IRIS AperTO

UNIVERSITÀ DEGLI STUDI DI TORINO

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

On Lipschitz continuous optimal stopping boundaries

This is the author's manuscript

Original Citation:

Availability:

This version is available http://hdl.handle.net/2318/1761919 since 2020-11-09T09:48:13Z

Published version:

DOI:10.1137/17M1113709

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

12 October 2024

On Lipschitz continuous optimal stopping boundaries

Tiziano De Angelis* & Gabriele Stabile†

November 9, 2020

Abstract. We obtain a probabilistic proof of the local Lipschitz continuity for the optimal stopping boundary of a class of problems with state space $[0,T]\times\mathbb{R}^d$, $d\geq 1$. To the best of our knowledge this is the only existing proof that relies exclusively upon stochastic calculus, all the other proofs making use of PDE techniques and integral equations. Thanks to our approach we obtain our result for a class of diffusions whose associated second order differential operator is not necessarily uniformly elliptic. The latter condition is normally assumed in the related PDE literature.

Keywords: optimal stopping, free boundary problems, Lipschitz free boundaries.

MSC2010 subject classification: 60G40, 35R35.

1 Introduction

In this work we deal with optimal stopping problems of the form

$$v(t,x)=\sup_{0\leq\tau\leq T-t}\mathsf{E}\left[\int_0^\tau h(t+s,X_s^x)ds+\mathbb{1}_{\{\tau<T-t\}}f(t+\tau,X_\tau^x)+\mathbb{1}_{\{\tau=T-t\}}g(X_\tau^x)\right] \tag{1.1}$$

where E denotes the expectation operator. For $d\geq 1$ and $d'\geq 1$, given a suitable $\mathbb{R}^d$-valued function μ and a $d\times d'$ matrix σ, the process $X\in\mathbb{R}^d$ follows the dynamic

$$X_t^x=x+\int_0^t\mu(X_s^x)ds+\sigma B_t,\quad t\geq 0,$$

with B a $\mathbb{R}^{d'}$-valued Brownian motion. The main focus of our study is the analysis of the regularity of the optimal stopping boundary, i.e. the boundary of the set in $[0,T)\times\mathbb{R}^d$ where $v=f$.

Under mild assumptions on μ, f, g and h we provide a probabilistic representation of the gradient of v. The latter is used, along with more technical requirements on f, g and h, to prove that the optimal stopping boundary may be expressed in terms of a locally Lipschitz continuous function $b:[0,T]\times\mathbb{R}^{d-1}\to\mathbb{R}$. One of the main features in our work is that we do not assume uniform non-degeneracy of the diffusion so that standard results based on PDE theory cannot be easily applied.

It is well known that optimal stopping theory goes hand in hand with the theory of free boundary problems in PDE and the question of regularity of optimal stopping boundaries (free boundaries) has been the object of intensive study. The one dimensional case $d=1$ attracted

*School of Mathematics, University of Leeds, Woodhouse Lane LS2 9JT, Leeds, United Kingdom. t.deangelis@leeds.ac.uk

†Dipartimento di Metodi e Modelli per l'Economia, il Territorio e la Finanza, Sapienza University of Rome, Via del Castro Laurenziano 9, 00161, Rome, Italy. gabriele.stabile@uniroma1.it

the interest of several mathematicians who developed approaches ranging from probability to analysis. Early contributions to the topic were made in [22], [26] and [37], among others. In [22] and [37] it was proven that the free-boundary b is differentiable in the open interval $(0, T)$ for a certain class of problems involving one-dimensional Brownian motion or solutions of one-dimensional SDEs with regular coefficents. Other papers employing PDE methods are for example [7], [32], where infinite differentiability of the free boundary in the Stefan problem is proved, and [17] where C^1 regularity of the boundary is obtained for a certain class of variational problems. The study of the optimal boundary of the American put option is perhaps one of the most renowned examples in this field and for an overview of existing results one may refer to [1], [3], [8], [9], [15], [18], [21] [24], and [29] among others. Finally it is worth recalling that a thorough discussion of analytical methods for free boundary problems on $[0, T] \times \mathbb{R}$ related to the heat operator may be found in the monograph [6] (see also [16, Ch. 8]). In the latter, as well as in several of the above references, the first step in the analysis of the regularity of the free boundary is to prove that it is Lipschitz continuous or at least Hölder continuous with constant $\alpha > 1/2$.

There is also a large body of literature addressing similar questions in higher dimensions. Accounting in full for these results is a difficult task and it falls outside the reach of our work. However for our purposes it is interesting to recall the following fact: Lipschitz regularity for the free boundary of certain Stefan problems (with $d \geq 1$) can be upgraded to $C^{1,\alpha}$ regularity for some $\alpha \in (0, 1)$ and eventually to C^∞ regularity, under suitable technical conditions. Detailed derivations of this informal statement may be found in the monographs [5] and [27] and references therein (see also [25] for the study of American options written on several assets and with convex payoff).

In the literature on optimal stopping the vast majority of papers studying problems of the form (1.1) with $d = 1$ addresses the question of continuity of the boundary without looking at higher regularity (of course with the exception of the works mentioned above; see [11] for some results and further references). Moreover, even the question of continuity becomes difficult to handle for $d > 1$. In the case of $d = 2$ and $T = \infty$, specific examples were addressed in [13] and [19], while a more complete answer was recently provided in [30].

Notably, Shreve and Soner [34, 36] address a problem of singular stochastic control which is equivalent to one of optimal stopping of the form (1.1), and characterise the optimal boundary as a real-valued, Lipschitz continuous function on $[0, T] \times \mathbb{R}^{d-1}$, $d \geq 1$. In their work they employ the equivalence between the problem of singular control and the one of optimal stopping, and study the latter purely by means of PDE methods similar to those in [2]. Regularity of the free boundary is used in [34, 36] to obtain a classical solution to a variational problem with gradient constraint related to the singular control problem. It is worth mentioning that the same authors had previously shown C^2 regularity for the optimal boundary of a two-dimensional singular control problem on an infinite time horizon [35]. However in the setting of [35] we are not aware of any direct link to an optimal stopping problem and therefore it is harder to draw a parallel with our work.

From the above discussion we learn that a reasonable attempt towards the study of regularity for optimal boundaries in optimal stopping theory should start form establishing their Lipschitz continuity. Of course this can be achieved in several instances by the PDE methods illustrated in the references above but instead we aim at finding a fully probabilistic approach. Under assumptions similar to those adopted in [34, 36], our work not only serves the purpose of bridging the PDE literature and the probabilistic one but it also contributes new results.

One of our main contributions is to prove that for $d > 1$ local Lipschitz continuity of the optimal boundary can be obtained without requiring uniform ellipticity of the operator $\sigma\sigma^\top$ (see Theorems 4.11 and 4.12, and Example 2 in Section 5). Relaxing this requirement makes it difficult to apply standard PDE results (including [5] and [27]) and the methods used in [34, 36]

are no longer valid. In the special case $d = 1$ (see Theorem 4.3) we are able to localize the assumptions made in [34, 36] and in particular the one relative to the running cost, i.e. our function h. Such relaxation allows us to apply our results to a wider class of examples than the one previously covered. For instance we can apply them in problems of irreversible capacity expansions where the running profit is expressed by a Cob-Douglas-type production function (see, e.g. [10] and Example 1 in Section 5). A more detailed comparison between our setting and the one in [34, 36] is provided in Remark 4.13.

We also notice that our functional (1.1) allows a rather generic time-space dependence of the functions f, g and h, while at the same time the dynamic of X allows state dependent drifts and correlations between the driving noises (i.e. σ is not necessarily diagonal). For $d = 1$ a generic time dependence of f and h makes it extremely hard and often impossible to establish monotonicity of the optimal boundary as a function of time. The latter is normally a key feature in the study of the boundary's continuity. One advantage of our approach is that instead we do not need such monotonicity to establish Lipschitz continuity (see [14] for a recent application in actuarial context). Moreover if the boundary is Lipschitz then $v \in C^1([0,T) \times \mathbb{R})$ (see Remark 4.5).

Our method consists of two main steps which we can formally summarise as follows. In the first step we find a probabilistic representation of the time/space derivatives of the value function. The latter is then used in the second step along with the implicit function theorem to obtain bounds on the gradient of the optimal boundary. Notice that, while the second step is somehow in line with ideas in [36], the first step is entirely new.

It is important to remark that despite the technical assumptions that we make, one of the main contributions of our work is the methodology. As it is often the case in optimal stopping and free boundary problems, in order to be able to give general results, one has to impose fairly strong conditions on the problem data. However, when considering specific examples it is possible to find ways around the technicalities and still apply the same methods. This is indeed true also for the theory that we are developing here and in Section 5 we provide some examples of such extensions.

The rest of the paper is organised as follows. In Section 2 we provide a rigorous formulation of the problem outlined in (1.1) along with the standing assumptions. In Section 3 we obtain a probabilistic representation formula for the gradient $\nabla_x v$ and for bounds on the time derivative $\partial_t v$ (see Theorem 3.1). Some other technical estimates are performed before passing to Section 4. In the latter we finally give our main results regarding existence of a locally-Lipschitz continuous optimal boundary for problem (1.1). This result is given under three different sets of assumptions: in Theorem 4.3 for $d = 1$ and in Theorems 4.11 and 4.12 for $d \geq 2$. In Section 5 we show some applications of our results and their extensions in specific examples.

2 Setup and problem formulation

Consider a complete probability space $(\Omega, \mathcal{F}, \mathsf{P})$ equipped with the natural filtration $\mathbb{F} := (\mathcal{F}_t)_{t\geq 0}$ generated by a $\mathbb{R}^{d'}$-valued Brownian motion $(B_t)_{t\geq 0}$. Assume that $\mathbb{F}$ is completed with P-null sets and let $X \in \mathbb{R}^d$ evolve according to

$$X_t^x = x + \int_0^t \mu(X_s^x)ds + \sigma B_t, \quad t \geq 0, \tag{2.1}$$

where $\mu \in C^1(\mathbb{R}^d; \mathbb{R}^d)$ with sub-linear growth and σ is a $d \times d'$ matrix. We denote by $\langle \cdot\,, \cdot \rangle$ the scalar product in $\mathbb{R}^d$ and by $\|\cdot\|_d$ the Euclidean norm in $\mathbb{R}^d$. Notice that $\sigma\sigma^\top$ is assumed to be non-negative but not necessarily uniformly elliptic. This means that it may exist $\xi \in \mathbb{R}^d$ such that $\langle \sigma\sigma^\top \xi, \xi \rangle = 0$.

Throughout the paper we will often use $\mathsf{P}_{t,x}(\,\cdot\,) = \mathsf{P}(\,\cdot\,|X_t = x)$ and $\mathsf{P}_x = \mathsf{P}_{0,x}$, so that $\mathsf{E}_{t,x}f(X_s) = \mathsf{E}f(X_s^{t,x})$, $s \geq t$, for any function f which is Borel-measurable and integrable. With no loss of generality we will assume $\Omega = C([0,T];\mathbb{R}^d)$ so that $t \mapsto \omega(t)$ is the canonical process and $\theta_\cdot$ the shifting operator such that $\theta_s\omega(t) = \omega(t+s)$.

For $T \in (0,+\infty)$ we consider optimal stopping problems of the form

$$v(t,x) = \sup_{0\leq\tau\leq T-t} \mathsf{E}\left[\int_0^\tau h(t+s,X_s^x)ds + \mathbb{1}_{\{\tau<T-t\}}f(t+\tau,X_\tau^x) + \mathbb{1}_{\{\tau=T-t\}}g(X_\tau^x)\right] \tag{2.2}$$

where f, g and h are real-valued with $f \in C^{1,2}([0,T]\times\mathbb{R}^d)$, $h \in C^{1,1}([0,T]\times\mathbb{R}^d)$ and $g \in C^2(\mathbb{R}^d)$. In the infinite horizon case, i.e. $T = +\infty$, we consider

$$v(t,x) = \sup_{\tau\geq 0} \mathsf{E}\left[\int_0^\tau h(t+s,X_s^x)ds + f(t+\tau,X_\tau^x)\right] \tag{2.3}$$

with f and h as above and, according to [33, Ch. 3], we set

$$\mathbb{1}_{\{\tau=+\infty\}}f(t+\tau,X_\tau^x) := \limsup_{s\to\infty} f(s,X_s^x), \quad \mathsf{P}\text{-a.s.}$$

In what follows conditions at T for the terminal value $g(X_T)$ are understood to hold only for $T < +\infty$ and can always be neglected for $T = +\infty$. From now on we assume that for all $(t,x) \in [0,T]\times\mathbb{R}^d$ it holds

$$\mathsf{E}\Big[\int_0^{T-t} |h(t+s,X_s^x)|ds + |g(X_{T-t}^x)| + \sup_{0\leq s\leq T-t} |f(t+s,X_s^x)|\Big] < +\infty. \tag{2.4}$$

Moreover, if $T = +\infty$ then we also assume

$$\limsup_{s\to\infty} f(t+s,X_s^x) = \lim_{s\to\infty} f(t+s,X_s^x) = 0, \quad \mathsf{P}\text{-a.s.} \tag{2.5}$$

Both assumptions are fulfilled in the examples of Section 5.

Remark 2.1. *Notice that the dynamic* (2.1) *and the optimisation problem* (2.2) *are general enough to include for example models involving geometric Brownian motion and Ornstein-Uhlenbeck.*

To avoid further technicalities we also assume that v is a lower semi-continuous function. Often such regularity (or even continuity) is easy to check in specific examples (e.g., those in Section 5). There also exist mild sufficient conditions that guarantee lower semi-continuity of v in more general settings (see for instance [33, Ch. 3]. See also Remark 2.10 and eq. (2.2.80) in [31, Ch.I, Sec. 2]).

The continuation set $\mathcal{C}$ and the stopping set $\mathcal{S}$ are given by

$$\mathcal{C} := \{(t,x) \in [0,T)\times\mathbb{R}^d \,:\, v(t,x) > f(t,x)\} \tag{2.6}$$

$$\mathcal{S} := \{(t,x) \in [0,T)\times\mathbb{R}^d \,:\, v(t,x) = f(t,x)\} \cup (\{T\}\times\mathbb{R}). \tag{2.7}$$

From standard optimal stopping theory we know that, in our setting, (2.4) and lower semi-continuity of v are sufficient for the optimality of

$$\tau_*(t,x) = \inf\{s \in [0,T-t] : (t+s,X_s^x) \in \mathcal{S}\} \tag{2.8}$$

provided that $f(T,x) \leq g(x)$, if $T < +\infty$ (see [31, Ch. I, Sec. 2, Cor. 2.9]). For the infinite horizon case notice that if $\mathsf{P}_{t,x}(\tau_* < +\infty) < 1$, then there is no optimal stopping time and τ_* is a

(optimal) Markov time (according to the terminology in [33, Ch. 3, Thm. 3]). However methods used in the next sections work for both finite and infinite values of τ_* thanks to (2.5).

For arbitrary $(t,x) \in [0,T] \times \mathbb{R}^d$ let

$$Y_s := v(t+s, X_s^x) + \int_0^s h(t+u, X_u^x)du.$$

Since v is lower semi-continuous and using standard results in optimal stopping (see [31, Ch. I, Sec. 2, Thm. 2.4]) we have that $(Y_s)_{0\le s\le T-t}$ is P-a.s. right-continuous and

$$Y_s \quad \text{is a supermartingale for } s \in [0, T-t], \tag{2.9}$$

$$Y_{s\wedge\tau_*} \quad \text{is a martingale for } s \in [0, T-t]. \tag{2.10}$$

Notice in particular that since Y is right-continuous then the process $s \mapsto v(t+s, X_s^x)$ is P-a.s. right continuous as well. As a note of caution we remark that if $T = +\infty$ then (2.9) continues to hold on $[0, +\infty]$ because Y is a uniformly integrable super-martingale thanks to (2.4) (see, e.g. [33, Thm. 9, Ch. 1]). Instead, (2.10) only holds on $[0, \infty)$.

We denote by $\mathcal{L}$ the infinitesimal generator associated to X and in particular we have

$$\mathcal{L}F(x) = \frac{1}{2}\sum_{i,j=1}^d (\sigma\sigma^\top)_{i,j}\frac{\partial^2 F}{\partial x_i \partial x_j}(x) + \sum_{i=1}^d \mu_i(x)\frac{\partial F}{\partial x_i}(x), \qquad F \in C^2(\mathbb{R}^d;\mathbb{R}). \tag{2.11}$$

For future frequent use we also introduce the following notation

$$m(t,x) := (\partial_t f + \mathcal{L}f)(t,x) \quad \text{and} \quad n(x) := \mathcal{L}g(x). \tag{2.12}$$

Since $\mu \in C^1(\mathbb{R}^d;\mathbb{R}^d)$ then the flow $x \mapsto X^x$ is differentiable ([28], Chapter V.7). Here we denote the initial point in (2.1) by $x = (x_1, \dots, x_d)$, the i-th component of X^x by $X^{x,i}$, the partial derivative with respect to x_k by $\partial_k = \frac{\partial}{\partial x_k}$, and the derivative of X^x with respect to the initial point x_k by $\partial_k X^x = (\partial_k X^{x,1}, \dots \partial_k X^{x,d})$. We define the process ∂X^x as a $d \times d$ matrix with entries $\partial_k X^{x,j}$ for $j,k = 1, \dots d$ and the maps $t \mapsto \partial_k X_t^{x,j}$ are P-a.s. continuous with dynamics given by

$$\partial_k X_t^{x,j} = \delta_{j,k} + \int_0^t \sum_{\ell=1}^d \partial_\ell \mu_j(X_s^x)\partial_k X_s^{x,\ell} ds = \delta_{j,k} + \int_0^t \langle \nabla_x \mu_j(X_s^x), \partial_k X_s^x\rangle ds. \tag{2.13}$$

In what follows we also assume that for any compact $K \subset \mathbb{R}^d$ it holds

$$\sup_{x\in K} \mathsf{E}\left[\sup_{0\le t\le T} \|\partial_k X_t^x\|_d^2\right] < +\infty \quad \text{for all } k = 1, \dots d. \tag{2.14}$$

The next will be a standing assumption throughout the paper

Assumption 2.2 (**Regularity** f, g, h.)**.** *For any compact $K \subset [0,T] \times \mathbb{R}^d$, there exists $c_K > 0$ such that for all $(t,x) \in K$ we have*

$$\mathsf{E}\left[\int_0^{T-t} \|\nabla_x h(t+s, X_s^x)\|_d^2 ds + \sup_{0\le s\le T-t} \|\nabla_x f(t+s, X_s^x)\|_d^2 + \|\nabla_x g(X_{T-t}^x)\|_d^2\right] \le c_K,$$

$$\mathsf{E}\left[\int_0^{T-t} |\partial_t h(t+s, X_s^x)|^2 ds + \sup_{0\le s\le T-t} |\partial_t f(t+s, X_s^x)|^2 + |h(T, X_{T-t}^x) + n(X_{T-t}^x)|^2\right] \le c_K.$$

It is important to remark that Assumption 2.2 and (2.14) are used in Theorem 3.1 in order to pass limits under expectations, thanks to uniform integrability. Therefore, in specific examples one can expect to find weaker sufficient conditions that allow this step in the proof.

3 Properties of the value function

In this section we provide useful bounds for the gradient of the value function v and some other technical results. These are obtained by often using the following condition

($\mathbf{A}_1$) Terminal value. If $T < +\infty$ we have $g(x) \geq f(T, x)$.

Before stating the next theorem it is useful to introduce the functions

$$\overline{v}(t,x) = \mathsf{E}\Big[\int_0^{\tau_*} \partial_t h(t+s, X_s^x)ds + \mathbb{1}_{\{\tau_*<T-t\}}\partial_t f(t+\tau_*, X_{\tau_*}^x) - \mathbb{1}_{\{\tau_*=T-t\}}\left(h(T, X_{T-t}^x) + n(X_{T-t}^x)\right)\Big] \tag{3.1}$$

and

$$\underline{v}(t,x) = \mathsf{E}\Big[\int_0^{\tau_*} \partial_t h(t+s, X_s^x)ds + \mathbb{1}_{\{\tau_*<T-t\}}\partial_t f(t+\tau_*, X_{\tau_*}^x) - \mathbb{1}_{\{\tau_*=T-t\}}\Big(|h(T, X_{T-t}^x) + n(X_{T-t}^x)| + |\partial_t f(T, X_{T-t}^x)|\Big)\Big]. \tag{3.2}$$

Theorem 3.1. *Assume condition (A_1). Then the value function v is locally Lipschitz continuous on $[0,T]\times\mathbb{R}^d$ and for a.e. (t,x) we have*

$$\partial_k v(t,x) = \mathsf{E}\Big[\int_0^{\tau_*} \langle\nabla_x h(t+s, X_s^x), \partial_k X_s^x\rangle ds + \mathbb{1}_{\{\tau_*<T-t\}}\langle\nabla_x f(t+\tau_*, X_{\tau_*}^x), \partial_k X_{\tau_*}^x\rangle + \mathbb{1}_{\{\tau_*=T-t\}}\langle\nabla_x g(X_{T-t}^x), \partial_k X_{T-t}^x\rangle\Big] \tag{3.3}$$

and

$$\underline{v}(t,x) \leq \partial_t v(t,x) \leq \overline{v}(t,x). \tag{3.4}$$

Proof. Step 1. (Spatial derivative). Here we show that $v(t,\cdot)$ is locally Lipschitz and (3.3) holds for a.e. $x\in\mathbb{R}^d$ and each given $t\in[0,T]$ (notice that the null set where $v(t,\cdot)$ is not differentiable may a priori depend on t). First we obtain bounds for the left and right derivative of $v(t,\cdot)$.

Fix $(t,x)\in[0,T]\times\mathbb{R}^d$ and take $\varepsilon>0$. For an arbitrary k we denote for simplicity $x_\varepsilon = (x_1,\dots x_k+\varepsilon,\dots x_d)$, and consider the processes $X^{x_\varepsilon} = (X^{x_\varepsilon,1},\dots\dots X^{x_\varepsilon,d})$ and $X^x = (X^{x,1},\dots X^{x,d})$. We remark that all components of the vector process X^{x_ε} are affected by the shift in the initial point.

We denote by $\tau = \tau_*(t,x)$ the optimal stopping time (independent of ε) for the problem with initial data (t,x). Using such optimality we first obtain

$$\begin{aligned} v(t,x_\varepsilon) - v(t,x) \\ \geq & \mathsf{E}\left[\int_0^\tau \left(h(t+s, X_s^{x_\varepsilon}) - h(t+s, X_s^x)\right)ds + \mathbb{1}_{\{\tau<T-t\}}\left(f(t+\tau, X_\tau^{x_\varepsilon}) - f(t+\tau, X_\tau^x)\right)\right] \\ & + \mathsf{E}\left[\mathbb{1}_{\{\tau=T-t\}}\left(g(X_{T-t}^{x_\varepsilon}) - g(X_{T-t}^x)\right)\right]. \end{aligned}$$

Dividing both sides of the above expression by ε and recalling Assumption 2.2 and (2.14) we can pass to the limit as $\varepsilon\to 0$ and use dominated convergence to conclude that

$$\begin{aligned} \liminf_{\varepsilon\to 0}\frac{v(t,x_\varepsilon)-v(t,x)}{\varepsilon} \\ \geq & \mathsf{E}\left[\int_0^\tau \langle\nabla_x h(t+s, X_s^x), \partial_k X_s^x\rangle ds + \mathbb{1}_{\{\tau<T-t\}}\langle\nabla_x f(t+\tau, X_\tau^x), \partial_k X_\tau^x\rangle\right] \\ & + \mathsf{E}\left[\mathbb{1}_{\{\tau=T-t\}}\langle\nabla_x g(X_{T-t}^x), \partial_k X_{T-t}^x\rangle\right]. \end{aligned} \tag{3.5}$$

To obtain a reverse inequality, pick $\delta > 0$ and denote $x_\delta = (x_1, \dots x_k - \delta, \dots x_d)$ and $X^{x_\delta} = (X^{x_\delta,1}, \dots X^{x_\delta,d})$. Since τ is optimal in $v(t,x)$ and sub-optimal in $v(t,x_\delta)$ we have

$$\begin{aligned} &v(t,x) - v(t,x_\delta) \\ &\quad \leq \mathsf{E}\left[\int_0^\tau \left(h(t+s,X_s^x) - h(t+s,X_s^{x_\delta})\right) ds + \mathbb{1}_{\{\tau<T-t\}}\left(f(t+\tau,X_\tau^x) - f(t+\tau,X_\tau^{x_\delta})\right)\right] \\ &\qquad + \mathsf{E}\left[\mathbb{1}_{\{\tau=T-t\}}\left(g(X_{T-t}^x) - g(X_{T-t}^{x_\delta})\right)\right]. \end{aligned}$$

Dividing both sides by δ, taking limits and using dominated convergence again we obtain

$$\begin{aligned} &\limsup_{\delta\to 0} \frac{v(t,x) - v(t,x_\delta)}{\delta} \\ &\quad \leq \mathsf{E}\left[\int_0^\tau \langle \nabla_x h(t+s,X_s^x), \partial_k X_s^x\rangle ds + \mathbb{1}_{\{\tau<T-t\}}\langle \nabla_x f(t+\tau,X_\tau^x), \partial_k X_\tau^x\rangle\right] \\ &\qquad + \mathsf{E}\left[\mathbb{1}_{\{\tau=T-t\}}\langle \nabla_x g(X_{T-t}^x), \partial_k X_{T-t}^x\rangle\right]. \end{aligned} \tag{3.6}$$

Now, (3.5) gives a lower bound for the right derivative with respect to x_k whereas (3.6) provides an upper bound for the corresponding left derivative. If x is a point of differentiability of $v(t,\cdot)$ then (3.5) and (3.6) imply that (3.3) holds at that point. It remains to show that $v(t,\cdot)$ is locally Lipschitz so that a.e. $x \in \mathbb{R}^d$ is a point of differentiability.

With the same notation as above let $\tau_\varepsilon = \tau_*(t,x_\varepsilon)$ be optimal for the problem with initial data (t,x_ε). By analogous arguments to those used previously and using Assumption 2.2 and (2.14) we find

$$\begin{aligned} &v(t,x_\varepsilon) - v(t,x) \\ &\leq \mathsf{E}\left[\int_0^{\tau_\varepsilon} \left(h(t+s,X_s^{x_\varepsilon}) - h(t+s,X_s^x)\right) ds + \mathbb{1}_{\{\tau_\varepsilon<T-t\}}\left(f(t+\tau_\varepsilon,X_{\tau_\varepsilon}^{x_\varepsilon}) - f(t+\tau_\varepsilon,X_{\tau_\varepsilon}^x)\right)\right] \\ &\quad + \mathsf{E}\left[\mathbb{1}_{\{\tau_\varepsilon=T-t\}}\left(g(X_{T-t}^{x_\varepsilon}) - g(X_{T-t}^x)\right)\right] \leq c(t,x)\varepsilon, \end{aligned} \tag{3.7}$$

for some $c(t,x) > 0$, which is uniform for (t,x) on a compact. Notice also that for the last inequality we have used

$$\left\|X_{\tau_\varepsilon}^{x_\varepsilon} - X_{\tau_\varepsilon}^x\right\|_d \leq \varepsilon \cdot \sum_k \sup_{0\leq s\leq T} \|\partial_k X_s^z\|_d$$

by the mean value theorem, with suitable $z \in \mathbb{R}^d$ such that $\|z-x\|_d \leq \varepsilon$.

The estimate in (3.7) and (3.5) imply $|v(t,x_\varepsilon) - v(t,x)| \leq \hat{c}(t,x)\varepsilon$, for some other constant $\hat{c}(t,x) > 0$ which can be taken uniform over compact sets. Symmetric arguments allow to prove also that $|v(t,x_\delta) - v(t,x)| \leq \hat{c}(t,x)\delta$, with x_δ as in (3.6).

Step 2. (Time derivative). Here we show that $t \mapsto v(t,x)$ is locally Lipschitz and (3.4) holds for a.e. $t \in [0,T]$ and each given $x \in \mathbb{R}^d$. We start by providing bounds for the left and right derivatives of $v(\,\cdot\,,x)$.

Fix $(t,x) \in [0,T] \times \mathbb{R}^d$ and let $\varepsilon > 0$. Then letting $\tau = \tau_*(t,x)$ optimal for the problem with initial data (t,x) we notice that τ is admissible for the problem with initial data $(t-\varepsilon,x)$. Using (2.9) and (2.10) we obtain the following upper bound.

$$\begin{aligned} v(t,x) - v(t-\varepsilon,x) \leq \mathsf{E}\Big[&\int_0^\tau \left(h(t+s,X_s^x) - h(t-\varepsilon+s,X_s^x)\right) ds \\ &+ v(t+\tau,X_\tau^x) - v(t-\varepsilon+\tau,X_\tau^x)\Big]. \end{aligned} \tag{3.8}$$

Now we notice that since $v \geq f$ on $[0,T]\times\mathbb{R}^d$ and $v = f$ in $\mathcal{S}$, by right continuity of $v(t+\cdot, X^x_\cdot)$ one has

$$\begin{aligned} &v(t+\tau, X^x_\tau) - v(t-\varepsilon+\tau, X^x_\tau) \leq f(t+\tau, X^x_\tau) - f(t-\varepsilon+\tau, X^x_\tau) \quad \text{on } \{\tau < T-t\} \\ &v(t+\tau, X^x_\tau) - v(t-\varepsilon+\tau, X^x_\tau) \leq g(X^x_{T-t}) - v(T-\varepsilon, X^x_{T-t}) \quad \text{on } \{\tau = T-t\}. \end{aligned}$$

Moreover from (2.2) we also have

$$\begin{aligned} v(T-\varepsilon, X^x_{T-t}) \geq& \mathsf{E}_{X^x_{T-t}}\left[\int_0^\varepsilon h(T-\varepsilon+s, X_s)ds + g(X_\varepsilon)\right] \\ =& g(X^x_{T-t}) + \mathsf{E}_{X^x_{T-t}}\left[\int_0^\varepsilon \left(h(T-\varepsilon+s, X_s) + n(X_s)\right) ds\right]. \end{aligned}$$

Collecting the above estimates and using the mean value theorem we conclude

$$\begin{aligned} &\frac{1}{\varepsilon}(v(t,x) - v(t-\varepsilon, x)) \\ \leq& \mathsf{E}\left[\int_0^\tau \partial_t h(t-\varepsilon'_s+s, X^x_s)ds + \mathbb{1}_{\{\tau<T-t\}}\partial_t f(t-\varepsilon''_\tau+\tau, X^x_\tau)\right] \\ &- \mathsf{E}_x\left[\mathbb{1}_{\{\tau=T-t\}}\mathsf{E}_{X_{T-t}}\left[\frac{1}{\varepsilon}\int_0^\varepsilon \left(h(T-\varepsilon+s, X_s) + n(X_s)\right) ds\right]\right] \end{aligned} \tag{3.9}$$

for ε'_s and ε''_τ in $[0,\varepsilon]$. Letting $\varepsilon \to 0$ and using Assumption 2.2 we get

$$\begin{aligned} &\limsup_{\varepsilon\to 0}\frac{v(t,x) - v(t-\varepsilon,x)}{\varepsilon} \\ &\quad\leq \mathsf{E}\Big[\int_0^\tau \partial_t h(t+s, X^x_s)ds + \mathbb{1}_{\{\tau<T-t\}}\partial_t f(t+\tau, X^x_\tau)\Big] \\ &\qquad - \mathsf{E}\Big[\mathbb{1}_{\{\tau=T-t\}}\left(h(T, X^x_{T-t}) + n(X^x_{T-t})\right)\Big]. \end{aligned} \tag{3.10}$$

To prove a reverse inequality we notice that $\tau \wedge (T-t-\varepsilon)$ is admissible for the problem with initial data $(t+\varepsilon, x)$, so that by using (2.9) and (2.10) and arguing as above we obtain

$$\begin{aligned} &v(t+\varepsilon, x) - v(t,x) \\ \geq& \mathsf{E}\Big[\int_0^{\tau\wedge(T-t-\varepsilon)} \left(h(t+\varepsilon+s, X^x_s) - h(t+s, X^x_s)\right) ds - \mathbb{1}_{\{\tau>T-t-\varepsilon\}}\int_{T-t-\varepsilon}^\tau h(t+s, X^x_s)ds\Big] \\ &+ \mathsf{E}\Big[\mathbb{1}_{\{\tau\leq T-t-\varepsilon\}}\left(f(t+\varepsilon+\tau, X^x_\tau) - f(t+\tau, X^x_\tau)\right)\Big] \\ &+ \mathsf{E}\Big[\mathbb{1}_{\{\tau>T-t-\varepsilon\}}\left(g(X^x_{T-t-\varepsilon}) - v(t+\tau, X^x_\tau)\right)\Big]. \end{aligned} \tag{3.11}$$

We can collect the two terms with the indicator of $\{\tau > T-t-\varepsilon\}$, use iterated conditioning and the martingale property (2.10) to get

$$\begin{aligned} &\mathsf{E}\Big[\mathbb{1}_{\{\tau>T-t-\varepsilon\}}\left(g(X^x_{T-t-\varepsilon}) - v(t+\tau, X^x_\tau) - \int_{T-t-\varepsilon}^\tau h(t+s, X^x_s)ds\right)\Big] \\ =& \mathsf{E}\Big[\mathbb{1}_{\{\tau>T-t-\varepsilon\}}\left(g(X^x_{T-t-\varepsilon}) - \mathsf{E}\left[v(t+\tau, X^x_\tau) + \int_{T-t-\varepsilon}^\tau h(t+s, X^x_s)ds\Big|\mathcal{F}_{T-t-\varepsilon}\right]\right)\Big] \\ =& \mathsf{E}\Big[\mathbb{1}_{\{\tau>T-t-\varepsilon\}}\left(g(X^x_{T-t-\varepsilon}) - v(T-\varepsilon, X^x_{T-t-\varepsilon})\right)\Big]. \end{aligned}$$

To estimate the last term we argue as follows

$$
\begin{aligned}
&v(T-\varepsilon, X^x_{T-t-\varepsilon})\\
&=\operatorname*{ess\,sup}_{0\le\sigma\le\varepsilon} \mathsf{E}_{X^x_{T-t-\varepsilon}}\left[\int_0^\sigma h(T-\varepsilon+s,X_s)ds + \mathbb{1}_{\{\sigma<\varepsilon\}} f(T-\varepsilon+\sigma,X_\sigma) + \mathbb{1}_{\{\sigma=\varepsilon\}} g(X_\varepsilon)\right]\\
&=\operatorname*{ess\,sup}_{0\le\sigma\le\varepsilon} \mathsf{E}_{X^x_{T-t-\varepsilon}}\left[\int_0^\sigma h(T-\varepsilon+s,X_s)ds + g(X_\sigma) + \mathbb{1}_{\{\sigma<\varepsilon\}}\left(f(T-\varepsilon+\sigma,X_\sigma)-g(X_\sigma)\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
&=g(X^x_{T-t-\varepsilon})\\
&\quad+\operatorname*{ess\,sup}_{0\le\sigma\le\varepsilon} \mathsf{E}_{X^x_{T-t-\varepsilon}}\Big[\int_0^\sigma \left(h(T-\varepsilon+s,X_s)+n(X_s)\right)ds + \mathbb{1}_{\{\sigma<\varepsilon\}}\Big(f(T,X_\sigma)-g(X_\sigma)\Big)\\
&\qquad\qquad - \mathbb{1}_{\{\sigma<\varepsilon\}}\int_{T-\varepsilon+\sigma}^T \partial_t f(u,X_\sigma)du\Big].
\end{aligned}
$$

Using that $f(T,x)\le g(x)$ by condition (A_1), we get

$$
\begin{aligned}
&v(T-\varepsilon, X^x_{T-t-\varepsilon})\\
&\le g(X^x_{T-t-\varepsilon})\\
&\quad+\operatorname*{ess\,sup}_{0\le\sigma\le\varepsilon} \mathsf{E}_{X^x_{T-t-\varepsilon}}\Big[\int_0^\sigma (h(T-\varepsilon+s,X_s)+n(X_s))\,ds + \int_{T-\varepsilon+\sigma}^T |\partial_t f(u,X_\sigma)|du\Big]\\
&\le g(X^x_{T-t-\varepsilon}) + \mathsf{E}_{X^x_{T-t-\varepsilon}}\Big[\int_0^\varepsilon |h(T-\varepsilon+s,X_s)+n(X_s)|ds\Big]\\
&\quad+\mathsf{E}_{X^x_{T-t-\varepsilon}}\Big[\int_0^\varepsilon \sup_{r\le\varepsilon}|\partial_t f(T-\varepsilon+s\wedge(\varepsilon-r)+r,X_r)|ds\Big].
\end{aligned}
\tag{3.12}
$$

Plugging the estimates above inside (3.11) we then obtain

$$
\begin{aligned}
&\frac{1}{\varepsilon}(v(t+\varepsilon,x)-v(t,x))\\
&\ge \mathsf{E}\left[\int_0^{\tau\wedge(T-t-\varepsilon)} \partial_t h(t+\varepsilon'_s+s,X^x_s)ds + \mathbb{1}_{\{\tau\le T-t-\varepsilon\}}\partial_t f(t+\varepsilon''_\tau+\tau,X^x_\tau)\right]\\
&\quad-\mathsf{E}_x\left[\mathbb{1}_{\{\tau>T-\varepsilon-t\}}\mathsf{E}_{X_{T-t-\varepsilon}}\Big[\frac{1}{\varepsilon}\int_0^\varepsilon |h(T-\varepsilon+s,X_s)+n(X_s)|ds\Big]\right]\\
&\quad-\mathsf{E}_x\left[\mathbb{1}_{\{\tau>T-\varepsilon-t\}}\mathsf{E}_{X_{T-t-\varepsilon}}\Big[\frac{1}{\varepsilon}\int_0^\varepsilon \sup_{r\le\varepsilon}|\partial_t f(T-\varepsilon+s\wedge(\varepsilon-r)+r,X_r)|ds\Big]\right]
\end{aligned}
$$

for suitable ε'_s and ε''_τ in $[0,\varepsilon]$. Taking limits as $\varepsilon\to 0$ we conclude

$$
\begin{aligned}
&\liminf_{\varepsilon\to 0}\frac{v(t+\varepsilon,x)-v(t,x)}{\varepsilon}\\
&\quad\ge \mathsf{E}\Big[\int_0^\tau \partial_t h(t+s,X^x_s)ds + \mathbb{1}_{\{\tau<T-t\}}\partial_t f(t+\tau,X^x_\tau)\Big]\\
&\qquad-\mathsf{E}\Big[\mathbb{1}_{\{\tau=T-t\}}\left(|h(T,X^x_{T-t})+n(X^x_{T-t})|+|\partial_t f(T,X^x_{T-t})|\right)\Big].
\end{aligned}
\tag{3.13}
$$

So far we have established a lower bound for the right-derivative and an upper bound for the left-derivative of $v(\cdot,x)$. Hence (3.4) holds for all $t\in[0,T]$ at which $v(\cdot,x)$ is differentiable, thanks to (3.10) and (3.13). Next we prove that $v(\,\cdot\,,x)$ is indeed locally Lipschitz so that (3.4) holds for a.e. $t\in[0,T]$ and each given $x\in\mathbb{R}^d$.

Let us set $\tau_\varepsilon := \tau_*(t-\varepsilon, x)$ and notice that $\tau_\varepsilon \wedge (T-t)$ is admissible for the problem with initial data (t,x). Therefore arguing as in (3.11) we get

$$\begin{aligned}
&v(t,x) - v(t-\varepsilon,x)\\
&\geq \mathsf{E}\Big[\int_0^{\tau_\varepsilon\wedge(T-t)} (h(t+s,X_s^x) - h(t-\varepsilon+s,X_s^x))\,ds - \mathbb{1}_{\{\tau_\varepsilon>T-t\}}\int_{T-t}^{\tau_\varepsilon} h(t-\varepsilon+s,X_s^x)ds\Big]\\
&\quad + \mathsf{E}\Big[\mathbb{1}_{\{\tau_\varepsilon\leq T-t\}}\left(f(t+\tau_\varepsilon,X_{\tau_\varepsilon}^x) - f(t-\varepsilon+\tau_\varepsilon,X_{\tau_\varepsilon}^x)\right)\Big]\\
&\quad + \mathsf{E}\Big[\mathbb{1}_{\{\tau_\varepsilon>T-t\}}\left(g(X_{T-t}^x) - v(t-\varepsilon+\tau_\varepsilon,X_{\tau_\varepsilon}^x)\right)\Big]. \qquad (3.14)
\end{aligned}$$

Repeating step by step the arguments that follow (3.11) we obtain

$$\begin{aligned}
&\frac{1}{\varepsilon}(v(t,x) - v(t-\varepsilon,x))\\
&\geq \mathsf{E}\left[\int_0^{\tau_\varepsilon\wedge(T-t)} \partial_t h(t-\varepsilon_s'+s,X_s^x)ds + \mathbb{1}_{\{\tau_\varepsilon\leq T-t\}}\partial_t f(t-\varepsilon_{\tau_\varepsilon}''+\tau_\varepsilon,X_{\tau_\varepsilon}^x)\right]\\
&\quad - \mathsf{E}_x\left[\mathbb{1}_{\{\tau_\varepsilon>T-t\}}\mathsf{E}_{X_{T-t}}\Big[\frac{1}{\varepsilon}\int_0^\varepsilon |h(T-\varepsilon+s,X_s)+n(X_s)|ds\Big]\right]\\
&\quad - \mathsf{E}_x\left[\mathbb{1}_{\{\tau_\varepsilon>T-t\}}\mathsf{E}_{X_{T-t}}\Big[\frac{1}{\varepsilon}\int_0^\varepsilon \sup_{r\leq\varepsilon}|\partial_t f(T-\varepsilon+s\wedge(\varepsilon-r)+r,X_r)|ds\Big]\right],
\end{aligned}$$

with ε_s' and $\varepsilon_{\tau_\varepsilon}''$ in $[0,\varepsilon]$. Using Assumption 2.2 and the above expression it is clear that we can find $c(t,x)>0$, which is uniform for (t,x) in a compact and such that $v(t,x)-v(t-\varepsilon,x) \geq -c(t,x)\,\varepsilon$. The latter, together with (3.10) imply that $|v(t,x)-v(t-\varepsilon,x)| \leq \hat{c}(t,x)\varepsilon$ for some other $\hat{c}(t,x)>0$ which is uniform on compact sets. A symmetric argument can be used to obtain an analogous bound for $|v(t+\varepsilon,x)-v(t,x)|$ and therefore $v(\,\cdot\,,x)$ is indeed locally Lipschitz.

Step 3. (Lipschitz property). To complete the proof it only remains to observe that, combining results in step 1 and 2 above, we obtain that v is locally Lipschitz on $[0,T]\times\mathbb{R}^d$. Hence, it is differentiable for a.e. $(t,x)\in[0,T]\times\mathbb{R}^d$ and (3.3) and (3.4) hold a.e. as claimed. □

Remark 3.2. *It is important to notice that the results of Theorem 3.1 hold in the same form when considering a state dependent diffusion coefficient $\sigma(x)$ in* (2.1)*, provided that $\sigma_{ij} \in C^1(\mathbb{R}^d;\mathbb{R})$. Indeed the proof remains exactly the same as we have never used the specific form of the dynamics of X in* (2.1)*.*

There is a simple and useful corollary to the theorem

Corollary 3.3. *Assume $T<+\infty$. Let condition (A_1) and one of the two conditions below hold*

(i) $g(x) = f(T,x)$, $x\in\mathbb{R}^d$,

(ii) $\exists\, c>0$ such that $h(T,x)+n(x) \geq -\partial_t f(T,x) - c$, for $x\in\mathbb{R}^d$.

Then for a.e. $(t,x)\in[0,T]\times\mathbb{R}^d$ and $\tau_* = \tau_*(t,x)$ we have

$$\partial_t v(t,x) \leq \mathsf{E}\left[\int_0^{\tau_*} \partial_t h(t+s,X_s^x)ds + \partial_t f(t+\tau_*,X_{\tau_*}^x)\right] + c\,\mathsf{P}(\tau_* = T-t) \qquad (3.15)$$

where $c=0$ if (i) holds.

Proof. Under (ii) the claim is trivial since $\partial_t v \leq \overline{v}$ and recalling (3.1). Under (i) instead, we notice that (3.8) in the proof of Theorem 3.1 may be bounded as follows

$$\begin{aligned} v(t,x) - v(t-\varepsilon,x) \leq & \mathsf{E}\Big[\int_0^\tau (h(t+s,X_s^x) - h(t-\varepsilon+s,X_s^x))\, ds \\ & + v(t+\tau,X_\tau^x) - v(t-\varepsilon+\tau,X_\tau^x)\Big] \\ \leq & \mathsf{E}\Big[\int_0^\tau (h(t+s,X_s^x) - h(t-\varepsilon+s,X_s^x))\, ds \\ & + f(t+\tau,X_\tau^x) - f(t-\varepsilon+\tau,X_\tau^x)\Big]. \end{aligned}$$

Then dividing by ε and taking limits as $\varepsilon \to 0$ we obtain (3.15), thanks to Assumption 2.2. □

Before concluding the section we provide two simple technical lemmas which will be useful in the next section.

Lemma 3.4. *For $k = 1, \dots d$ one has P-almost surely*

$$\sup_{0\leq t\leq T} \|\partial_k X_t^x\|_d^2 \leq 2\exp\left(2T\int_0^T \sum_{j=1}^d \|\nabla_x \mu_j(X_s^x)\|_d^2 ds\right). \tag{3.16}$$

Proof. By using $|a+b|^2 \leq 2(|a|^2+|b|^2)$ and Hölder inequality applied to (2.13) we get

$$\begin{aligned} \|\partial_k X_t^x\|_d^2 = & \sum_{j=1}^d \left(\delta_{j,k} + \int_0^t \langle \nabla_x \mu_j(X_s^x), \partial_k X_s^x\rangle ds\right)^2 \\ \leq & 2\left(1 + T\sum_{j=1}^d \int_0^t \|\nabla_x \mu_j(X_s^x)\|_d^2 \|\partial_k X_s^x\|_d^2 ds\right). \end{aligned}$$

An application of Gronwall's inequality concludes the proof. □

Lemma 3.5. *Let $q : \mathbb{R}^d \to \mathbb{R}$ be Borel-measurable and with growth*

$$|q(x)| \leq q_0(1 + \|x\|_d^p) \tag{3.17}$$

for some $q_0 > 0$ and $p \geq 1$. Assume $\|\mu(x)\|_d \leq C$ for all $x \in \mathbb{R}^d$ and a given constant $C > 0$. Then for $T < +\infty$ and any stopping time $\tau \in [t,T]$ we have

$$\mathsf{E}_{t,x}\left[\mathbb{1}_{\{\tau=T\}} q(X_\tau)\right] \leq K\,(1+\|x\|_d^p)(T-t)^{-1}\mathsf{E}_{t,x}(\tau - t), \quad (t,x) \in [0,T]\times\mathbb{R}^d \tag{3.18}$$

where $K > 0$ depends only on q_0, d, p, C and T.

Proof. Using polynomial growth and Hölder inequality, and by letting $c > 0$ be a constant that changes from line to line, we get

$$\begin{aligned} & \mathsf{E}_{t,x}\left[\mathbb{1}_{\{\tau=T\}} q(X_\tau)\right] \\ & \leq c\,\mathsf{E}_{t,x}\left[\mathbb{1}_{\{\tau=T\}}\left(1+\|X_\tau\|_d^p\right)\right] \\ & \leq c\,\mathsf{E}_{t,x}\left[\mathbb{1}_{\{\tau=T\}}\left(1+\|x\|_d^p + \left\|\int_t^\tau \mu(X_s)ds\right\|_d^p + \|\sigma(B_\tau - B_t)\|_d^p\right)\right] \\ & \leq c\Big((1+\|x\|_d^p)\mathsf{P}_{t,x}(\tau=T) + \mathsf{E}_{t,x}\Big[\mathbb{1}_{\{\tau=T\}}(\tau-t)^p\Big] \\ & \qquad + \sqrt{\mathsf{P}_{t,x}(\tau=T)}\sqrt{\mathsf{E}_{t,x}\left[\|\sigma(B_\tau-B_t)\|_d^{2p}\right]}\Big) \\ & \leq c\Big((1+\|x\|_d^p)\mathsf{P}_{t,x}(\tau=T) + \sqrt{\mathsf{P}_{t,x}(\tau=T)}\sqrt{\mathsf{E}_{t,x}\Big[(\tau-t)^p\Big]}\Big). \end{aligned}$$

Finally, by using Markov inequality

$$\mathsf{P}_{t,x}(\tau = T) = \mathsf{P}_{t,x}(\tau - t \geq T - t) \leq (T-t)^{-p}\mathsf{E}_{t,x}\left[(\tau - t)^p\right] \tag{3.19}$$

we obtain

$$\begin{aligned}
&\mathsf{E}_{t,x}\left[\mathbb{1}_{\{\tau=T\}}q(X_T)\right] \\
&\leq c\left((1+\|x\|_d^p)(T-t)^{-1}\mathsf{E}_{t,x}(\tau-t) + \tfrac{1}{(T-t)^{p/2}}\mathsf{E}_{t,x}\Big[(\tau-t)^p\Big]\right) \\
&\leq c\left(1+\|x\|_d^p + (T-t)^{p/2}\right)(T-t)^{-1}\mathsf{E}_{t,x}(\tau-t).
\end{aligned}$$

It is now immediate to obtain (3.18). □

4 Properties of the optimal boundary

In order to analyse the shape of the continuation set and later on the regularity of its boundary we need to restrict our assumptions. In particular in the following we will often need

(A_2) Terminal value. If $T < +\infty$ we have $\partial_1 g(x) \geq \partial_1 f(T,x)$.

(B) Drift. For $k = 2, \dots, d$ it holds $\mu_k(x) = \mu_k(x_2, \dots x_d)$, hence $\partial_1 X^{x,k} \equiv 0$ and from (2.13)

$$\partial_1 X_t^{x,1} = 1 + \int_0^t \partial_1\mu_1(X_s^x)\partial_1 X_s^{x,1}ds. \tag{4.1}$$

(C) Spatial monotonicity for $m+h$. Let $\partial_1 f \in C^{1,2}([0,T]\times\mathbb{R}^d)$ so that $m(t,x)$ in (2.12) is continuously differentiable with respect to x_1. Assume also that

$$\partial_1(h+m)(t,x) \geq 0, \qquad \text{for } (t,x) \in [0,T]\times\mathbb{R}^d, \tag{4.2}$$

and that for any compact $K \subseteq [0,T]\times\mathbb{R}^d$ we have

$$\sup_{(t,x)\in K} \mathsf{E}\left[\int_0^{T-t} |\partial_1 m(t+s, X_s^x)|^2 ds\right] < +\infty.$$

Condition (B) allows a tractable expression for $\partial_1 X^{x,1}$ (see (4.1)), which together with condition (C) provide a simple way of determining the shape of the continuation region (see the next proposition).

Proposition 4.1. *Assume conditions (A_1), (A_2), (B) and (C). Then the stopping region is characterised by a free boundary*

$$b : [0,T]\times\mathbb{R}^{d-1} \to \mathbb{R}\cup\{\pm\infty\} \tag{4.3}$$

such that

$$\mathcal{S} = \{(t, x_1, x_2, \dots x_d) \in [0,T]\times\mathbb{R}^d : x_1 \leq b(t, x_2, \dots x_d)\} \tag{4.4}$$

Proof. In order to prove the claim it is sufficient to show that $\partial_1(v-f)(t,x) \geq 0$. The latter indeed implies

$$(t, x_1, x_2, \dots x_d) \in \mathcal{S} \implies (t, x_1', x_2, \dots x_d) \in \mathcal{S} \quad \text{for all } x_1' \leq x_1.$$

The task is relatively easy thanks to (3.3). Notice that $\partial_1 X^x = (\partial_1 X^{x,1}, 0, \dots 0)$ due to condition (B) and

$$\partial_1 X_t^{x,1} = \exp\left(\int_0^t \partial_1 \mu_1(X_s^x) ds\right).$$

Then, using that $\partial_1 g(x) \geq \partial_1 f(T,x)$ due to condition (A_2), for a.e. $(t,x) \in [0,T] \times \mathbb{R}^d$ we get

$$\begin{aligned}\partial_1 v(t,x) =& \mathsf{E}\Big[\int_0^{\tau_*} \partial_1 h(t+s, X_s^x)\partial_1 X_s^{x,1} ds + \mathbb{1}_{\{\tau_* < T-t\}}\partial_1 f(t+\tau_*, X_{\tau_*}^x)\partial_1 X_{\tau_*}^{x,1} \\ &+ \mathbb{1}_{\{\tau_* = T-t\}}\partial_1 g(X_{T-t}^x)\partial_1 X_{T-t}^{x,1}\Big] \\ \geq& \mathsf{E}\Big[\int_0^{\tau_*} \partial_1 h(t+s, X_s^x)\partial_1 X_s^{x,1} ds + \partial_1 f(t+\tau_*, X_{\tau_*}^x)\partial_1 X_{\tau_*}^{x,1}\Big].\end{aligned} \tag{4.5}$$

Now an application of Dynkin's formula gives

$$\begin{aligned}&\mathsf{E}\left[\partial_1 X_{\tau_*}^{x,1}\partial_1 f(t+\tau_*, X_{\tau_*}^x)\right] \\ =&\partial_1 f(t,x) + \mathsf{E}\Big[\int_0^{\tau_*} e^{\int_0^s \partial_1\mu_1(X_u^x)du}\left(\partial_t\partial_1 f + \mathcal{L}(\partial_1 f) + \partial_1\mu_1\partial_1 f\right)(t+s, X_s^x)ds\Big] \\ =&\partial_1 f(t,x) + \mathsf{E}\Big[\int_0^{\tau_*} e^{\int_0^s \partial_1\mu_1(X_u^x)du}\partial_1 m(t+s, X_s^x)ds\Big]\end{aligned}$$

where we have used the easily verifiable equality $\partial_1 m(t,x) = \left(\partial_t\partial_1 f + \mathcal{L}(\partial_1 f) + \partial_1\mu_1\partial_1 f\right)(t,x)$.

Plugging the last expression in (4.5) and rearranging terms gives

$$\partial_1(v-f)(t,x) \geq \mathsf{E}\Big[\int_0^{\tau_*} e^{\int_0^s \partial_1\mu_1(X_u^x)du}\,\partial_1(h+m)(t+s, X_s^x)ds\Big] \geq 0 \tag{4.6}$$

thanks to (4.2). □

Remark 4.2. *It should be clear that a completely symmetric result holds if* (4.2) *is replaced by* $\partial_1(h+m) \leq 0$ *and similarly one assumes* $\partial_1 g(x) \leq \partial_1 f(T,x)$. *In such case arguments analogous to the ones employed to prove Proposition 4.1 may be instead used to prove that* $\mathcal{S} = \{(t, x_1, x_2, \dots x_d) \in [0,T] \times \mathbb{R}^d : x_1 \geq b(t, x_2, \dots x_d)\}$.

If (4.2) holds, then for each $(t, x_2, \dots x_d) \in [0,T] \times \mathbb{R}^{d-1}$ we can define

$$\gamma(t, x_2, \dots x_d) := \inf\{x_1 \in \mathbb{R} : (h+m)(t,x) > 0\} \tag{4.7}$$

with $\gamma(t, x_2, \dots x_d) = +\infty$ if the set is empty. It follows from standard arguments that the set

$$\mathcal{R} := \{(t,x) \in [0,T] \times \mathbb{R}^d : x_1 > \gamma(t, x_2, \dots x_d)\}$$

is contained in $\mathcal{C}$. Obviously if $d = 1$ then γ is a function of time only.

In the next sections we state the results concerning regularity of the optimal boundary.

4.1 Lipschitz boundary for $d = 1$

Here we prove Lipschitz continuity of b in the case $d = 1$ and for that we are going to need

(D) Bounds I. Let $f \in C^{2,3}([0,T] \times \mathbb{R})$ so that $m \in C^1([0,T] \times \mathbb{R})$. There exists $c > 0$ such that $\partial_t(h+m)(t,x) \leq c(1 + \partial_1(h+m)(t,x))$ for $(t,x) \in [0,T] \times \mathbb{R}$ and, for any compact $K \subseteq [0,T] \times \mathbb{R}$, we have

$$\sup_{(t,x)\in K} \mathsf{E}\left[\int_0^{T-t} |\partial_t m(t+s, X_s^x)|^2 ds\right] < +\infty.$$

We would like to remark that some of the assumptions we make for the proof below may be relaxed when the structure of f, g, h and μ is known explicitly. This fact will be illustrated in Example 1 in Section 5. Notice also that in this setting $\sigma \in \mathbb{R}^{d'}$ and $\sigma\sigma^\top = \|\sigma\|^2_{d'}$. In what follows $\overline{\mathcal{I}}$ denotes the closure of a set $\mathcal{I}$.

Theorem 4.3. *Assume that $d = 1$ and $\sigma\sigma^\top > 0$. Assume (A_1), (A_2), (C), (D) and, if $T < +\infty$, either (i) or (ii) from Corollary 3.3. Assume further that $\partial_1\mu(x) > -\mu$ for some $\mu > 0$ and there exists an interval $\mathcal{I} := (t_1, t_2)$ with $\overline{\mathcal{I}} \subset [0, T)$ and such that*

(i) $\overline{\gamma} := \sup_{t\in\overline{\mathcal{I}}} \gamma(t) < +\infty$ (see (4.7));

(ii) there exists $r \in (\overline{\gamma}, +\infty)$ and $\alpha_r > 0$ such that $\partial_1(h + m)(t, x) = \partial_x(h + m)(t, x) \geq \alpha_r$ for $(t, x) \in \overline{\mathcal{I}} \times (-\infty, r)$.

Then, for any $\varepsilon > 0$ there is $K_\varepsilon > 0$ such that b is (bounded) Lipschitz on $[t_1 + \varepsilon, t_2 - \varepsilon]$ with Lipschitz constant K_ε.

First we need a technical lemma, whose proof will be given after that of the theorem.

Lemma 4.4. *Under the same assumptions as in Theorem 4.3 we have*

(a) for any $t'_1 < t'_2$ such that $[t'_1, t'_2] \subseteq \overline{\mathcal{I}}$ it holds $\big((t'_1, t'_2) \times \mathbb{R}\big) \cap \mathcal{S} \neq \varnothing$;

(b) $\lim_{x\to-\infty}(v - f)(t, x) = 0$ for all $t \in \mathcal{I}$.

Proof of Theorem 4.3. We provide a full proof only for $T < +\infty$ but the same arguments hold for $T = +\infty$ up to simple minor changes. We set $w := v - f$ and we use the notation ∂_x in place of ∂_1 and $\mu'(x) := \partial_x\mu(x)$.

In this setting the diffusion is uniformly non degenerate. Therefore, letting U be an open rectangle in the (t, x)-plane whose closure $\overline{U}$ is contained in $\mathcal{C}$, and denoting its parabolic boundary by $\partial_P U$, we have that $v \in C^{1,2}(U) \cap C(\overline{U})$ is the unique classical solution of the boundary value problem

$$\partial_t u + \mathcal{L}\, u = 0, \quad \text{on } U \text{ and } u|_{\partial_P U} = v|_{\partial_P U}. \tag{4.8}$$

For a proof of this standard result one can consult the proof of Theorem 2.7.7 in [20]. In particular, we will use below that $\partial_x v$ and $\partial_t v$ are continuous in $\mathcal{C}$ (away from the boundary).

The free boundary b is the zero-level set of w. Since $\partial_x w \geq 0$ (see (4.6)), w is continuous and (b) in Lemma 4.4 holds, we can find $\delta > 0$ sufficiently small so that the equation $w(t, x) = \delta$ has a solution $x = b_\delta(t) > -\infty$ for all $t \in \mathcal{I}$. Further, by assumption (ii) we have $\partial_x w(t, x) > 0$ in $[\mathcal{I} \times (-\infty, r)] \cap \mathcal{C}$. Therefore the δ-level set of w is locally given by a function $b_\delta \in C(\mathcal{I})$. Moreover $b_\delta(t) > b(t)$ for all $t \in \mathcal{I}$ (this trivially holds if $b(t) = -\infty$ for some $t \in \mathcal{I}$).

On $\mathcal{I}$ the family $(b_\delta)_{\delta>0}$ decreases as $\delta \to 0$ so that its limit b_0 exists (possibly equal to $-\infty$). The mapping $t \mapsto b_0(t)$, is upper semi-continuous (on the extended real line), as decreasing limit of continuous functions, and $b_0(t) \geq b(t)$. Since $w(t, b_\delta(t)) = \delta$ it is clear that taking limits as $\delta \to 0$ we get $w(t, b_0(t)) = 0$ and therefore $b_0(t) \leq b(t)$ so that we conclude

$$\lim_{\delta\to 0} b_\delta(t) = b(t) \quad \text{for } t \in \mathcal{I}. \tag{4.9}$$

Since v is continuously differentiable in $\mathcal{C}$ and for all $t \in \mathcal{I}$ it holds $(t, b_\delta(t)) \in \mathcal{C}$ with $\partial_x w(t, b_\delta(t)) > 0$, then the implicit function theorem gives

$$b'_\delta(t) = -\frac{\partial_t w(t, b_\delta(t))}{\partial_x w(t, b_\delta(t))}, \qquad t \in \mathcal{I}. \tag{4.10}$$

Now we aim at showing that for arbitrary $\varepsilon > 0$, letting $\mathcal{I}_\varepsilon := (t_1 + \varepsilon, t_2 - \varepsilon)$ we have

$$b_\delta \text{ is bounded from below on } \mathcal{I}_\varepsilon, \text{ uniformly in } \delta \text{ and} \tag{4.11}$$

$$\text{there exists } K_\varepsilon > 0 \text{ such that } |b'_\delta(t)| \le K_\varepsilon \text{ on } \mathcal{I}_\varepsilon \text{ uniformly in } \delta. \tag{4.12}$$

If (4.11)–(4.12) hold, then by Ascoli-Arzelà's theorem we can extract a sequence $(b_{\delta_j})_{j\ge1}$ such that $b_{\delta_j} \to g$ uniformly on $\mathcal{I}_\varepsilon$ as $j \to \infty$, where g is Lipschitz continuous with constant K_ε. Uniqueness of the limit implies $g = b$ on $\mathcal{I}_\varepsilon$ and therefore b is (bounded) Lipschitz on $\mathcal{I}_\varepsilon$ with constant K_ε.

The rest of the proof will be devoted to verifying (4.11)–(4.12) or equivalently to finding a uniform bound for b_δ on $\mathcal{I}_\varepsilon$ and a uniform bound for

$$|b'_\delta(t)| = \frac{|\partial_t w(t, b_\delta(t))|}{\partial_x w(t, b_\delta(t))} \qquad \text{for } t \in \mathcal{I}_\varepsilon.$$

The proof is divided in steps.

Step 1. (Upper bound for $b(t)$). Due to (i) we have $b(t) < r$ for $t \in \overline{\mathcal{I}}$ for r as in (ii).

Step 2. (Lower bound for $b(t)$ and $b'(t)$). For any $(t, x) \in [0, T] \times \mathbb{R}$ and any stopping time $\tau \in [0, T - t]$ we have

$$\partial_t f(t, x) = \mathsf{E}\left[\partial_t f(t + \tau, X^x_\tau) - \int_0^\tau \partial_t m(t + s, X^x_s) ds\right]. \tag{4.13}$$

Take $(t, x) \in [0, T) \times \mathbb{R}$ and $\tau_* = \tau_*(t, x)$. Using (3.15), condition (D), (4.6) and $\mu'(x) \ge -\mu$ we easily obtain

$$\begin{aligned}
\partial_t w(t, x) \le& \mathsf{E}\Big[\int_0^{\tau_*} \partial_t (h + m)(t + s, X^x_s) ds\Big] + c' \,\mathsf{P}(\tau_* = T - t) \\
\le& c\left((1 + (T - t)^{-1})\mathsf{E}[\tau_*] + \mathsf{E}\Big[\int_0^{\tau_*} \partial_x (h + m)(t + s, X^x_s) ds\Big]\right) \\
\le& c\left((1 + (T - t)^{-1})\mathsf{E}[\tau_*] + e^{\mu(T-t)} \partial_x w(t, x)\right)
\end{aligned} \tag{4.14}$$

where c and c' are constants and in the second line we have also used Markov inequality as in (3.19). Using (4.14) in (4.10) gives us a lower bound for b'_δ of the form

$$b'_\delta(t) \ge -C\left(1 + \frac{\varphi(t, b_\delta(t))}{\partial_x w(t, b_\delta(t))}\right) \qquad \text{for } t \in \mathcal{I} \tag{4.15}$$

where $\varphi(t, x) := \mathsf{E}[\tau_*(t, x)] = \mathsf{E}_{t,x}[\tau_* - t]$. Next we want to find a bound for $(\varphi/\partial_x w)(t, b_\delta(t))$.

Recalling (4.6) we denote $\hat{w}$ the function

$$\hat{w}(t, x) = \mathsf{E}_{t,x}\left[\int_t^{\tau_*} e^{\int_t^s \mu'(X_u) du} \partial_x (h + m)(s, X_s) ds\right] \tag{4.16}$$

so that $\hat{w}(t, x) \le \partial_x w(t, x)$. It is important to notice that

$$\hat{w}_r := \inf_{t \in \mathcal{I}} \hat{w}(t, r) > 0 \tag{4.17}$$

since the segment $[t_1, t_2] \times \{r\}$ is strictly above the optimal boundary and $\partial_x(h + m) > 0$ in a neighborhood of the segment (see also Lemma 4.6).

Fix $t \in \mathcal{I}_{\varepsilon/2} = (t_1 + \varepsilon/2, t_2 - \varepsilon/2)$ and, to simplify notation, set $x_\delta := b_\delta(t)$ and $\tau_\delta := \tau_*(t, x_\delta)$. Let

$$\tau_r := \inf\{s \ge t : (s, X_s) \notin (t_1, t_2) \times (-\infty, r)\} \qquad \mathsf{P}_{t,x_\delta}\text{-a.s.} \tag{4.18}$$

and notice that $\tau_r < T$, P_{t,x_δ}-a.s. since $\overline{\mathcal{I}} \subset [0,T)$. Tower property of conditional expectation and (4.16) give

$$\begin{aligned}
&\hat{w}(t,x_\delta)\\
=&\mathsf{E}_{t,x_\delta}\Big[\int_t^{\tau_r\wedge\tau_\delta} e^{\int_t^s \mu'(X_u)du}\partial_x(h+m)(s,X_s)ds + \mathbb{1}_{\{\tau_r<\tau_\delta\}}\int_{\tau_r}^{\tau_\delta} e^{\int_t^s \mu'(X_u)du}\partial_x(h+m)(s,X_s)ds\Big]\\
=&\mathsf{E}_{t,x_\delta}\Big[\int_t^{\tau_r\wedge\tau_\delta} e^{\int_t^s \mu'(X_u)du}\partial_x(h+m)(s,X_s)ds\\
&\quad + \mathbb{1}_{\{\tau_r<\tau_\delta\}}e^{\int_t^{\tau_r} \mu'(X_u)du}\mathsf{E}_{t,x_\delta}\Big(\int_{\tau_r}^{\tau_r+\tau_\delta\circ\theta_{\tau_r}} e^{\int_t^s \mu'(X_u)du}\partial_x(h+m)(s,X_s)ds\Big|\mathcal{F}_{\tau_r}\Big)\Big],
\end{aligned}$$

where we recall that $\theta_\cdot$ is the shift operator on the canonical space. Strong Markov property gives

$$\begin{aligned}
&\hat{w}(t,x_\delta)\\
=&\mathsf{E}_{t,x_\delta}\Big[\int_t^{\tau_r\wedge\tau_\delta} e^{\int_t^s \mu'(X_u)du}\partial_x(h+m)(s,X_s)ds\\
&\quad + \mathbb{1}_{\{\tau_r<\tau_\delta\}}e^{\int_t^{\tau_r} \mu'(X_u)du}\mathsf{E}_{\tau_r,X_{\tau_r}}\Big(\int_{\tau_r}^{\tau_\delta} e^{\int_{\tau_r}^s \mu'(X_u)du}\partial_x(h+m)(s,X_s)ds\Big)\Big] \qquad (4.19)\\
=&\mathsf{E}_{t,x_\delta}\Big[\int_t^{\tau_r\wedge\tau_\delta} e^{\int_t^s \mu'(X_u)du}\partial_x(h+m)(s,X_s)ds + \mathbb{1}_{\{\tau_r<\tau_\delta\}}e^{\int_t^{\tau_r} \mu'(X_u)du}\hat{w}(\tau_r,X_{\tau_r})\Big].
\end{aligned}$$

For $\hat{w}$ we therefore have the following lower bound

$$\hat{w}(t,x_\delta) \geq C'\left(\mathsf{E}_{t,x_\delta}\left[\alpha_r(\tau_\delta\wedge\tau_r - t) + \mathbb{1}_{\{\tau_r<\tau_\delta\}}\hat{w}(\tau_r,X_{\tau_r})\right]\right) \qquad (4.20)$$

with $C' = e^{-\mu T}$. The same argument applied to φ gives

$$\varphi(t,x_\delta) = \mathsf{E}_{t,x_\delta}\left[(\tau_\delta\wedge\tau_r - t) + \mathbb{1}_{\{\tau_r<\tau_\delta\}}\varphi(\tau_r,X_{\tau_r})\right]. \qquad (4.21)$$

Now from (4.21) and (4.20), and by noticing that $0 \leq \varphi(t,x) \leq T$ we obtain

$$\begin{aligned}
0 \leq& \frac{\varphi(t,x_\delta)}{\partial_x w(t,x_\delta)} \leq \frac{\varphi(t,x_\delta)}{\hat{w}(t,x_\delta)}\\
\leq& \frac{1}{C'\alpha_r} + \frac{\mathsf{E}_{t,x_\delta}\left[\mathbb{1}_{\{\tau_r<\tau_\delta\}}\varphi(\tau_r,X_{\tau_r})\right]}{C'\left(\mathsf{E}_{t,x_\delta}\left[\alpha_r(\tau_\delta\wedge\tau_r - t) + \mathbb{1}_{\{\tau_r<\tau_\delta\}}\hat{w}(\tau_r,X_{\tau_r})\right]\right)}\\
\leq& C_r\left(1 + \frac{\mathsf{P}_{t,x_\delta}(\tau_r<\tau_\delta,\tau_r<t_2) + \mathsf{P}_{t,x_\delta}(\tau_r<\tau_\delta,\tau_r=t_2)}{\mathsf{E}_{t,x_\delta}\left[\alpha_r(\tau_\delta\wedge\tau_r - t) + \mathbb{1}_{\{\tau_r<\tau_\delta\}}\hat{w}(\tau_r,X_{\tau_r})\right]}\right), \qquad (4.22)
\end{aligned}$$

where $C_r := (\alpha_r^{-1}\vee T)/C'$. Since $t \in \mathcal{I}_{\varepsilon/2} = (t_1+\varepsilon/2, t_2-\varepsilon/2)$ we have

$$\begin{aligned}
&\frac{\mathsf{P}_{t,x_\delta}(\tau_r<\tau_\delta,\tau_r<t_2)}{\mathsf{E}_{t,x_\delta}\left[\alpha_r(\tau_\delta\wedge\tau_r - t) + \mathbb{1}_{\{\tau_r<\tau_\delta\}}\hat{w}(\tau_r,X_{\tau_r})\right]}\\
&\leq \frac{\mathsf{P}_{t,x_\delta}(\tau_r<\tau_\delta,\tau_r<t_2)}{\mathsf{E}_{t,x_\delta}\left[\mathbb{1}_{\{\tau_r<\tau_\delta\}\cap\{\tau_r<t_2\}}\hat{w}(\tau_r,r)\right]} \leq \frac{\mathsf{P}_{t,x_\delta}(\tau_r<\tau_\delta,\tau_r<t_2)}{\hat{w}_r\mathsf{P}_{t,x_\delta}(\tau_r<\tau_\delta,\tau_r<t_2)} \leq \frac{1}{\hat{w}_r} \qquad (4.23)
\end{aligned}$$

by using that $\mathbb{1}_{\{\tau_r<t_2\}}X_{\tau_r} = r$ and recalling (4.17). Similarly, for $t \in \mathcal{I}_{\varepsilon/2}$, we have

$$\begin{aligned}
&\frac{\mathsf{P}_{t,x_\delta}(\tau_r<\tau_\delta,\tau_r=t_2)}{\mathsf{E}_{t,x_\delta}\left[\alpha_r(\tau_\delta\wedge\tau_r - t) + \mathbb{1}_{\{\tau_r<\tau_\delta\}}\hat{w}(\tau_r,X_{\tau_r})\right]}\\
&\leq \frac{\mathsf{P}_{t,x_\delta}(\tau_r<\tau_\delta,\tau_r=t_2)}{\mathsf{E}_{t,x_\delta}\left[\mathbb{1}_{\{\tau_r<\tau_\delta\}\cap\{\tau_r=t_2\}}\alpha_r(\tau_\delta\wedge\tau_r - t)\right]} \leq \frac{1}{\alpha_r(t_2-t)} \leq \frac{2}{\alpha_r\varepsilon}. \qquad (4.24)
\end{aligned}$$

Now plugging the last two estimates back into (4.22) and recalling (4.15) we finally conclude

$$b'_\delta(t) \geq -C\left(1 + C_r\left(1 + \hat{w}_r^{-1} + 2(\alpha_r\varepsilon)^{-1}\right)\right) =: -\beta_{\varepsilon,r} \quad \text{for all } t \in \overline{\mathcal{I}_{\varepsilon/2}}. \tag{4.25}$$

Thanks to (a) in Lemma 4.4 we can find $t_0 \in \mathcal{I}$, arbitrarily close to t_1 and such that $|b(t_0)| < \infty$. So with no loss of generality we assume $t_0 = t_1 + \varepsilon/2$ and $b(t_0) > -\infty$. Since the bound in (4.25) is uniform in δ, then (4.9) implies that $b(t) \geq b(t_0) - \beta_{\varepsilon,r}|t - t_0|$ for all $t \in \overline{\mathcal{I}_{\varepsilon/2}}$ and proves (4.11). Hence, there exits $r^\varepsilon < r$ such that $b(t) \in (r^\varepsilon, r)$ for all $t \in \overline{\mathcal{I}_{\varepsilon/2}}$. This fact will be used in the next step of the proof.

Step 3. (Upper bound for $b'(t)$). It remains to find an upper bound for b'_δ on $\overline{\mathcal{I}_\varepsilon}$ which is uniform in δ. For that it is convenient to denote

$$\widetilde{g}(x) := |h(T,x) + n(x)| + 2|\partial_t f(T,x)| \quad \text{for } x \in \mathbb{R}. \tag{4.26}$$

and recall that $\partial_t v(t,x) \geq \underline{v}(t,x)$ given in (3.2). Using again (4.13) we immediately find

$$\partial_t w(t,x) \geq \overline{w}(t,x) := \mathsf{E}\left[\int_0^{\tau_*} \partial_t(h+m)(t+s, X_s^x)ds - \mathbb{1}_{\{\tau_* = T-t\}}\widetilde{g}(X_{T-t}^x)\right] \tag{4.27}$$

for $(t,x) \in [0,T] \times \mathbb{R}$. Now, for $t \in \mathcal{I}_\varepsilon$ we set

$$\tau'_r := \inf\{s \geq t : (s, X_s) \notin (t_1 + \varepsilon, t_2 - \varepsilon/2) \times (-\infty, r)\} \qquad \mathsf{P}_{t,x_\delta}\text{-a.s.}, \tag{4.28}$$

and, using the strong Markov property as in step 2 above, we find

$$\overline{w}(t,x_\delta) = \mathsf{E}_{t,x_\delta}\left[\int_t^{\tau'_r \wedge \tau_\delta} \partial_t(h+m)(s,X_s)ds + \mathbb{1}_{\{\tau'_r < \tau_\delta\}}\overline{w}(\tau'_r, X_{\tau'_r})\right]. \tag{4.29}$$

Since $b(t) \in [r^\varepsilon, r]$ for $t \in \overline{\mathcal{I}_{\varepsilon/2}}$ then under P_{t,x_δ} we have $X_s \in [r^\varepsilon, r]$ for all $s \in (t, \tau'_r \wedge \tau_\delta)$ and therefore there exists $\nu_\varepsilon > 0$ such that $\partial_t(h+m)(s,X_s) \geq -\nu_\varepsilon$ for all $s \in (t, \tau'_r \wedge \tau_\delta)$. On the other hand, from the definition of $\overline{w}$ in (4.27) and properties of $\partial_t(h+m)$ and $\widetilde{g}$ (see Assumption 2.2) it also follows

$$\overline{w}_{r,\varepsilon} := \sup_{t \in \mathcal{I}} |\overline{w}(t,r)| + \sup_{x \in [r^\varepsilon, r]} |\overline{w}(t_2 - \varepsilon/2, x)| < \infty. \tag{4.30}$$

In conclusion, from (4.29) we have

$$\overline{w}(t,x_\delta) \geq -\nu_\varepsilon \mathsf{E}_{t,x_\delta}\left[(\tau_\delta \wedge \tau'_r - t)\right] - \overline{w}_{r,\varepsilon}\mathsf{P}_{t,x_\delta}(\tau'_r < \tau_\delta). \tag{4.31}$$

Using (4.19) with τ_r replaced by τ'_r we obtain (4.20) with τ_r replaced by τ'_r. Hence, recalling (4.10), for $t \in \mathcal{I}_\varepsilon$ we have

$$\begin{aligned} b'_\delta(t) \leq & \frac{\nu_\varepsilon \mathsf{E}_{t,x_\delta}\left[(\tau_\delta \wedge \tau'_r - t)\right] + \overline{w}_{r,\varepsilon}\mathsf{P}_{t,x_\delta}(\tau'_r < \tau_\delta)}{C'\left(\mathsf{E}_{t,x_\delta}\left[\alpha_r(\tau_\delta \wedge \tau'_r - t) + \mathbb{1}_{\{\tau'_r < \tau_\delta\}}\hat{w}(\tau'_r, X_{\tau'_r})\right]\right)} \\ \leq & \frac{1}{C'}\left(\frac{\nu_\varepsilon}{\alpha_r} + \frac{\overline{w}_{r,\varepsilon}\mathsf{P}_{t,x_\delta}(\tau'_r < \tau_\delta)}{\mathsf{E}_{t,x_\delta}\left[\alpha_r(\tau_\delta \wedge \tau'_r - t) + \mathbb{1}_{\{\tau'_r < \tau_\delta\}}\hat{w}(\tau'_r, X_{\tau'_r})\right]}\right). \end{aligned} \tag{4.32}$$

The last term on the right-hand side above may be estimated by using the same arguments as in (4.23) and (4.24), upon noticing that the argument in (4.24) carries over to this case since for $t \in \mathcal{I}_\varepsilon$ we have $\mathbb{1}_{\{\tau'_r < \tau_\delta\} \cap \{\tau'_r = t_2 - \varepsilon/2\}}(\tau_\delta \wedge \tau'_r - t) \geq \varepsilon/2$.

Therefore we conclude

$$b'_\delta(t) \leq C'_r\left(1 + 2(\alpha_r\varepsilon)^{-1} + \hat{w}_r^{-1}\right) \tag{4.33}$$

where $C'_r = [(\nu_\varepsilon/\alpha_r) \vee \overline{w}_{r,\varepsilon}]/C'$. Now (4.33) and (4.25) imply (4.12) and the proof is complete. □

Proof of Lemma 4.4. We set $w := v - f$ and use the notation ∂_x in place of ∂_1. By a simple application of Dynkin's formula we can write w as

$$w(t,x) = \mathsf{E}\left[\int_0^{\tau_*} (h+m)(t+s, X_s^x)ds + \mathbb{1}_{\{\tau_*=T-t\}}\big(g(X_{\tau_*}) - f(t+\tau_*, X_{\tau_*}^x)\big)\right]. \tag{4.34}$$

For future reference we notice that under (ii) of Theorem 4.3 we have $\gamma \in C^1(\overline{\mathcal{I}})$ by the implicit function theorem, since $(h+m)(t,\gamma(t)) = 0$. Therefore

$$\underline{\gamma} := \inf_{t\in\overline{\mathcal{I}}} \gamma(t) > -\infty.$$

Step 1. Here we prove (a) by contradiction. Assume that we can find $t_1' < t_2'$ in $\overline{\mathcal{I}}$ such that $\big((t_1', t_2') \times \mathbb{R}\big) \cap \mathcal{S} = \varnothing$. Fix $t \in (t_1', t_2')$, take $x \le \underline{\gamma}$ and define

$$\rho_\gamma(t,x) := \inf\{s \ge 0 \,:\, X_s^x \ge \underline{\gamma}\} \wedge (t_2' - t).$$

By assumption we have $\tau_*(t,x) \ge (t_2' - t)$, hence $\tau_*(t,x) \ge \rho_\gamma(t,x)$, P-a.s. Recall that $(h+m)(s,y) \le 0$ for $y \le \gamma(s)$, $s \in [0,T]$ and notice that

$$(h+m)(s,y) = (h+m)(s,\underline{\gamma}) - \int_y^{\underline{\gamma}} \partial_x(h+m)(s,z)dz \le -\alpha_r(\underline{\gamma} - y), \tag{4.35}$$

thanks to (ii) in Theorem 4.3, for $s \in \mathcal{I}$ and $y \le \underline{\gamma}$. Denote $\rho_\gamma = \rho_\gamma(t,x)$. Similarly to, e.g., (4.19), we can use the tower property of conditional expectation and strong Markov property in (4.34), along with (4.35). This gives

$$\begin{aligned} w(t,x) =& \mathsf{E}\left[\int_0^{\rho_\gamma} (h+m)(t+s, X_s^x)ds + w(t+\rho_\gamma, X_{\rho_\gamma}^x)\right] \\ \le& -\alpha_r \mathsf{E}\left[\int_0^{\rho_\gamma} \big(\underline{\gamma} - X_s^x\big)ds\right] + w_\gamma, \end{aligned} \tag{4.36}$$

where $w_\gamma := \sup_{t\le s\le t_2'} w(s,\underline{\gamma}) < +\infty$, since w is continuous, and we used that $w(t+\rho_\gamma, X_{\rho_\gamma}^x) \le w_\gamma$ since $\partial_x w \ge 0$ by (4.6). Finally, letting $x \downarrow -\infty$ we reach a contradiction because $\rho_\gamma(t,x) \uparrow (t_2' - t)$ and the first term in the last expression of (4.36) goes to $-\infty$.

Step 2. Here we prove (b). Pick any $t \in \mathcal{I}$. From the previous step we know that we can find $t' \in \mathcal{I}$ with $t' > t$ such that $|b(t')| < +\infty$. In particular, this implies that $\{t'\} \times (-\infty, b(t')] \subseteq \mathcal{S}$. Now, for $x \le \underline{\gamma} \wedge b(t')$ let

$$\rho_\gamma'(t,x) := \inf\{s \ge 0 \,:\, X_s^x \ge \underline{\gamma} \wedge b(t')\}$$

and notice that $\rho_\gamma'(t,x) \wedge \tau_*(t,x)$ is smaller than the first time the process $(t+s, X_s^x)$ leaves the the set $[t,t') \times (-\infty, \underline{\gamma} \wedge b(t'))$. Hence, $\rho_\gamma'(t,x) \wedge \tau_*(t,x) \le (t'-t)$, P-a.s.

Using the tower property of conditional expectation and strong Markov property in (4.34), we obtain

$$\begin{aligned} w(t,x) =& \mathsf{E}\left[\int_0^{\rho_\gamma' \wedge \tau_*} (h+m)(t+s, X_s^x)ds + w(t+\rho_\gamma' \wedge \tau_*, X_{\rho_\gamma' \wedge \tau_*}^x)\right] \\ \le& \mathsf{E}\left[\mathbb{1}_{\{\rho_\gamma' < \tau_*\}} w(t+\rho_\gamma', X_{\rho_\gamma'}^x)\right], \end{aligned}$$

where we have also used that $(h+m)(s,y) \le 0$ for $y \le \gamma(s)$, and $\mathbb{1}_{\{\tau_* \le \rho_\gamma'\}} w(t+\tau_*, X_{\tau_*}^x) = 0$, P-a.s. Next we notice that

$$\{\rho_\gamma' < \tau_*\} = \{\rho_\gamma' < \tau_*, \rho_\gamma' < t' - t\}$$

because $\{\rho'_\gamma < \tau_*, \rho'_\gamma \geq t'-t\} = \varnothing$. Moreover, we recall that $w(t+\rho'_\gamma, X^x_{\rho'_\gamma}) \leq \sup_{t\leq s\leq t'} w(s,\underline{\gamma}) =: w'_\gamma < +\infty$ as $\partial_x w \geq 0$, by (4.6). Then we obtain

$$w(t,x) \leq w'_\gamma \mathsf{P}\big(\rho'_\gamma(t,x) < t'-t\big) = w'_\gamma \mathsf{P}\big(\sup_{0\leq s\leq t'-t} X^x_s > \underline{\gamma} \wedge b(t')\big).$$

Letting $x \to -\infty$ gives us (b), as claimed. □

Remark 4.5. *One can use local Lipschitz continuity of $t \mapsto b(t)$ and the law of iterated logarithm to show that τ_* is actually equal to the first time X goes strictly below the boundary, i.e., for all $(t,x) \in [0,T) \times \mathbb{R}$, it holds P-a.s.*

$$\inf\{s \geq 0 \,:\, X^x_s \leq b(t+s)\} \wedge (T-t) = \inf\{s \geq 0 \,:\, X^x_s < b(t+s)\} \wedge (T-t).$$

This is an important fact that can be used to prove that $(t,x) \mapsto \tau_*(t,x)$ is continuous P-a.s., and it is zero at all boundary points, hence implying $v \in C^{1,1}([0,T) \times \mathbb{R})$ (see for example [12, Sec. 5 and 6]).

4.2 Lipschitz boundary for $d \geq 2$

Lipschitz regularity for optimal boundaries when $d \geq 2$ requires slightly stronger assumptions on the functions f, g, h and μ which, however, are in line with those originally used in [36]. We give the result under two different sets of assumptions, namely conditions (F) and (G) below. The main difference between the two is that in (F) we do not need a positive lower bound for the quantity $\partial_1(h+m)$ but we need to compensate by imposing stronger bounds on the remaining quantities. Under conditions (G) we have a uniform lower bound on $\partial_1(h+m)$ so that other inequalities may be relaxed.

One should compare (G) to the assumptions in [36] and notice that we are in a similar setting (see also Remark 4.13 below for further details). We stress here that our theorems below *do not require uniform ellipticity* of the operator $\sigma\sigma^\top$ and therefore could not be obtained by PDE methods employed in [36]. We illustrate an application of our results and methodology in Example 2 of Section 5, which addresses the case of a degenerate diffusion.

The main idea of the proofs below is again to use the implicit function theorem as in the case of $d = 1$ (Theorem 4.3). However here we cannot rely upon continuity of $\nabla_x v$ and $\partial_t v$ due to the lack of uniform ellipticity of the parabolic operator $\partial_t + \mathcal{L}$ and therefore the arguments from Theorem 4.3 do not carry over. To overcome this additional difficulty we provide some notation and some technical lemmas, and throughout the section we use

(E) Regularity $\partial_t m$ and $\partial_k m$. Let $f \in C^{2,3}([0,T] \times \mathbb{R}^d)$ so that $m \in C^1([0,T] \times \mathbb{R}^d)$. For any compact $K \subseteq [0,T] \times \mathbb{R}^d$ we have

$$\sup_{(t,x)\in K} \mathsf{E}\left[\int_0^{T-t} |\partial_k m(t+s, X^x_s)|^2 + |\partial_t m(t+s, X^x_s)|^2 ds\right] < +\infty.$$

From now on we denote $w := v - f$. An application of Dynkin formula to $f(t, X_t)$ gives

$$w(t,x) = \sup_{0\leq\tau\leq T-t} \mathsf{E}\left[\int_0^\tau (h+m)(t+s, X^x_s)ds + \mathbb{1}_{\{\tau=T-t\}}\left(g(X^x_{T-t}) - f(T, X^x_{T-t})\right)\right]. \tag{4.37}$$

For $k = 1, \dots d$ we introduce the functions

$$\begin{aligned} w^\circ_k(t,x) :=& \mathsf{E}\Big[\int_0^{\tau_*} \langle \nabla_x (h+m)(t+s, X^x_s), \partial_k X^x_s\rangle ds \\ &+ \mathbb{1}_{\{\tau_*=T-t\}} \langle \nabla_x \left(g(X^x_{T-t}) - f(T, X^x_{T-t})\right), \partial_k X^x_{T-t}\rangle\Big] \end{aligned} \tag{4.38}$$

and in particular under condition (B) we notice that

$$w_1^\circ(t,x) = \mathsf{E}\Big[\int_0^{\tau_*} \partial_1(h+m)(t+s,X_s^x)\partial_1 X_s^{x,1} ds + \mathbb{1}_{\{\tau_*=T-t\}}\partial_1\left(g(X_{T-t}^x) - f(T,X_{T-t}^x)\right)\partial_1 X_{T-t}^{x,1}\Big]. \tag{4.39}$$

Thanks to Theorem 3.1 we have $\partial_1 w = w_1^\circ$ almost everywhere and therefore under condition (A_1), for a.e. $(t,x)\in[0,T]\times\mathbb{R}^d$, we also have

$$\partial_1 w(t,x) \geq \mathsf{E}\Big[\int_0^{\tau_*} \partial_1(h+m)(t+s,X_s^x)\partial_1 X_s^{x,1} ds\Big]. \tag{4.40}$$

In analogy with (3.1) and (3.2) we also introduce

$$\overline{w}(t,x) := \mathsf{E}\left[\int_0^{\tau_*} \partial_t(h+m)(t+s,X_s^x)ds - \mathbb{1}_{\{\tau_*=T-t\}}\left((h+\partial_t f)(T,X_{T-t}^x) + n(X_{T-t}^x)\right)\right] \tag{4.41}$$

$$\underline{w}(t,x) := \mathsf{E}\left[\int_0^{\tau_*} \partial_t(h+m)(t+s,X_s^x)ds - \mathbb{1}_{\{\tau_*=T-t\}}\left|(h+\partial_t f)(T,X_{T-t}^x) + n(X_{T-t}^x)\right|\right]. \tag{4.42}$$

The next technical lemma will be used in the proofs of Proposition 4.10 below.

Lemma 4.6. *Let $\mathcal{O}$ be a bounded open set in $[0,T]\times\mathbb{R}^d$ and $K\subset\mathcal{O}\cap\mathcal{C}$ a compact. Then if (A_2) and (B) hold, and $\partial_1(h+m)$ is continuous and strictly positive on $\mathcal{O}$, we have*

$$\inf_{(t,x)\in K} w_1^\circ(t,x) > 0 \tag{4.43}$$

Proof. Letting $\tau_0 := \inf\{s\geq t : (s,X_s)\notin\mathcal{O}\cap\mathcal{C}\}$, $\mathsf{P}_{t,x}$-a.s., we easily obtain from (4.39), (A_2) and (B) that

$$w_1^\circ(t,x) \geq c_{\mathcal{O}}\mathsf{E}_{t,x}\left[\tau_0 - t\right] \tag{4.44}$$

for some $c_{\mathcal{O}}>0$ only depending on $\mathcal{O}$. Choose now an arbitrary function $\psi\in C^2([0,T]\times\mathbb{R}^d)$ such that $\psi(\cdot)\geq 1$ on the complement of $\mathcal{O}\cap\mathcal{C}$ (denoted $(\mathcal{O}\cap\mathcal{C})^c$) and $\psi(\cdot)\leq\frac{1}{2}$ on K. With no loss of generality $\mathcal{L}\psi(\cdot)\leq c'_{\mathcal{O}}$ on $\mathcal{O}$ for some $c'_{\mathcal{O}}>0$ since $\psi\in C^2(\mathcal{O})$ and $\mathcal{O}$ is bounded. Then by an application of Dynkin formula, for any $(t,x)\in K$ we obtain

$$\begin{aligned}\mathsf{E}_{t,x}\left[\tau_0 - t\right] &= \mathsf{E}_{t,x}\left[\int_t^{\tau_0} dt\right] \geq \frac{1}{c'_{\mathcal{O}}}\mathsf{E}\left[\int_t^{\tau_0}\mathcal{L}\psi(s,X_s)dt\right]\\ &= \frac{1}{c'_{\mathcal{O}}}\left(\mathsf{E}_{t,x}\left[\psi(\tau_0,X_{\tau_0})\right] - \psi(t,x)\right) \geq \frac{1}{2c'_{\mathcal{O}}}\end{aligned} \tag{4.45}$$

where the final inequality uses that $\mathsf{P}_{t,x}\left[(\tau_0,X_{\tau_0})\in(\mathcal{O}\cap\mathcal{C})^c\right]=1$ so that $\psi(\tau_0,X_{\tau_0})\geq 1$, while $\psi(t,x)\leq\frac{1}{2}$ on K. Plugging (4.45) in (4.44) concludes the proof. □

From now on we let

$$c : [0,T]\times\mathbb{R}^{d-1}\to\mathbb{R}$$

be a given arbitrary function. The next lemma is an application of the chain rule and its proof is omitted.

Lemma 4.7. *Assume that c is differentiable at $(t,x_2,\dots x_d)$ and denote*

$$c := (c(t,x_2,\dots x_d),x_2,\dots x_d).$$

Then under condition (B) we have P*-a.s.*

$$\partial_t(X_s^{c,1}) = \partial_t c(t, x_2, \dots x_d)\partial_1 X_s^{x,1}|_{x=c} \tag{4.46}$$

$$\partial_k(X_s^{c,1}) = \partial_k c(t, x_2, \dots x_d)\partial_1 X_s^{x,1}|_{x=c} + \partial_k X_s^{x,1}|_{x=c}, \tag{4.47}$$

$$\partial_k(X_s^{c,j}) = \partial_k X_s^{x,j}|_{x=c}, \quad \text{and} \quad \partial_t(X_s^{c,j}) = 0 \tag{4.48}$$

for all $s \in [0, T]$ *and all* $k, j \geq 2$.

Notice that (4.48) holds because the j-th component of X^x, with $j \geq 2$, is not affected by changes in the initial point $x_1 = c(t, x_2, \dots x_d)$ of the first component $X^{x,1}$, due to condition (B). With a slight abuse of notation we are often going to use $\partial_k X^{c,j} = \partial_k X^{x,j}|_{x=c}$ when no confusion shall arise. For future reference we also give a straightforward corollary.

Corollary 4.8. *Let* $G \in C^{1,1}([0, T] \times \mathbb{R}^d)$ *and* $F \in C^1(\mathbb{R}^d)$*, then under the assumptions of Lemma 4.7 and with the same notation we have* P*-a.s., for all* $s \in [0, T - t]$*,*

$$\begin{aligned}\frac{\partial}{\partial t}G(t+s, X_s^c) =& \partial_t G(t+s, X_s^c) + \partial_1 G(t+s, X_s^c)\partial_1 X^{x,1}|_{x=c}\partial_t c(t, x_2, \dots x_d),\\ \frac{\partial}{\partial x_k}F(X_s^c) =& \langle \nabla_x F(X_s^c), \partial_k X_s^x|_{x=c}\rangle + \partial_1 F(X_s^c)\partial_1 X_s^{x,1}|_{x=c}\partial_k c(t, x_2, \dots x_d).\end{aligned}$$

The next lemma provides useful bounds which will be then employed to prove the main theorems below.

Lemma 4.9. *Let conditions* (A_1)*,* (A_2)*, (B) and (E) hold. Assume that* c *is differentiable at* $(t, x_2, \dots x_d)$ *and denote*

$$\begin{aligned}c &:= (c(t, x_2, \dots x_d), x_2, \dots x_d),\\ c_{t,\varepsilon}^+ &:= (c(t+\varepsilon, x_2, \dots x_d), x_2, \dots x_d),\\ c_{t,\varepsilon}^- &:= (c(t-\varepsilon, x_2, \dots x_d), x_2, \dots x_d),\\ c_{k,\varepsilon}^+ &:= (c(t, x_2, \dots x_k + \varepsilon, \dots x_d), x_2, \dots x_k + \varepsilon, \dots x_d),\\ c_{k,\varepsilon}^- &:= (c(t, x_2, \dots x_k - \varepsilon, \dots x_d), x_2, \dots x_k - \varepsilon, \dots x_d),\end{aligned}$$

for $k = 2, \dots d$*. Then for any* $k \geq 2$ *we have*

$$\begin{aligned}&\limsup_{\varepsilon\to 0}\frac{w(t,c) - w(t, c_{k,\varepsilon}^-)}{\varepsilon} \\ &\leq w_1^\circ(t,c)\partial_k c(t, x_2, \dots x_d) + w_k^\circ(t,c) \leq \liminf_{\varepsilon\to 0}\frac{w(t, c_{k,\varepsilon}^+) - w(t,c)}{\varepsilon}.\end{aligned} \tag{4.49}$$

Moreover we also have

$$\limsup_{\varepsilon\to 0}\frac{w(t,c) - w(t-\varepsilon, c_{t,\varepsilon}^-)}{\varepsilon} \leq w_1^\circ(t,c)\partial_t c(t, x_2, \dots x_d) + \overline{w}(t,c), \tag{4.50}$$

$$\liminf_{\varepsilon\to 0}\frac{w(t+\varepsilon, c_{t,\varepsilon}^+) - w(t,c)}{\varepsilon} \geq w_1^\circ(t,c)\partial_t c(t, x_2, \dots x_d) + \underline{w}(t,c). \tag{4.51}$$

Proof. The proof relies on Lemma 4.7 and on arguments similar to those used to prove Theorem 3.1. Denote $\tau = \tau_*(t,c)$ and let us consider the first inequality in (4.49). Notice that τ is optimal for $w(t,c)$ and sub-optimal for $w(t, c_{k,\varepsilon}^-)$ for $\varepsilon > 0$. Therefore we may estimate

$$\begin{aligned}w(t,c) - w(t, c_{k,\varepsilon}^-) \leq& \mathsf{E}\left[\int_0^\tau \left((h+m)(t+s, X_s^c) - (h+m)(t+s, X_s^{c_{k,\varepsilon}^-})\right) ds\right]\\ &+ \mathsf{E}\left[\mathbb{1}_{\{\tau = T-t\}}\left(g(X_{T-t}^c) - g(X_{T-t}^{c_{k,\varepsilon}^-}) + f(T, X_{T-t}^{c_{k,\varepsilon}^-}) - f(T, X_{T-t}^c)\right)\right].\end{aligned}$$

Dividing by ε, taking limits as $\varepsilon \to 0$ and using Corollary 4.8 and (2.14) we obtain the first inequality in (4.49) upon recalling the definitions of w_1° and w_k°.

A symmetric argument may be applied to obtain the second inequality in (4.49). This time we notice that τ is sub-optimal for $w(t, c_{k,\varepsilon}^+)$ for $\varepsilon > 0$ so that

$$\begin{aligned} w(t, c_{k,\varepsilon}^+) - w(t,c) \geq & \mathsf{E}\left[\int_0^\tau \left((h+m)(t+s, X_s^{c_{k,\varepsilon}^+}) - (h+m)(t+s, X_s^c)\right) ds\right] \\ & + \mathsf{E}\left[\mathbb{1}_{\{\tau = T-t\}}\left(g(X_{T-t}^{c_{k,\varepsilon}^+}) - g(X_{T-t}^c) + f(T, X_{T-t}^c) - f(T, X_{T-t}^{c_{k,\varepsilon}^+})\right)\right] \end{aligned}$$

holds. Dividing by ε and passing to the limit the claim follows thanks to Corollary 4.8.

For (4.50) we repeat arguments similar to those that led to (3.9) in step 2 of the proof of Theorem 3.1. These give

$$\begin{aligned} & w(t,c) - w(t-\varepsilon, c_{t,\varepsilon}^-) \\ & \quad \leq \mathsf{E}\left[\int_0^\tau \left((h+m)(t+s, X_s^c) - (h+m)(t-\varepsilon+s, X_s^{c_{t,\varepsilon}^-})\right) ds\right] \\ & + \mathsf{E}\left[\mathbb{1}_{\{\tau=T-t\}}\left(\hat{g}(X_{T-t}^c) - \hat{g}(X_{T-t}^{c_{t,\varepsilon}^-})\right)\right] \\ & - \mathsf{E}_{c_{t,\varepsilon}^-}\left[\mathbb{1}_{\{\tau=T-t\}}\mathsf{E}_{X_{T-t}}\left(\int_0^\varepsilon [(h+\partial_t f)(T-\varepsilon+s, X_s) + n(X_s)] ds\right)\right], \end{aligned}$$

where we have set $\hat{g}(x) := g(x) - f(T,x)$ to simplify the notation. Now dividing both sides of the above expression by ε, letting $\varepsilon \to 0$ and using Corollary 4.8 we get (4.50).

Similar arguments hold for (4.51) and following step 2 in the proof of Theorem 3.1, using that $\hat{g} \geq 0$, we have

$$\begin{aligned} & w(t+\varepsilon, c_{t,\varepsilon}^+) - w(t,c) \\ & \quad \geq \mathsf{E}\left[\int_0^{\tau\wedge(T-t-\varepsilon)} \left((h+m)(t+\varepsilon+s, X_s^{c_{t,\varepsilon}^+}) - (h+m)(t+s, X_s^c)\right) ds\right] \\ & \quad + \mathsf{E}\left[\mathbb{1}_{\{\tau > T-t-\varepsilon\}}\left(\hat{g}(X_{T-t-\varepsilon}^{c_{t,\varepsilon}^+}) - \hat{g}(X_{T-t-\varepsilon}^c)\right)\right] \\ & \quad - \mathsf{E}_c\left[\mathbb{1}_{\{\tau > T-t-\varepsilon\}}\mathsf{E}_{X_{T-t-\varepsilon}}\left(\int_0^\varepsilon |(h+\partial_t f)(T-\varepsilon+s, X_s) + n(X_s)| ds\right)\right]. \end{aligned}$$

Finally, dividing by ε and passing to the limit we get (4.51) thanks to Corollary 4.8. □

Under the assumptions of Proposition 4.1 we have $\partial_1 w \geq 0$ almost everywhere and therefore, for each $\delta > 0$, and for fixed $(t, x_2, \dots x_d)$, the equation $w(t, \cdot, x_2, \dots x_d) = \delta$ has at most a unique solution which we denote by $b_\delta(t, x_2, \dots x_d)$.

The next proposition states that b_δ is Lipschitz whenever finite and provides an important representation of its gradient at the points of differentiability. Below we use $\overline{U}$ for the closure of a set U.

Proposition 4.10. *Assume conditions (A_1), (A_2), (B), (C) and (E). Fix $(\hat{t}, \hat{x}_2, \dots \hat{x}_d) \in (0,T) \times \mathbb{R}^{d-1}$ and assume that there exists an open bounded neighbourhood U of $(\hat{t}, \hat{x}_2, \dots \hat{x}_d)$ and numbers $-\infty < \underline{b} < \overline{b} < +\infty$ such that*

$$\underline{b} < b(t, x_2, \dots x_d) < \overline{b} \quad \text{for } (t, x_2, \dots x_d) \in \overline{U}, \tag{4.52}$$

$$\partial_1(h+m)(t, x_1, x_2, \dots x_d) > 0 \quad \text{for } (t, x_2, \dots x_d) \in \overline{U} \text{ and } x_1 \in [\underline{b}, \overline{b}]. \tag{4.53}$$

Then there exists $\delta_U > 0$ such that b_δ is Lipschitz in U for all $\delta \in (0, \delta_U]$. Moreover, for all $k \geq 2$, and for a.e. $(t, y) \in U$ we have

$$\partial_k b_\delta(t, y) = -\frac{w_k^\circ(t, b_\delta(t, y), y)}{w_1^\circ(t, b_\delta(t, y), y)}, \tag{4.54}$$

$$-\frac{\overline{w}(t, b_\delta(t, y), y)}{w_1^\circ(t, b_\delta(t, y), y)} \leq \partial_t b_\delta(t, y) \leq -\frac{\underline{w}(t, b_\delta(t, y), y)}{w_1^\circ(t, b_\delta(t, y), y)}. \tag{4.55}$$

Proof. From now on we denote

$$U_b := \{(t, x) \,:\, x_1 \in (\underline{b}, \overline{b}),\ (t, x_2, \dots x_d) \in U\}.$$

Since $\overline{U}_b$ is compact and w is continuous, then there exists $\delta_U > 0$ sufficiently small and such that b_δ is bounded on U for all $\delta \leq \delta_U$, due to (4.52). With no loss of generality we then assume that $b_\delta(t, y) \in (\underline{b}, \overline{b})$ for $(t, y) \in U$ and all $\delta \leq \delta_U$. Next we show Lipschitz regularity of b_δ.

For all $(t, y) \in U$ the map $x_1 \mapsto w(t, x_1, y)$ is Lipschitz (Theorem 3.1). Then it is differentiable with $\partial_1 w(t, x_1, y) > 0$, for a.e. x_1 such that $(t, x_1, y) \in \mathcal{C} \cap U_b$, by (4.40) and (4.53). It follows that for all $(t, y) \in U$ the mapping $x_1 \mapsto w(t, x_1, y)$ is strictly increasing on $(b(t, y), \overline{b})$ and therefore a version of the implicit function theorem (see [23]) implies that b_δ is continuous in U.

For $\varepsilon \in \mathbb{R}$ we denote $b_\delta^\varepsilon := b_\delta(\hat{t}, \hat{x}_2 + \varepsilon, \hat{x}_3, \dots \hat{x}_d)$ and $b_\delta^0 = b_\delta$. With no loss of generality we assume that $(\hat{t}, b_\delta^\varepsilon, \hat{x}_2 + \varepsilon, \hat{x}_3 \dots \hat{x}_d)$ and $(\hat{t}, b_\delta^0, \hat{x}_2, \hat{x}_3 \dots \hat{x}_d)$ lie in U_b. Since $b_\delta \in C(U)$ and we are interested in small ε, there is again no loss of generality in assuming that

$$\zeta \mapsto (\hat{t}, b_\delta^\varepsilon, \hat{x}_2 + \zeta, \hat{x}_3, \dots \hat{x}_d),\ \ \eta \mapsto (\hat{t}, \eta, \hat{x}_2, \dots \hat{x}_d)$$

lie in a compact $K \subset U_b \cap \mathcal{C}$, for $\zeta \in (0, \varepsilon)$ and $\eta \in (b_\delta \wedge b_\delta^\varepsilon, b_\delta \vee b_\delta^\varepsilon)$.

Using that w is Lipschitz in U_b (see Theorem 3.1), for $\varepsilon \in \mathbb{R}$ we have

$$\begin{aligned}
0 =& w(\hat{t}, b_\delta^\varepsilon, \hat{x}_2 + \varepsilon, \hat{x}_3, \dots \hat{x}_d) - w(\hat{t}, b_\delta, \hat{x}_2, \dots \hat{x}_d) \\
=& w(\hat{t}, b_\delta^\varepsilon, \hat{x}_2 + \varepsilon, \dots \hat{x}_d) - w(\hat{t}, b_\delta^\varepsilon, \hat{x}_2, \dots \hat{x}_d) + w(\hat{t}, b_\delta^\varepsilon, \hat{x}_2, \dots \hat{x}_d) - w(\hat{t}, b_\delta, \hat{x}_2, \dots \hat{x}_d) \\
=& \int_0^\varepsilon \partial_2 w(\hat{t}, b_\delta^\varepsilon, \hat{x}_2 + \zeta, \hat{x}_3, \dots \hat{x}_d) d\zeta + \int_{b_\delta}^{b_\delta^\varepsilon} \partial_1 w(\hat{t}, \zeta, \hat{x}_2, \dots \hat{x}_d) d\zeta.
\end{aligned} \tag{4.56}$$

Then, using Lemma 4.6 with $\mathcal{O} = U_b$, we have $\partial_1 w \geq c_{K,\delta} > 0$ on K for a suitable constant $c_{K,\delta}$ depending on K and δ. From the last expression in (4.56) we get

$$\left| \int_0^\varepsilon \partial_2 w(\hat{t}, b_\delta^\varepsilon, \hat{x}_2 + \zeta, \hat{x}_3, \dots \hat{x}_d) d\zeta \right| = \left| \int_{b_\delta}^{b_\delta^\varepsilon} \partial_1 w(\hat{t}, \zeta, \hat{x}_2, \dots \hat{x}_d) d\zeta \right| \geq c_{K,\delta} |b_\delta^\varepsilon - b_\delta|$$

and using that $|\partial_2 w| \leq c_U'$ a.e. on U_b for a suitable constant c_U' (see Theorem 3.1), we conclude

$$|b_\delta^\varepsilon - b_\delta| \leq c_U' / c_{K,\delta} \cdot |\varepsilon|. \tag{4.57}$$

The same argument may be repeated for the remaining variables x_k, $k \geq 3$, and for t. Then, for any $(t', x') := (t', x_2', \dots x_d')$ and $(t, x) := (t, x_2, \dots x_d)$ belonging to a small ball in U, we have

$$|b_\delta(t', x') - b_\delta(t, x)| \leq c_\delta''(|t - t'| + \|x - x'\|_{d-1}) \tag{4.58}$$

for a suitable constant c_δ'' which depends on the small ball as well. This proves that b_δ is locally Lipschitz in U, hence it is differentiable almost everywhere therein.

Next we obtain probabilistic bounds for the gradient of b_δ. For $\varepsilon > 0$ we adopt the notation of Lemma 4.9. To simplify the exposition we set $c := (b_\delta(t, x_2, \dots x_d), x_2, \dots x_d)$, so that $c_{k,\varepsilon}^\pm$ and

$c^{\pm}_{t,\varepsilon}$ have the same meaning as in Lemma 4.9 but with $b_\delta(\cdot)$ instead of the function $c(\cdot)$. Recall that $b_\delta \in (\underline{b}, \overline{b})$ on U. Then, for all $k \geq 2$

$$\delta = w(t,c) = w(t, c^-_{k,\varepsilon}) = w(t, c^+_{k,\varepsilon}) = w(t-\varepsilon, c^-_{t,\varepsilon}) = w(t+\varepsilon, c^+_{t,\varepsilon}). \tag{4.59}$$

Hence by (4.49)–(4.51) we obtain that if b_δ is differentiable at $(t, x_2, \dots x_d)$, then

$$\partial_k b_\delta(t, x_2, \dots x_d) = -\frac{w^\circ_k(t, b_\delta(t, x_2, \dots), x_2, \dots x_d)}{w^\circ_1(t, b_\delta(t, x_2, \dots), x_2, \dots x_d)} \tag{4.60}$$

$$-\frac{\overline{w}(t, b_\delta(t, x_2, \dots), x_2, \dots x_d)}{w^\circ_1(t, b_\delta(t, x_2, \dots), x_2, \dots x_d)} \leq \partial_t b_\delta(t, x_2, \dots x_d) \leq -\frac{\underline{w}(t, b_\delta(t, x_2, \dots), x_2, \dots x_d)}{w^\circ_1(t, b_\delta(t, x_2, \dots), x_2, \dots x_d)}. \tag{4.61}$$

Since b_δ is differentiable almost everywhere in U then (4.54) and (4.55) follow. □

Using the bounds obtained for $\partial_t b_\delta$ and $\nabla_x b_\delta$ we can now prove the main theorems of this section. In what follows T may be infinite unless stated otherwise. The first theorem uses the next condition.

(F) Bounds II. For $(t,x) \in [0,T] \times \mathbb{R}^d$ there exists $c > 0$ such that

$$\sum_{j=1}^d |\partial_j (h+m)(t,x)| + |\partial_t (h+m)(t,x)| \leq c\, \partial_1 (h+m)(t,x), \tag{4.62}$$

and if $T < +\infty$ then also

$$|h(T,x) + n(x)| + 2|\partial_t f(T,x)| + \sum_{j=1}^d |\partial_j (g(x) - f(T,x))| \leq c\, \partial_1 (g(x) - f(T,x)). \tag{4.63}$$

Now we can state the theorem and give its proof.

Theorem 4.11 (Statement under (F)). *Assume that $d \geq 2$ and conditions (A_1), (A_2), (B), (C), (E) and (F) hold. Assume also that the bound $\sum_{j=1}^d \|\nabla_x \mu_j(x)\|_d \leq c$ holds for all $x \in \mathbb{R}^d$ and for some $c > 0$.*

If there exists $(\hat{t}, \hat{x}_2, \dots \hat{x}_d)$ and an open bounded neighbourhood U of the point such that (4.52) *and* (4.53) *hold then b is Lipschitz on U.*

Proof. We provide a full proof only for $T < +\infty$ but the same arguments hold for $T = +\infty$ up to simple minor changes.

The main idea of this proof is to show that

$$\|\partial_t b_\delta\|_{U,\infty} + \sum_{k=2}^d \|\partial_k b_\delta\|_{U,\infty} \leq c \tag{4.64}$$

for a uniform $c > 0$. Here $\|\cdot\|_{U,\infty}$ is the usual $L^\infty(U)$ norm and we are going to use the expressions for w°_k, $\overline{w}$ and $\underline{w}$ (see (4.38)–(4.42)) to find bounds in (4.54) and (4.55) (notice that the latter hold in the a.e. sense).

Recalling (3.16), the fact that $\partial_1 \mu_j(x) = 0$ for $j > 1$ due to (B), and the bounds in (F) it is not hard to verify that (4.38) gives

$$|w^\circ_k(t,x)| \leq \beta_0 w^\circ_1(t,x), \qquad (t,x) \in [0,T] \times \mathbb{R}^d \tag{4.65}$$

for all $k = 2, \dots d$, and a suitable $\beta_0 > 0$ which is independent of t, x and k. Similarly, using the bounds (F) in (4.41)–(4.42) we can find $\beta_1 > 0$ such that

$$\max\{|\underline{w}|, |\overline{w}|\}(t,x) \leq \beta_1 w_1^\circ(t,x), \qquad (t,x) \in [0,T] \times \mathbb{R}^d. \tag{4.66}$$

Now we argue as in the proof of Theorem 4.3 and since $|b| < +\infty$ on U we can find $\delta_U > 0$ sufficiently small and such that $w(t, \cdot, x_2, \dots x_d) = \delta$ has a solution $x_1 = b_\delta(t, x_2, \dots x_d)$, which is finite in U for all $\delta \leq \delta_U$. Then b_δ is Lipschitz in U by Proposition 4.10. Moreover from (4.54), (4.55) and (4.66) we obtain, for a.e. $(t,y) \in U$

$$|\partial_t b_\delta(t,y)| \leq \frac{\max\{|\underline{w}|, |\overline{w}|\}(t, b_\delta(t,y), y)}{w_1^\circ(t, b_\delta(t,y), y)} \leq \beta_1 \tag{4.67}$$

and from (4.65)

$$|\partial_k b_\delta(t,y)| = \frac{|w_k^\circ(t, b_\delta(t,y), y)|}{w_1^\circ(t, b_\delta(t,y), y)} \leq \beta_0, \quad k = 2, \dots d. \tag{4.68}$$

As in the proof of Theorem 4.3 (see (4.9)) we have pointwise convergence $b_\delta \downarrow b$ as $\delta \to 0$ and therefore, by using again Ascoli-Arzelà's theorem, we conclude that $b_\delta \to b$ uniformly on U. Hence b is Lipschitz on U. □

For the next theorem we are going to use a different technical condition.

(G) Bounds III. There exists $c_1, c_2 > 0$ such that

$$\partial_1(h+m)(t,x) \geq c_1 \tag{4.69}$$

$$\sum_{j=1}^d |\partial_j(h+m)(t,x)| + |\partial_t(h+m)(t,x)| \leq c_2\,[1 + \partial_1(h+m)(t,x)] \tag{4.70}$$

for $(t,x) \in [0,T] \times \mathbb{R}^d$. Moreover, if $T < +\infty$, at least one of the two conditions below holds:

(a) for $(t,x) \in [0,T] \times \mathbb{R}^d$ it holds

$$\begin{aligned} &|h(T,x) + n(x)| + 2|\partial_t f(T,x)| \\ &\qquad + \sum_{j=1}^d |\partial_j(g(x) - f(T,x))| \leq c_2\,[1 + \partial_1(g(x) - f(T,x))]\,; \end{aligned} \tag{4.71}$$

(b) for some $p \geq 1$ and for all $(t,x) \in [0,T] \times \mathbb{R}^d$ it holds

$$\sum_{j=1}^d |\partial_j(g(x) - f(T,x))| \leq c_2\,[1 + \partial_1(g(x) - f(T,x))]\,, \tag{4.72}$$

$$|h(T,x) + n(x)| + |\partial_t f(t,x)| \leq c_2(1 + \|x\|_d^p). \tag{4.73}$$

Now we can state the theorem and provide its proof.

Theorem 4.12 (Statement under (G)). *Assume that $d \geq 2$ and conditions (A_1), (A_2), (B), (C), (E) and (G) hold. Let $\sum_{j=1}^d \|\nabla_x \mu_j(x)\|_d \leq c$ hold true for all $x \in \mathbb{R}^d$ and a given $c > 0$ and, if $T < +\infty$ and only (b) of (G) holds, assume also $\|\mu(x)\|_d \leq c$ for $x \in \mathbb{R}^d$.*

If there exists $(\hat{t}, \hat{x}_2, \dots \hat{x}_d)$ and an open bounded neighbourhood U of the point such that (4.52) holds, then b is Lipschitz on U.

Proof. We provide a full proof only for $T < +\infty$ but the same arguments hold for $T = +\infty$ up to simple minor changes.

The idea of the proof is the same as in the previous theorem. Recalling (3.16), the fact that $\partial_1 \mu_j(x) = 0$ for $j > 1$ due to (B), and the bounds in (G) it is not hard to verify that (4.38) gives

$$\begin{aligned} |w_k^\circ(t,x)| \leq & \beta_0 \left(\mathsf{E}_{t,x}(\tau_* - t) + \mathsf{P}_{t,x}(\tau_* = T) + w_1^\circ(t,x)\right) \\ \leq & \beta_0 (1 + (T-t)^{-1}) \mathsf{E}_{t,x}(\tau_* - t) + \beta_0 w_1^\circ(t,x), \qquad (t,x) \in [0,T] \times \mathbb{R}^d \end{aligned} \tag{4.74}$$

for all $k = 2, \dots d$, and a suitable $\beta_0 > 0$ which is independent of t, x and k. Notice that the second inequality is just an application of Markov inequality (see (3.19)).

For the bounds on $\underline{w}$ and $\overline{w}$ we treat separately the case in which condition (a) of (G) holds and (b) of (G) holds. Starting with the former and recalling (4.41)–(4.42) it is not hard to show that there exists a constant $\beta_1 > 0$ such that for all $(t,x) \in [0,T) \times \mathbb{R}^d$

$$\begin{aligned} & \max\{|\underline{w}|, |\overline{w}|\}(t,x) \\ & \leq \mathsf{E}_{t,x}\left[\int_t^{\tau_*} |\partial_t (h+m)(s, X_s)| ds + \mathbb{1}_{\{\tau_* = T\}} |(h + \partial_t f)(T, X_T) + n(X_T)|\right] \\ & \leq \beta_1 \left(\mathsf{E}_{t,x}(\tau_* - t) + \mathsf{P}_{t,x}(\tau_* = T) + w_1^\circ(t,x)\right) \\ & \leq \beta_1 (1 + (T-t)^{-1}) \mathsf{E}_{t,x}(\tau_* - t) + \beta_1 w_1^\circ(t,x). \end{aligned} \tag{4.75}$$

Under condition (b) instead we use Lemma 3.5 to get

$$\mathsf{E}_{t,x}\left[\mathbb{1}_{\{\tau_* = T\}} |(h + \partial_t f)(T, X_T) + n(X_T)|\right] \leq \beta_1' (1 + \|x\|_d^p)(T-t)^{-1} \mathsf{E}_{t,x}(\tau_* - t)$$

for some $\beta_1' > 0$, while the estimate for the running cost $\partial_t(h+m)$ is the same as in (4.75). Therefore we can find again $\beta_1 > 0$ such that, for $(t,x) \in [0,T) \times \mathbb{R}^d$

$$\begin{aligned} & \max\{|\underline{w}|, |\overline{w}|\}(t,x) \\ & \leq \beta_1 \left[1 + (1 + \|x\|_d^p)(T-t)^{-1}\right] \mathsf{E}_{t,x}(\tau_* - t) + \beta_1 w_1^\circ(t,x). \end{aligned} \tag{4.76}$$

We notice that thanks to (G) we also have

$$w_1^\circ(t,x) \geq c\, \mathsf{E}_{t,x}(\tau_* - t), \quad (t,x) \in [0,T] \times \mathbb{R}^d.$$

To show the Lipschitz property let b_δ be the δ-level set of w and let us find bounds for (4.54) and (4.55). In particular for a.e. $(t,y) \in [0,T) \times \mathbb{R}^{d-1}$ we estimate

$$\begin{aligned} |\partial_t b_\delta(t,y)| \leq & \frac{\max\{|\underline{w}|, |\overline{w}|\}(t, b_\delta(t,y), y)}{w_1^\circ(t, b_\delta(t,y), y)} \\ \leq & \beta_1 + \frac{\beta_1}{c}\left[1 + (1 + \|(b_\delta, y)\|_d^p)(T-t)^{-1}\right], \end{aligned} \tag{4.77}$$

with $(b_\delta, y) := (b_\delta(t,y), y)$ for simplicity of notation, and

$$|\partial_k b_\delta(t,y)| = \frac{|w_k^\circ(t, b_\delta(t,y), y)|}{w_1^\circ(t, b_\delta(t,y), y)} \leq \beta_0 + \frac{\beta_0}{c}(1 + (T-t)^{-1}) \tag{4.78}$$

for $k = 2, \dots d$. The rest of the proof follows by letting $\delta \downarrow 0$ and using Ascoli-Arzelà's theorem as in the proof of Theorem 4.3. □

Remark 4.13. *The last two theorems above work under weaker technical conditions on f, g and h than those imposed in [36]. It is worth drawing a precise parallel between the two contributions and it is important to notice that some inequalities are reversed just because [36] considers a*

problem of optimal stopping with minimisation of costs, and in which the stopping set lies above the continuation set.

As for the notation, [36] takes a state-independent obstacle, i.e. $f(t,x) = f(t)$*, and a different ordering of the space coordinates. Indeed our* $\partial_k(h+m)(t,x)$ *should be associated to* $h_{n,n+1-k}(t,x)$ *in [36] for* $k = 1, \dots n$*. Similarly our* $\partial_k(g(x) - f(T,x))$ *corresponds to* $g_{n,n+1-k}(x)$ *of [36]. The setting we adopted in Theorem 4.12, with (G) using the specifications in (b), is more general than the setting in [36]; in particular conditions in (2.2) of [36] imply our* (4.69)–(4.70) *and* (4.72)*. The polynomial growth* (4.73) *is the same as in [36] and the requirement* $g_n(x) \leq f(0)$ *of [36] corresponds to our* $g(x) \geq f(T,x)$*.*

Finally we notice that results in [36] are obtained for $\mu(x) \equiv 0$ *and* $\sigma = diag\, 1$ *in* (2.1)*.*

5 Some examples and further extensions

Here we illustrate a couple of applications of our results to problems studied in the literature on stochastic control [10], [13]. The Lipschitz regularity of the free boundary in such problems is new and was not discovered in [10] and [13]. In all the examples below it is not difficult to check that the standing assumptions (2.4), (2.5), (2.14) and Assumption 2.2 hold.

Example 1. Here we consider a problem of optimal stopping arising in connection with one of irreversible investment (see [10]), under a Cobb-Douglas type production function. The state space is $[0,T] \times \mathbb{R}$ and the optimisation problem reads (see (3.15) and Section 4 in [10])

$$v(t,x) = \sup_{0 \leq \tau \leq T-t} \mathsf{E}\left[-\int_0^\tau e^{-rs-(1-\alpha)X_s^x} ds - c_1 e^{-r\tau} \mathbb{1}_{\{\tau < T-t\}} - c_2 e^{-r\tau} \mathbb{1}_{\{\tau = T-t\}} \right] \tag{5.1}$$

where $r > 0$, $\alpha \in (0,1)$, $c_1 \geq c_2 \geq 0$ and

$$X_t^x = x + \mu t + \sigma B_t, \qquad x \in \mathbb{R}.$$

Notice that, due to discounting, in this example we must replace the infinitesimal generator $\mathcal{L}$ by $\mathcal{L} - r$, which corresponds to the diffusion X killed at the constant rate r. In this setting we have

$$\begin{aligned} &h(x) = -e^{-(1-\alpha)x}, \quad f(x) = -c_1, \quad g(x) = -c_2, \quad m(x) = rc_1, \quad n(x) = rc_2, \\ &\partial_t(h+m) = 0, \quad \partial_x(h+m)(x) = (1-\alpha)e^{-(1-\alpha)x}. \end{aligned}$$

Here we want to use Theorem 4.3 and we start by noticing that conditions (A_1), (A_2), (C) and (D) hold. Moreover, the curve γ (see (4.7)) is simply given by

$$\gamma(t) = \gamma = (1-\alpha)^{-1} \ln(1/rc_1),$$

so that (i) of Theorem 4.3 holds. As for (ii) it is immediate to check that for any $x_0 > \gamma$ one has $\partial_x(h+m) \geq (1-\alpha)e^{-(1-\alpha)x_0}$ for $x \leq x_0$. It only remains to check the requirements of Corollary 3.3. If $c_2 = c_1$ then (i) in the latter corollary holds and therefore we have

Proposition 5.1. *For* $c_1 = c_2$ *Theorem 4.3 is true for problem* (5.1)*.*

We notice that (ii) in Corollary 3.3 is too strong in this setting and can never be verified. Moreover, in [10] they consider $c_1 > 0$ and $c_2 = 0$, so that the assumptions of Corollary 3.3 do not hold. However, as already mentioned, the key point in our method is the probabilistic representation for the bounds of $\partial_t v$ and $\partial_x v$. In particular, the explicit nature of problem (5.1) allows us to refine (3.2) and this turns out to be sufficient to prove Lipschitz regularity of b. In what follows we achieve this task.

Proposition 5.2. *If $c_1 > 0$ and $c_2 = 0$ then b is Lipschitz on $[0, T)$.*

Proof. First notice that (3.3) and (3.1) give

$$\partial_x v(t,x) = (1-\alpha)\mathsf{E}\left[\int_0^{\tau_*} e^{-rs-(1-\alpha)X_s^x} ds\right] \tag{5.2}$$

and

$$\overline{v}(t,x) = e^{-r(T-t)}\mathsf{E}\left[\mathbb{1}_{\{\tau_*=T-t\}} e^{-(1-\alpha)X_{T-t}^x}\right]. \tag{5.3}$$

For the lower bound of $\partial_t v$ we follow the proof of Theorem 3.1 up to (3.11), which now reads

$$\begin{aligned} v(t+\varepsilon,x) - v(t,x) &\geq -\mathsf{E}\left[\mathbb{1}_{\{\tau>T-t-\varepsilon\}}\left(e^{-r\tau}v(t+\tau,X_\tau^x) + \int_{T-t-\varepsilon}^{\tau} e^{-rs}h(X_s^x)ds\right)\right] \\ &= -\mathsf{E}\left[\mathbb{1}_{\{\tau>T-t-\varepsilon\}} e^{-r(T-t-\varepsilon)} v(T-\varepsilon, X_{T-t-\varepsilon}^x)\right] \geq 0 \end{aligned}$$

since $v \leq 0$ by (5.1). Since $\partial_t v$ exists at all points of $\mathcal{C}$, the above gives $\partial_t v \geq 0$.

The latter is useful in estimating b_δ' in (4.10). In fact we immediately obtain (noticing that here $\partial_t v = \partial_t w$ and $\partial_x v = \partial_x w$)

$$0 \geq b_\delta'(t) \geq -\frac{\overline{v}(t, b_\delta(t))}{\partial_x v(t, b_\delta(t))}. \tag{5.4}$$

To estimate the right-hand side of the above we observe that

$$e^{-(1-\alpha)X_t^x} = \frac{d\widetilde{\mathsf{P}}}{d\mathsf{P}}\Big|_{\mathcal{F}_t} \vartheta(t,x)$$

with

$$\frac{d\widetilde{\mathsf{P}}}{d\mathsf{P}}\Big|_{\mathcal{F}_t} := e^{(\alpha-1)\sigma B_t - \frac{1}{2}(\alpha-1)^2\sigma^2 t} \quad \text{and} \quad \vartheta(t,x) := e^{(\alpha-1)(x+\mu t+\frac{1}{2}(\alpha-1)\sigma^2 t)}.$$

Using the new probability measure in (5.2) and (5.3) we get

$$0 \geq b_\delta'(t) \geq -\frac{\vartheta(T-t,x)\widetilde{\mathsf{P}}(\tau_* = T-t)}{(1-\alpha)\widetilde{\mathsf{E}}\left[\int_0^{\tau_*} e^{-rs}\vartheta(s,x)ds\right]} \geq -\frac{(T-t)^{-1}\vartheta(T-t,x)\widetilde{\mathsf{E}}[\tau_*]}{(1-\alpha)\underline{\vartheta}(t,x)\,\widetilde{\mathsf{E}}[\tau_*]} \tag{5.5}$$

by using Markov inequality in the numerator and setting $\underline{\vartheta}(t,x) := \inf_{0\leq s\leq T-t} e^{-rs}\theta(s,x)$. Hence we conclude

$$0 \geq b_\delta'(t) \geq -\frac{(T-t)^{-1}\vartheta(T-t,x)}{(1-\alpha)\underline{\vartheta}(t,x)}$$

□

Example 2. Now we consider a multi-dimensional problem which arises in connection with irreversible investment with stochastic prices and demand (see [13]). The state variables are $(t,x,y,z) \in [0,T] \times \mathbb{R}^3$ and we treat separately the case $T < +\infty$ and $T = +\infty$. The problem is degenerate because there is no dynamics in the z direction (however a dynamic with no Brownian part could be included without altering our analysis). This can also be seen as a family of problems depending on a parameter z. The *most interesting feature* is that we can prove Lipschitz continuity of the boundary also with respect to the parameter z.

The optimisation problem reads

$$v(t,x,y,z) = \sup_{0\le\tau\le T-t} \mathsf{E}\left[\int_0^\tau e^{-rs}(z - X_t^x)dt - e^{-r\tau}Y_\tau^y\right] \tag{5.6}$$

where $r > 0$. It is worth mentioning that the running cost above corresponds to taking $c(x,z) = \frac{1}{2}(x-z)^2$ in the control problem studied by [13]. As discussed in the latter paper, this choice is very natural in that context (see Remark 2.5 in [13]). Notice also that the form of the payoff prevents a dimensionality reduction, hence also the infinite horizon case $T = +\infty$ is truly 2-dimensional and parameter dependent.

A typical application of this model is for electricity generation in presence of renewable sources. Here X is associated to a stochastic demand net of generation from renewables, and Y to the stochastic spot price of electricity. The variable Z is related to the level of production capacity from conventional generation. In this model both demand and price can take negative values, which is consistent with real market observations. We consider three different cases to illustrate our methodology in full.

Example 2-(a). Let us start with a finite-horizon problem with simple dynamics. Let $T < +\infty$ and

$$X_t^x = x + \alpha t + \beta B_t \quad \text{and} \quad Y_t^y = y + \mu t + \sigma W_t \tag{5.7}$$

where $\alpha, \beta, \sigma, \mu$ are constants while B and W are (possibly correlated) Brownian motions.

Again we are in presence of a killing at a rate r and therefore we use $\mathcal{L} - r$ instead of $\mathcal{L}$. We have

$$\begin{aligned} &h(x,z) = z - x, \quad f(y) = g(y) = -y, \quad m(y) = n(y) = ry - \mu, \\ &\partial_t(h+m) = 0, \quad \partial_x(h+m)(x,y,z) = -1, \quad \partial_z(h+m)(x,y,z) = 1, \quad \partial_y(h+m)(x,y,z) = r. \end{aligned}$$

Considering $\partial_1 = \partial_y$, it is immediate to check that Proposition 4.1 holds and we have

$$\mathcal{S} = \{(t,x,y,z) \,:\, y \le b(t,x,z)\}.$$

Finiteness of the boundary was proved in [13] for the infinite horizon case, and therefore holds as well for $T < +\infty$. Moreover all assumptions of Theorem 4.12 hold under condition (b) of (G). Hence we have

Proposition 5.3. *For $T < +\infty$ and with the dynamics* (5.7)*, Theorem 4.12 is true for problem* (5.6)*.*

Example 2-(b). Let $T = +\infty$ and the dynamic of (X,Y) be

$$X_t^x = x + \alpha\int_0^t (\zeta - X_s^x)ds + \beta B_t \quad \text{and} \quad Y_t^y = y + \mu t + \sigma W_t \tag{5.8}$$

where $\alpha, \beta, \sigma, \mu, \zeta$ are constants while B and W are (possibly correlated) Brownian motions. In this setting we assume a mean reverting dynamic for the demand. For the finiteness of (5.6) and to guarantee (2.5) we pick $r > \alpha$. We have

$$\begin{aligned} &h(x,z) = z - x, \quad f(y) = g(y) = -y, \quad m(y) = n(y) = ry - \mu, \\ &\partial_t(h+m) = 0, \quad \partial_x(h+m)(x,y,z) = -1, \quad \partial_z(h+m)(x,y,z) = 1, \quad \partial_y(h+m)(x,y,z) = r. \end{aligned}$$

As in case (a) above, taking $\partial_1 = \partial_y$, we see immediately that Proposition 4.1 holds and we have

$$\mathcal{S} = \{(x,y,z) \,:\, y \le b(x,z)\}. \tag{5.9}$$

Now Theorems 4.11 and 4.12 hold, because both conditions (F) and (G) hold due to infinite horizon (hence the unbounded drift for X is admissible). So we have

Proposition 5.4. *For $T = +\infty$ and with the dynamics* (5.8)*, Theorems 4.11 and 4.12 are true for problem* (5.6)*.*

Example 2-(c). Here we want to consider non-negative prices and this will require to adapt our methods as we did in Example 1. Let $T = +\infty$ and the dynamic of (X, Y) be

$$X_t^x = x + \alpha \int_0^t (\zeta - X_s^x) ds + \beta B_t \quad \text{and} \quad Y_t^y = y + \mu \int_0^t Y_s^y ds + \sigma \int_0^t Y_s^y dW_s, \tag{5.10}$$

where $\alpha, \beta, \sigma, \mu, \zeta$ are constants while B and W are independent Brownian motions. We take $r > \alpha \vee \mu$ to guarantee finiteness of the value function and (2.5).

To fit in the framework of (2.1) we must consider the new state variable $\pi := \ln y$ so that a new process Π_t can be defined as

$$Y_t^y = \exp\{\pi + (\mu - \tfrac{\sigma^2}{2})t + \sigma W_t\} =: \exp\{\Pi_t^\pi\}$$

In the state variables (x, π, z) we have

$$\begin{aligned} &h(x,z) = z - x, \quad f(\pi) = g(\pi) = -e^\pi, \quad m(\pi) = n(\pi) = (r-\mu)e^\pi, \quad \partial_t(h+m) = 0, \\ &\partial_x(h+m)(x,\pi,z) = -1, \quad \partial_z(h+m)(x,\pi,z) = 1, \quad \partial_\pi(h+m)(x,\pi,z) = (r-\mu)e^\pi. \end{aligned}$$

Now the stopping set reads $\mathcal{S} = \{(x, \pi, z) : \pi \le b(x, z)\}$ and the boundary is finite. It was shown in [13] that $w := v + e^\pi$ is differentiable in x and π inside the continuation region. Moreover, with a slight abuse of notation, we also define $\partial_z w := w_3^\circ$ (see (4.38)).

Proposition 5.5. *For $T = +\infty$ and with the dynamics* (5.10) *we have b Lipschitz on $\mathbb{R}^2$.*

Proof. For $(x, \pi, z) \in \mathbb{R}^3$ let $\tau_* = \tau_*(x, \pi, z)$ for simplicity. From (3.3) and (4.38) we get

$$\partial_x w(x,\pi,z) = -\mathsf{E}\left[\int_0^{\tau_*} e^{-(r+\alpha)t} dt\right], \quad \partial_\pi w(x,\pi,z) = (r-\mu)\widetilde{\mathsf{E}}\left[\int_0^{\tau_*} e^{-(r-\mu)t} dt\right] \tag{5.11}$$

$$\partial_z w(x,\pi,z) = \mathsf{E}\left[\int_0^{\tau_*} e^{-rt} dt\right] \tag{5.12}$$

where the measure $\widetilde{\mathsf{P}}$ is defined by

$$\frac{d\widetilde{\mathsf{P}}}{d\mathsf{P}}\Big|_{\mathcal{F}_t} := e^{\sigma W_t - \frac{\sigma^2}{2}t}, \qquad t \ge 0. \tag{5.13}$$

The crucial observation now is that, while the dynamic of X is unaffected by the change of measure (due to the independence of B and W) the dynamic of Π under $\widetilde{\mathsf{P}}$ becomes

$$\Pi_t^\pi = \pi + (\mu + \tfrac{\sigma^2}{2})t + \sigma \widetilde{W}_t$$

where $\widetilde{W}_t = W_t - \sigma t$ is the new Brownian motion under $\widetilde{\mathsf{P}}$.

Now, if on $(\Omega, \mathcal{F}, \mathsf{P})$ we define

$$\widetilde{\Pi}_t^\pi = \pi + (\mu + \tfrac{\sigma^2}{2})t + \sigma W_s$$

and $\widetilde{\tau}_* := \inf\{t \ge 0 : \widetilde{\Pi}_t^\pi \le b(X_t^x, z)\}$, then we immediately see that

$$\mathsf{Law}(X^x, \Pi^\pi | \widetilde{\mathsf{P}}) = \mathsf{Law}(X^x, \widetilde{\Pi}^\pi | \mathsf{P}) \quad \text{and} \quad \mathsf{Law}(\tau_* | \widetilde{\mathsf{P}}) = \mathsf{Law}(\widetilde{\tau}_* | \mathsf{P}).$$

Moreover by comparison principles for SDEs we have that $\widetilde{\Pi}_t^\pi \geq \Pi_t^\pi$, P-a.s. for all $t \geq 0$ and therefore we get $\widetilde{\tau}_* \geq \tau_*$, P-a.s., and

$$\widetilde{\mathsf{E}}\left[\int_0^{\tau_*} e^{-(r-\mu)t}dt\right] = \mathsf{E}\left[\int_0^{\widetilde{\tau}_*} e^{-(r-\mu)t}dt\right] \geq \mathsf{E}\left[\int_0^{\tau_*} e^{-(r-\mu)t}dt\right] \tag{5.14}$$

Using (5.11)–(5.14) and setting $\tau_* = \tau_*(x, b_\delta(x,z), z)$ for simplicity, we find a uniform bound for ∇b_δ for any $\delta > 0$, that is

$$|\partial_z b_\delta(x,z)| = \left|\frac{\partial_z w(x, b_\delta(x,z), z)}{\partial_\pi w(x, b_\delta(x,z), z)}\right| \leq (r-\mu)^{-1}\frac{\mathsf{E}\left[\int_0^{\tau_*} e^{-rt}dt\right]}{\mathsf{E}\left[\int_0^{\tau_*} e^{-(r-\mu)t}dt\right]} \leq \frac{1}{r-\mu}$$

$$|\partial_x b_\delta(x,z)| = \left|\frac{\partial_x w(x, b_\delta(x,z), z)}{\partial_\pi w(x, b_\delta(x,z), z)}\right| \leq (r-\mu)^{-1}\frac{\mathsf{E}\left[\int_0^{\tau_*} e^{-(r+\alpha)t}dt\right]}{\mathsf{E}\left[\int_0^{\tau_*} e^{-(r-\mu)t}dt\right]} \leq \frac{1}{r-\mu}.$$

Thus taking $\delta \to 0$ we find that b is Lipschitz as claimed.

□

Acknowledgements:
This work was financially supported by Sapienza University of Rome, research project "*Modelli stocastici a tempo continuo per le scelte previdenziali*", grant no. C26A14HXBR. T. De Angelis was also partially supported by EPSRC grant EP/R021201/1. We thank G. Peskir for helpful comments about general theory of optimal stopping. Finally, we thank the Associate Editor and an anonymous Referee for pertinent comments that helped improve the exposition of the paper.

References

[1] Bayraktar, E., Xing, H. (2009). Analysis of the Optimal Exercise Boundary of American Options for Jump Diffusions. SIAM J. Math. Anal. **41** (2), pp. 825–860.

[2] Bensoussan, A., Lions, J.L. (1982). Applications of Variational Inequalities in Stochastic Control. North-Holland.

[3] Blanchet, A. (2006). On the Regularity of the Free-Boundary in Parabolic Obstacle Problems. Application to American Options. *Nonlinear Anal.* **65**, pp. 1362–1378.

[4] Blumenthal, R.M., Getoor, R.K. (1960). Markov processes and potential theory. Academic press, New York-London.

[5] Caffarelli, L.A., Salsa, S. (2005). A geometric approach to free boundary problems. Graduate Studies in Mathematics, Vol. 68, American Mathematical Society, Providence (RI).

[6] Cannon, J.R. (1984). The one-dimensional heat equation. Encyclopedia of Mathematics and its Applications, Vol. 23. Cambridge University Press.

[7] Cannon, J. R., Denson Hill, C. (1967). On the Infinite Differentiability of the Free Boundary in a Stefan problem. *J. Math. Anal. Appl.* **22**, pp. 385–397.

[8] Chen, X., Chadam, J. (2007). A Mathematical Analysis of The Optimal Exercise Boundary for American Put Options. *SIAM J. Math. Anal.* **38** (5), pp. 1613–1641.

[9] Chen, X., Chadam, J., Jiang, L.,Zheng, W. (2008). Convexity of the Exercise Boundary of the American Put Options on a Zero Dividend Asset. *Math. Finance* **18** (1), pp. 185–197.

[10] Chiarolla, M. B. *and* Haussmann, U. G. (2009). On a Stochastic, Irreversible Investment Problem. *SIAM J. Control Optim.* **48** (2), pp 438–462.

[11] De Angelis, T. (2015). A note on the continuity of free-boundaries in finite-horizon optimal stopping problems for one dimensional diffusions. *SIAM J. Control Optim.* **53** (1), pp. 167-184

[12] De Angelis, T., Ekström, E. (2017). The dividend problem with a finite horizon. *Ann. Appl. Probab.* **27** (6), pp. 3525-3546.

[13] De Angelis, T., Federico, S., Ferrari, G. (2014). On the optimal boundary of a three-dimensional singular stochastic control problem arising in irreversible investment. *Math. Oper. Res.* **42** (4), pp. 1135–1161.

[14] De Angelis, T., Stabile, G. (2017). On the free boundary of an annuity purchase. *Accepted in Finance Stoch.* **arXiv**:1707.09494.

[15] Ekström, E. (2004). Convexity of the Optimal Stopping Boundary for the American Put Option. *J. Math. Anal. Appl.* **299**, pp. 147–156.

[16] Friedman, A. (2008). Partial Differential Equations of Parabolic Type. Dover Publications, New York.

[17] Friedman, A. (1975). Parabolic Variational Inequalities in one Space Dimension and Smoothness of the Free-Boundary. *J. Funct. Anal.* **18**, pp. 151–176.

[18] Jacka, S.D. (1991). Optimal stopping and the American put. *Math. Finance* **1** (2), pp. 1–14.

[19] Johnson, P., Peskir, G. (2017). Quickest detection problems for Bessel processes. *Ann. Appl. Probab.* **27** (2), pp. 1003–1056.

[20] Karatzas, I., Shreve, S.E. (1998). Methods of Mathematical Finance. Applications of Mathematics 39, Springer-Verlag, New York.

[21] Kim, J. (1990). The analytic valuation of American options. *Rev. Financ. Stud.* **3** (4), pp. 547–572.

[22] Kotlow, D.B. (1973). A Free-Boundary Problem Connected with the Optimal Stopping Problem for Diffusion Processes. *Trans. Amer. Math. Soc.* **184**, pp. 457–478.

[23] Kumagai, S. (1980). An implicit function theorem: Comment. *J. Optim. Theory Appl.* **31**(2), pp. 285–288.

[24] Lamberton, D., Villeneuve, S. (2003). Critical price near maturity for an American option on a dividend-paying stock. *Ann. Appl. Probab.* **13** (2), pp. 800–815.

[25] Laurence, P., Salsa, S. (2009). Regularity of the Free Boundary of an American Option on Several Assets. *Comm. Pure Appl. Math.* **62**, pp. 969–994.

[26] McKean, H. P., Jr. (1965). Appendix: A Free Boundary problem for the Heat Equation arising from a problem of Mathematical Economics. *Ind. Management Rev.* **6**, pp. 32–39.

[27] Petrosyan, A., Shahgholian, H., Uraltseva, N. (2012). Regularity of free boundaries in obstacle-type problems. Graduate Studies in Mathematics, Vol. 136, American Mathematical Society, Providence (RI).

[28] Protter, P.E. (2005). Stochastic Integration and Differential Equations (Second Edition). Springer Berlin Heidelberg.

[29] Peskir, G. (2005). On the American option problem. *Math. Finance* **15** (1), pp. 169–181.

[30] Peskir, G. (2017). Continuity of the optimal stopping boundary for two-dimensional diffusions. *To appear in Ann. Appl. Probab.*

[31] Peskir, G., Shiryaev, A.N. (2006). Optimal Stopping and Free-Boundary Problems. Lectures in Mathematics, ETH Zürich, Birkhäuser.

[32] Schaeffer, D. G. (1976). A New Proof of the Infinite Differentiability of the Free Boundary in the Stefan Problem. *J. Differential Equations* **20**, pp. 266–269.

[33] Shiryaev, A.N. (1978). Optimal stopping rules. Springer, New York-Heidelberg.

[34] Shreve, S.E., Soner, H.M. (1991). A free boundary problem related to singular stochastic control. *Applied stochastic analysis (London, 1989)*, pp. 265–301.

[35] Soner, H.M., Shreve S.E. (1989). Regularity of the value function for a two-dimensional singular stochastic control problem. *SIAM J. Control Optim.* **27** (4), pp. 876–907.

[36] Soner, H.M., Shreve S.E. (1991). A free boundary problem related to singular stochastic control: the parabolic case. *Comm. Partial Differential Equations* **16** (2-3), pp. 373–424.

[37] Van Moerbeke, P. (1976). On optimal stopping and free-boundary problems. *Arch. Ration. Mech. Anal.* **60**, pp. 101–148.