}_m)$ for all n by [29, Thm. 3.5.11 and Cor. 3.5.1]. □

Remark B.2: Thanks to [8, Thm. 2.6.5 and Rem. 2.6.4] the bound (B.5) can be replaced by

$$\|\hat{u}_n\|_{W^{1,2,p}(\mathcal{O}_m)} \leq K\|F(\phi,\nabla\varphi) - F_n(\phi_n,\nabla\varphi_n)\|_{L^p(\mathcal{O}_m)}, \quad p \in (1,\infty).$$

Hence, stability of solutions of (B.3) also holds in the space $W^{1,2,p}(\mathcal{O}_m)$, i.e., $\lim_n u_n = u$ in $W^{1,2,p}(\mathcal{O}_m)$. ■

B.4 Maximum Principle

The maximum principle is a well-known result in the theory of PDE. We give an alternative proof of the maximum principle by a probabilistic argument. Consider the second-order differential operator $\mathcal{L}$ defined as

$$(\mathcal{L}\varphi)(x) = \frac{1}{2}\mathrm{tr}(a(x)D^2\varphi(x)) + \langle b(x), \nabla\varphi(x)\rangle$$

for any $\varphi \in C^2(\mathbb{R}^d)$.

Lemma B.3 (Maximum Principle) *Let $\varphi \in C^{1,2,\alpha}(\mathcal{O})$ for some $\mathcal{O} \subseteq \mathbb{R}^{d+1}_{0,T}$. Let*

$(t,x) \in \mathcal{O}$ be an internal point of (local) maximum for φ. Then

$$\partial_t \varphi(t,x) + \mathcal{L}\varphi(t,x) \leq 0$$

where $\mathcal{L}$ is a second order differential operator locally elliptic (see (3.10)) defined as above.

Proof. Let (t,x) be an internal of (local) maximum. There exists a $\delta > 0$ with $t + \delta \leq T$, and a $B_\delta(t,x) := \{(s,y) \in [t, t+\delta) \times \mathbb{R}^d : |x-y|_d < \delta\}$ such that $\varphi(t,x) \geq \varphi(s,y)$ for all $(s,y) \in B_\delta(t,x)$. Since the matrix $a(x)$ is locally elliptic, it is uniformly elliptic in a neighbourhood of (t,x) and thus it is positive-definite. It means that we can decompose $a(x) = \sigma(x)\sigma^\top(x)$ for some $\sigma \in \mathbb{R}^{d\times d}$ (see [55, page 558]); this σ is not necessarily unique, but it would be if we required symmetry of it. We can now consider the stochastic process

$$\begin{aligned} \mathrm{d}X_s &= b(X_s)\mathrm{d}s + \sigma(X_s)\mathrm{d}W_s, \quad \text{for } s \in [0,\infty), \\ X_0 &= x, \end{aligned}$$

where $(W_s)_{s\in[0,\infty)}$ is a d-dimensional Brownian motion. Let τ_δ the first exit time of the process X_s from the ball $B_\delta(x) := \{y \in \mathbb{R}^d : |x-y|_d < \delta\}$. By an application of Dynkin's formula, justified by the boundedness of the spatial derivatives of φ, we have

$$\mathsf{E}_x\big[\varphi(t+\tau_\delta, X_{\tau_\delta})\big] = \varphi(t,x) + \mathsf{E}_x\bigg[\int_0^{\tau_\delta} (\partial_t + \mathcal{L})\varphi(t+s, X_s)\,\mathrm{d}s\bigg].$$

The functions attains a maximum in (t,x), thus the left-hand side above is less or equal than the first term on the right-hand side above and we get

$$\mathsf{E}_x\bigg[\int_0^{\tau_\delta} (\partial_t + \mathcal{L})\varphi(t+s, X_s)\,\mathrm{d}s\big] \leq 0.$$

Dividing by δ and passing to the limit as $\delta \downarrow 0$, we are allow to pass the limit under expectation by dominated convergence theorem. Thus, we have

$$(\partial_t + \mathcal{L})\varphi(t,x) \leq 0,$$

and the lemma is proved. □

B.5 Convergence of the Sequence $(t^\lambda_{n_k}, x^\lambda_{n_k})_{k\in\mathbb{N}}$ in Proposition 3.17

Here we prove that $(\tilde t, \tilde x) \in \arg\max_{\overline{\mathcal{O}}_m} v^\lambda$. Arguing by contradiction let us assume $(\tilde t, \tilde x) \notin \arg\max_{\overline{\mathcal{O}}_m} v^\lambda$. Then there exists a $\epsilon > 0$ such that $v^\lambda(\tilde t, \tilde x) \le \max_{\overline{\mathcal{O}}_m} v^\lambda - \epsilon$ and so there exists a neighbourhood U_ϵ of $(\tilde t, \tilde x)$ such that $v^\lambda(t, x) \le \max_{\overline{\mathcal{O}}_m} v^\lambda - \frac{\epsilon}{2}$ for all $(t, x) \in \overline{U}_\epsilon$. For all sufficiently large k's we also have $(t^\lambda_{n_k}, x^\lambda_{n_k}) \in U_\epsilon$ and by uniform convergence

$$|v^{\lambda,n_k} - v^\lambda|(t, x) \le \frac{\epsilon}{4}, \quad \text{for } (t, x) \in \overline{\mathcal{O}}_m. \tag{B.6}$$

Hence

$$\begin{aligned}\max_{(t,x)\in\overline{\mathcal{O}}_m} v^{\lambda,n_k}(t, x) &= v^{\lambda,n_k}(t^\lambda_{n_k}, x^\lambda_{n_k})\\ &\le v^\lambda(t^\lambda_{n_k}, x^\lambda_{n_k}) + \frac{\epsilon}{4}\\ &\le \max_{(t,x)\in\overline{\mathcal{O}}_m} v^\lambda(t, x) - \frac{\epsilon}{4},\end{aligned} \tag{B.7}$$

where the first equality is by definition of $(t^\lambda_{n_k}, x^\lambda_{n_k})$, the first inequality by (B.6) and the final inequality follows by $(t^\lambda_{n_k}, x^\lambda_{n_k}) \in \overline{U}_\epsilon$.

With no loss of generality we can assume v^λ and $v^{\lambda,n}$ be positive. Otherwise we apply our argument to $\tilde v^\lambda = v^\lambda - \min_{\overline{\mathcal{O}}_m} v^\lambda + 1$ and the associated sequence $\tilde v^{\lambda,n} = v^{\lambda,n} - \min_{\overline{\mathcal{O}}_m} v^\lambda + 1$. By triangular inequality and positivity of v^λ and $v^{\lambda,n}$ we have

$$\max_{\overline{\mathcal{O}}_m} |v^\lambda - v^{\lambda,n_k}| \ge \max_{\overline{\mathcal{O}}_m} |v^\lambda| - \max_{\overline{\mathcal{O}}_m} |v^{\lambda,n_k}| = \max_{\overline{\mathcal{O}}_m} v^\lambda - \max_{\overline{\mathcal{O}}_m} v^{\lambda,n_k} \ge \frac{\epsilon}{4},$$

for all k's sufficiently large, where the inequality is due to (B.7). This contradicts uniform convergence and therefore $(\tilde t, \tilde x) \in \arg\max_{\overline{\mathcal{O}}_m} v^\lambda$ as claimed.

B.6 Existence of X^ε in the proof of Theorem 3.33

For $m \in \mathbb{N}$ let (b^m, σ^m) be functions $\mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^{d\times d'}$ that are equal to (b, σ) on B_m and extend continuously to be constant on $\mathbb{R}^d \setminus B_m$. Similarly, $\alpha^m : \mathbb{R}^{d+1}_{0,T} \to \mathbb{R}^d$

is defined as

$$\alpha^m(t,x) = -2\psi'_\varepsilon\left(|\nabla u^\varepsilon(t,x)|_d^2 - f^2(t,x)\right)\nabla u^\varepsilon(t,x), \quad \text{on } B_m,$$

and extended continuously to be constant on $\mathbb{R}^d \setminus B_m$. Since $u^\varepsilon \in C^{0,1,\alpha}_{\ell oc}(\mathbb{R}^{d+1}_{0,T})$, then thanks to [65, Thm. 1] there exists a unique strong solution of

$$X^m_s = x + \int_0^s \left(b^m(X^m_u) + \alpha^m(t+u, X^m_u)\right)\mathrm{d}u + \int_0^s \sigma^m(X^m_u)\mathrm{d}W_u, \quad s \in [0, T-t].$$

Notice that $X^m = X^{m;t}$ depends on t via the time-inhomogeneous drift α^m but we omit it for simplicity. Notice also that here we should understand $n^m_s \dot{\nu}^m_s = \alpha^m(t+s, X^m_s)$ and $X^m = X^{m;[n^m,\nu^m]}$. Letting $\zeta_{m,k} = \inf\{s \geq 0 : |X^m_s|_d \geq k\}$, for any $m \geq k$ we have $X^m_{s\wedge\zeta_{m,k}} = X^k_{s\wedge\zeta_{k,k}}$ for all $s \in [0, T-t]$, P_x-a.s. (i.e., the two processes are indistinguishable). Thus, setting $X^\varepsilon_s(\omega) := X^k_s(\omega)$ for $s < \zeta_{k,k}(\omega)$ and denoting $\tau_k = \inf\{s \geq 0 : |X^\varepsilon_s|_d \geq k\}$ it is clear that, by uniqueness of strong solutions and the definition of the pair $(n^\varepsilon, \nu^\varepsilon)$, the process X^ε satisfies

$$X^\varepsilon_{s\wedge\tau_k} = x + \int_0^{s\wedge\tau_k} \left(b(X^\varepsilon_u) + n^\varepsilon_u \dot{\nu}^\varepsilon_u\right)\mathrm{d}u + \int_0^{s\wedge\tau_k} \sigma(X^\varepsilon_u)\mathrm{d}W_u, \quad s \in [0, T-t].$$

By continuity of paths $\tau_k \leq \tau_{k+1}$ and $\tau_\infty := \lim_{k\to\infty} \tau_k$ is well-defined. Moreover, ν^ε satisfies the same bound as in (3.102) thanks to (3.126). Therefore, linear growth of the coefficients of the SDE and the same arguments as those at the end of the proof of Proposition 3.25 imply that $\mathsf{P}_x(\tau_\infty \leq T-t) = 0$. Then X^ε is well-defined on $[0, T-t]$.

B.7 Shaefer's fixed point theorem

Theorem B.4 (Thm. 9.2.4 in [25]) *Suppose $T : \mathcal{D} \to \mathcal{D}$ is a continuous and compact mapping on a Banach space $\mathcal{D}$. Assume further that the set*

$$\{f \in \mathcal{D} \,|\, f = \rho T[f] \text{ for some } \rho \in [0,1]\}$$

is bounded. Then T has a fixed point.

C Chapter 4

In this section, we give the auxiliary results for Chapter 4, i.e., an estimate on the fist moment of the local time process.

C.1 Upper bound for the Local Time

We give below an extension of the result in [20, Lemma 5.1]

Lemma C.1 *Let X be a real valued càdl'ag semimartingale with jumps of bounded variation and starting from x, let $L_t^0(X)$ be its local time at 0 in the time-interval $[0,t]$. Then, for any $\epsilon \in (0,1)$ we have*

$$\begin{aligned} \mathsf{E}_x[L_t^0(X)] \leq 4\epsilon - 2\mathsf{E}_x\Big[\int_0^t \big(\mathbb{1}_{\{X_s\in[0,\epsilon)\}} + \mathbb{1}_{\{X_s\geq\epsilon\}} e^{1-\frac{X_s}{\epsilon}}\big)\,\mathrm{d}X_s^c\Big] \\ + \frac{1}{\epsilon}\mathsf{E}_x\Big[\int_0^t \mathbb{1}_{\{X_s>\epsilon\}} e^{1-\frac{X_s}{\epsilon}}\,\mathrm{d}\langle X\rangle_s^c\Big] \\ + \mathsf{E}_x\Big[\sum_{0<s\leq t} |\Delta X_s|\Big] \end{aligned}$$

where X_s^c and $\langle X\rangle_s^c$ are the continuous parts of X and of the quadratic variation of X, respectively, and $\Delta X_s := X_s - X_{s-}$.

Proof. For $\epsilon \in (0,1)$ and $y \in \mathbb{R}$ we define

$$g_\epsilon(y) = 0 \cdot \mathbb{1}_{\{y<0\}} + y\mathbb{1}_{\{0\leq y<\epsilon\}} + \epsilon\big(2 - e^{1-\frac{y}{\epsilon}}\big)\mathbb{1}_{\{y\geq\epsilon\}}.$$

Following the idea from [20, Lemma 5.1] we have that $g_\epsilon \in C^1(\mathbb{R}\setminus\{0\})$, it is semi-concave, i.e., $y \mapsto g_\epsilon(y) - y^2$ is concave. Moreover, g_ϵ is such that

$$\begin{aligned} & 0 \leq g_\epsilon(y) \leq 2\epsilon, \quad \text{for } y \in \mathbb{R}; \\ & g_\epsilon'(y) = \mathbb{1}_{\{0\leq y\leq\epsilon\}} + e^{1-\frac{y}{\epsilon}}\mathbb{1}_{\{y\geq\epsilon\}}, \quad \text{for } y \in \mathbb{R}; \\ & g_\epsilon''(y) = 0, \quad \text{for } y \in (-\infty,0)\cup(0,\epsilon); \\ & g_\epsilon''(y) = -\epsilon^{-1}e^{1-\frac{y}{\epsilon}}, \quad \text{for } y > \epsilon. \end{aligned}$$

Applying [57, Thm. IV.70 and Cor. IV.70.1] to $g_\epsilon(X_t)$ we get

$$\begin{aligned} g_\epsilon(X_t) - g_\epsilon(X_0) = &\int_0^t g_\epsilon'(X_{s-})\,\mathrm{d}X_s^c + \frac{1}{2}\int_0^t g_\epsilon''(X_s)\mathbb{1}_{\{X_s\neq 0\}\cap\{X_s\neq\epsilon\}}\,\mathrm{d}\langle X\rangle_s^c \\ &+ L_t^0(X) + \sum_{0<s\leq t} g_\epsilon(X_s) - g_\epsilon(X_{s-}). \end{aligned}$$

Rearranging terms and multiplying by 2, using $|g_\epsilon(X_s) - g_\epsilon(X_{s-})| \leq |X_s - X_{s-}| = |\Delta X_s|$ by Lipschitz property of g_ϵ and that X has jumps of finite variation, we get

$$L_t^0(X) \leq 4\epsilon - 2\int_0^t g_\epsilon'(X_{s-})\,\mathrm{d}X_s^c - \int_0^t g_\epsilon''(X_s)\mathbb{1}_{\{X_s\neq 0\}\cap\{X_s\neq\epsilon\}}\,\mathrm{d}\langle X\rangle_s^c + 2\sum_{0<s\leq t}|\Delta X_s|$$

Using the properties of g_ϵ listed above (see also [20, Lemma 5.1]) and applying expectation we get the result. □

Bibliography

[1] L.H.R. Alvarez. Singular stochastic control in the presence of a state-dependent yield structure. *Stochastic Process. Appl.*, 86(2):323–343, 2000.

[2] L.H.R. Alvarez and L.A. Shepp. Optimal harvesting of stochastically fluctuating populations. *Journal of Mathematical Biology*, 37(2):155–177, 1998.

[3] J. Bather and H. Chernoff. Sequential decisions in the control of a space-ship. In *Fifth Berkeley Symposium on Mathematical Statistics and Probability*, volume 3, page 181. University of California Press, 1967.

[4] J. Bather and H. Chernoff. Sequential decisions in the control of a space-ship (finite fuel). *J. Appl. Probab.*, 4(3):584–604, 1967.

[5] E. Bayraktar and Y.-J. Huang. On the multidimensional controller-and-stopper games. *SIAM J. Control Optim.*, 51(2):1263–1297, 2013.

[6] E. Bayraktar and V.R. Young. Proving regularity of the minimal probability of ruin via a game of stopping and control. *Finance Stoch.*, 15(4):785–818, 2011.

[7] V.E. Beneš, L.A. Shepp, and H.S. Witsenhausen. Some solvable stochastic control problemst. *Stochastics*, 4(1):39–83, 1980.

[8] A. Bensoussan and J.-L. Lions. *Applications of variational inequalities in stochastic control*, volume 12 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam-New York, 1982.

[9] F.E. Benth, K.H. Karlsen, and K. Reikvam. On the existence of optimal controls for a singular stochastic control problem in finance. In *Mathematical*

finance (Konstanz, 2000), Trends Math., pages 79–88. Birkhäuser, Basel, 2001.

[10] F. Boetius and M. Kohlmann. Connections between optimal stopping and singular stochastic control. *Stochastic Processes and their Applications*, 77(2):253–281, 1998.

[11] A. Bovo, T. De Angelis, and E. Issoglio. Variational inequalities on unbounded domains for zero-sum singular-controller vs. stopper games, 2022.

[12] A. Budhiraja and K. Ross. Existence of optimal controls for singular control problems with state constraints. *Ann. Appl. Probab.*, 16(4):2235–2255, 2006.

[13] M.B. Chiarolla and T. De Angelis. Optimal stopping of a hilbert space valued diffusion: An infinite dimensional variational inequality. *Appl. Math. Optim.*, 73(2):271–312, 2016.

[14] M.B. Chiarolla and U.G. Haussmann. Optimal control of inflation: A central bank problem. *SIAM J. Control Optim.*, 36(3):1099–1132, 1998.

[15] M.B. Chiarolla and U.G. Haussmann. On a stochastic, irreversible investment problem. *SIAM J. Control Optim.*, 48(2):438–462, 2009.

[16] S. Choukroun, A. Cosso, and H. Pham. Reflected BSDEs with nonpositive jumps, and controller-and-stopper games. *Stochastic Process. Appl.*, 125(2):597–633, 2015.

[17] P.-L. Chow, J.-L. Menaldi, and M. Robin. Additive control of stochastic linear systems with finite horizon. *SIAM J. Control Optim.*, 23(6):858–899, 1985.

[18] M.H.A. Davis and A.R. Norman. Portfolio selection with transaction costs. *Math. Oper. Res.*, 15(4):676–713, 1990.

[19] T. De Angelis and G. Ferrari. A stochastic partially reversible investment problem on a finite time-horizon: Free-boundary analysis. *Stochastic Process. Appl.*, 124(12):4080–4119, 2014.

[20] T. De Angelis, M. Germain, and E. Issoglio. A numerical scheme for stochastic differential equations with distributional drift, 2022.

[21] A.K. Dixit and R.S. Pindyck. *Investment under Uncertainty.* Number 5474 in Economics Books. Princeton University Press, 1994.

[22] E.B. Dynkin. Optimal choice of the stopping moment of a markov process. *Dokl. Akad. Nauk SSSR*, 150(2):238–240, 1963.

[23] N. El Karoui. Les aspects probabilistes du contrôle stochastique. In P.L. Hennequin, editor, *Ninth Saint Flour Probability Summer School—1979 (Saint Flour, 1979)*, volume 876, pages 73–238. Springer Berlin Heidelberg, 1981.

[24] L.C. Evans. A second order elliptic equation with gradient constraint. *Comm. Partial Differential Equations*, 4(5):555–572, 1979.

[25] L.C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, Providence, R.I., second edition, 2010.

[26] W.H. Fleming and R.W. Rishel. *Deterministic and stochastic optimal control*, volume 1 of *Applications of Mathematics.* Springer-Verlag, Berlin-New York, 1975.

[27] W.H. Fleming and H.M. Soner. *Controlled Markov processes and viscosity solutions*, volume 25 of *Stochastic Modelling and Applied Probability.* Springer, New York, second edition, 2006.

[28] A. Friedman. Stochastic games and variational inequalities. *Arch. Ration. Mech. Anal.*, 51(5):321–346, 1973.

[29] A. Friedman. *Partial differential equations of parabolic type.* Courier Dover Publications, 2008.

[30] D. Gilbarg and N.S. Trudinger. *Elliptic partial differential equations of second order.* Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[31] X. Guo and P. Tomecek. A class of singular control problems and the smooth fit principle. *SIAM J. Control Optim.*, 47(6):3076–3099, 2009.

[32] S. Hamadène and J.-P. Lepeltier. Reflected bsdes and mixed game problem. *Stochastic Process. Appl.*, 85(2):177–188, 2000.

[33] S. Hamadène. Mixed zero-sum stochastic differential game and American game options. *SIAM J. Control Optim.*, 45(2):496–518, 2006.

[34] U.G. Haussmann and W. Suo. Singular optimal stochastic controls I: Existence. *SIAM J. Control Optim.*, 33(3):916–936, 1995.

[35] D. Hernandez-Hernandez, R.S. Simon, and M. Zervos. A zero-sum game between a singular stochastic controller and a discretionary stopper. *Ann. Appl. Probab.*, 25(1):46–80, 2015.

[36] D. Hernandez-Hernandez and K. Yamazaki. Games of singular control and stopping driven by spectrally one-sided lévy processes. *Stochastic Process. Appl.*, 125(1):1–38, 2015.

[37] G. A. Hunt. Markov processes and potentials I. *Illinois J. Math.*, 1(1):44 – 93, 1957.

[38] R. Hynd. Analysis of Hamilton-Jacobi-Bellman equations arising in stochastic singular control. *ESAIM Control Optim. Calc. Var.*, 19(1):112–128, 2013.

[39] H. Ishii and S. Koike. Boundary regulatity and uniqueness for an elliptic equations with gradient constraint. *Comm. Partial Differential Equations*, 8(4):317–346, 1983.

[40] I. Karatzas and S.E. Shreve. Connections between optimal stopping and singular stochastic control i. monotone follower problems. *SIAM J. Control Optim.*, 22(6):856–877, 1984.

[41] I. Karatzas and S.E. Shreve. Connections between optimal stopping and singular stochastic control ii. reflected follower problems. *SIAM J. Control Optim.*, 23(3):433–451, 1985.

[42] I. Karatzas and W.D. Sudderth. The controller-and-stopper game for a linear diffusion. *Ann. Probab.*, 29(3):1111–1127, 2001.

[43] I. Karatzas and H. Wang. A barrier option of american type. *Appl. Math. Optim.*, 42(3):259–279, 2000.

[44] I. Karatzas and I.-M. Zamfirescu. Martingale approach to stochastic differential games of control and stopping. *Ann. Probab.*, 36(4):1495–1527, 2008.

[45] M. Kelbert and H.A. Moreno-Franco. HJB equations with gradient constraint associated with controlled jump-diffusion processes. *SIAM J. Control Optim.*, 57(3):2185–2213, 2019.

[46] L. Kruk. Optimal policies for n-dimensional singular stochastic control problems part I: The Skorohod problem. *SIAM J. Control Optim.*, 38(5):1603–1622, 2000.

[47] N.V. Krylov. *Lectures on elliptic and parabolic equations in Sobolev spaces*, volume 96 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2008.

[48] N.V. Krylov. *Controlled diffusion processes*, volume 14 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2009. Translated from the 1977 Russian original by A.B. Aries, Reprint of the 1980 edition.

[49] J. Li and G. Zitkovic. Existence, characterization, and approximation in the generalized monotone-follower problem. *SIAM J. Control Optim.*, 55(1):94–118, 2017.

[50] P.C. Lon and M. Zervos. A model for optimally advertising and launching a product. *Math. Oper. Res.*, 36(2):363–376, 2011.

[51] J. Ma. On the principle of smooth fit for a class of singular stochastic control problems for diffusions. *SIAM J. Control Optim.*, 30(4):975–999, 1992.

[52] A.P. Maitra and W.D. Sudderth. The gambler and the stopper. In *Statistics, probability and game theory*, volume 30 of *IMS Lecture Notes Monogr. Ser.*, pages 191–208. Inst. Math. Statist., Hayward, CA, 1996.

[53] H.P. McKean. A free boundary problem for the heat equation arising from a problem in mathematical economics. *Industrial Management Review*, 6(2):32–39, 1965.

[54] J.L. Menaldi and M.I. Taksar. Optimal correction problem of a multidimensional stochastic system. *Automatica J. IFAC*, 25(2):223–232, 1989.

[55] C.D. Meyer. *Matrix analysis and applied linear algebra*, volume 71. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000.

[56] G. Peskir and A. Shiryaev. *Optimal stopping and free-boundary problems*. Lectures in Mathematics. ETH Zürich. Birkhäuser Verlag, Basel, 2006.

[57] P.E. Protter. *Stochastic integration and differential equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, second edition, 2005.

[58] A.N. Shiryaev. On optimum methods in quickest detection problems. *Theory Probab. Appl.*, 8(1):22–46, 1963.

[59] A.N. Shiryaev. *Optimal stopping rules*, volume 8 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2008. Translated from the 1976 Russian second edition by A.B. Aries, Reprint of the 1978 translation.

[60] J.L. Snell. Applications of martingale system theorems. *Trans. Amer. Math. Soc.*, 73(2):293–312, 1952.

[61] H.M. Soner and S.E. Shreve. Regularity of the value function for a two-dimensional singular stochastic control problem. *SIAM J. Control Optim.*, 27(4):876–907, 1989.

[62] H.M. Soner and S.E. Shreve. A free boundary problem related to singular stochastic control: the parabolic case. *Comm. Partial Differential Equations*, 16(2-3):373–424, 1991.

[63] L. Stettner. Penalty method for finite horizon stopping problems. *SIAM J. Control Optim.*, 49(3):1078–1099, 2011.

[64] M.I. Taksar. Skorohod problems with nonsmooth boundary conditions. *J. Comput. Appl. Math.*, 40(2):233–251, 1992.

[65] A.Y. Veretennikov. On strong solutions and explicit formulas for solutions of stochastic integral equations. *Mat. Sb.*, 39(3):387–403, 1981.

[66] A. Wald. Sequential tests of statistical hypotheses. *Ann. Math. Statist.*, 16(2):117–186, 1945.

[67] A. Weerasinghe. A controller and a stopper game with degenerate variance control. *Electron. Commun. Probab.*, 11:89–99, 2006.

[68] M. Wiegner. The C1,1–character of solutions of second order elliptic equations with gradient constraint. *Comm. Partial Differential Equations*, 6(3):361–371, 1981.

[69] H. Zhu. Generalized solution in singular stochastic control: the nondegenerate problem. *Appl. Math. Optim.*, 25(3):225–245, 1992.