Ann. Appl. Probab. Vol. 30, No. 3, 2020, (1007–1031)
Research Report No. 13, 2016, *Probab. Statist. Group Manchester* (29 pp)

Global C^1 Regularity of the Value Function in Optimal Stopping Problems

T. De Angelis & G. Peskir

We show that if either the process is strong Feller and the boundary point is probabilistically regular for the stopping set, or the process is strong Markov and the boundary point is probabilistically regular for the interior of the stopping set, then the boundary point is Green regular for the stopping set. Combining this implication with the existence of a continuously differentiable flow of the process we show that the value function is continuously differentiable at the optimal stopping boundary whenever the gain function is so. The derived fact holds both in the parabolic and elliptic case of the boundary value problem under the sole hypothesis of probabilistic regularity of the optimal stopping boundary, thus improving upon known analytic results in the PDE literature, and establishing the fact for the first time in the case of integro-differential equations. The method of proof is purely probabilistic and conceptually simple. Examples of application include the first known probabilistic proof of the fact that the time derivative of the value function in the American put problem is continuous across the optimal stopping boundary.

1. Introduction

A challenging question in boundary value problems is to establish *regularity* of the solution up to the boundary. By regularity we mean continuity, differentiability, and/or higher degrees of smoothness. The problem has a long and venerable history. Continuity results can be traced back to Poincaré [43] and the references therein. Differentiability results date back to Gevrey [22] for parabolic equations and Kellog [29] for elliptic equations (see also [30]). Extensions to more general parabolic and elliptic equations were made possible using the techniques developed by Schauder [48] (see [32] for further details). As a rule of thumb in the PDE literature it is known that (probabilistic) regularity of the boundary implies continuity of the solution up to the boundary, and smoothness (or Hölder continuity) of the boundary implies smoothness of the solution up to the boundary (see e.g. [19, Theorem 7, p. 64] for parabolic equations and [23, Lemma 6.18, p. 111] for elliptic equations). This common belief translates to free boundary problems for parabolic and elliptic equations as well (see e.g. [20, Lemma 4.5, p. 167] for a definite result of this kind dating back to Gevrey [22] as well as [4] and [5, Chapter 8] for related results in higher dimensions). The analytic method of variational inequalities removes the focus from the free boundary itself and derives a global continuity of the *space* derivative (for parabolic and elliptic equations of *diffusion* processes) when the obstacle function is *globally*

Mathematics Subject Classification 2010. Primary 60G40, 60J25, 60H30. Secondary 35R35, 35J15, 35K10.
Key words and phrases: Optimal stopping problem, strong Markov/Feller process, free boundary problem, regularity of the value function, regularity of a boundary point, regularity of a stochastic flow, smooth fit.

C^1 while establishing that the time derivative exists in a *weak* sense only (see [1, Corollary 1.3, p. 207] and [21, Theorem 3.2, p. 26; Theorem 8.2, p. 77; Theorem 8.4, p. 80]). The latter fact is not surprising since the time derivative can fail to exist in the absence of probabilistic regularity of the free boundary (see e.g. [40, Example 14]). A probabilistic approach in [36] returns to a probabilistic regularity of the free boundary by assuming moreover that the free boundary is twice continuously differentiable and thus making the assumption 'intractable' as the paper points out itself.

In this paper we develop a conceptually simple/direct probabilistic method which shows that the differentiability results for free boundary problems can be derived solely from a probabilistic regularity of the boundary i.e. with no need for its smoothness (or Hölder continuity) of any kind. This applies to (i) both the space derivative and the time derivative, (ii) more general strong Markov/Feller processes (not just diffusions), and (iii) both smooth and non-smooth obstacle functions. Free boundary problems (in analysis) are known to be equivalent to optimal stopping problems (in probability) and we derive the differentiability results in the context of optimal stopping problems which are also of interest in themselves. We do that by establishing a *continuous* smooth fit between the value function and the gain (obstacle) function at the optimal stopping (free) boundary that is traditionally derived using probabilistic methods in a directional sense only (see Section 2 for details).

In Section 2 we formulate the optimal stopping problem (2.1)/(2.2) and explain its background in terms of (i) strong Markov/Feller processes, (ii) boundary point regularity (probabilistic, Green, barrier, Dirichlet), (iii) stochastic flow regularity, and (iv) infinitesimal generator regularity (including continuous and smooth fit). In Section 3 we show that if either the process is strong Feller and the boundary point is probabilistically regular for the stopping set, or the process is strong Markov and the boundary point is probabilistically regular for the interior of the stopping set, then the boundary point is Green regular for the stopping set (in the sense that the expected waiting time for entering the stopping set vanishes as the initial point of the process approaches the boundary point from within the continuation set). Combining this implication with the existence of a continuously differentiable flow of the process we show in Sections 4 and 5 that the value function is continuously differentiable at the optimal stopping boundary whenever the gain function is so. Theorems 8 and 10 deal with the space derivative (in infinite and finite horizon respectively) and Theorems 13 and 15 deal with the time derivative (in infinite and finite horizon respectively). Examples 12 and 17 derive the analogous regularity results for the space derivative and the time derivative respectively, when the gain function is not smooth away from the optimal stopping boundary, using the local time of the process on the singular points at which the smoothness breaks down.

The advantage of the probabilistic method employed in the derived results is that the only hypothesis on the optimal stopping boundary used is its probabilistic regularity for the stopping set or its interior (which is implied by monotonicity of the optimal stopping boundary for instance). This level of generality is insufficient for the PDE methods as they require at least a Lipschitz (or Hölder) continuity of the optimal stopping boundary. The derived results hold both in the parabolic and elliptic case of the free boundary problem, thus improving upon known analytic results in the PDE literature, and establishing the fact for the first time in the case of integro-differential equations. Moreover, the 'lifting' method of Example 17 to our knowledge is applied for the first time in the literature. It enables one to 'lift' a Lipschitz continuity of the superharmonic/value function to its C^1 regularity at Green regular boundary points. Among other implications this yields the first known probabilistic proof of the fact that the time derivative of the value function in the American put problem is continuous across the optimal stopping boundary.

In parallel to producing a first draft of the present paper we have also applied/tested some parts of the method of proof in specific examples. This includes [10] for the time derivative in the Brownian motion case and [27] for the space derivative in the Bessel process case. For further/existing applications to (singular) stochastic control problems and optimal stopping games we refer to [11] and [12] respectively. Among intermediate references we note that the paper [2] studies continuity of the time derivative of solutions to parabolic free-boundary problems in one (spatial) dimension under the hypotheses that $G = 0$ on the stopping set (with $G > 0$ at the end of time) and $H < 0$ globally in the optimal stopping problem (2.2) below. These hypotheses are rarely satisfied in the mainstream examples of optimal stopping problems studied in the literature (including the American put problem where $G > 0$ on the stopping set and $H = 0$ globally) and the present paper fills this gap as well.

2. Problem formulation

In this section we introduce the setting of the problem and explain its background in terms of the general hypotheses imposed and sufficient conditions that imply them.

1. *Optimal stopping problem.* We consider the optimal stopping problem

$$V(x) = \sup_{\tau} \mathsf{E}_x \Big[e^{-\Lambda_\tau} G(X_\tau) + \int_0^\tau e^{-\Lambda_t} H(X_t)\, dt \Big] \tag{2.1}$$

for $x \in I\!R^d$ with $d \geq 1$ where $X = (X^1, \ldots, X^d)$ is a standard Markov process (in the sense of [3, p. 45]) taking values in $I\!R^d$. Thus X is strong Markov, right-continuous with left limits, and left-continuous over stopping times. The process X starts at x under the probability measure P_x for $x \in I\!R^d$ (or its measurable subset identified with $I\!R^d$ in the sequel for simplicity). The supremum in (2.1) is taken over all stopping times τ of X (i.e. stopping times with respect to the natural filtration of X), or equivalently, over all stopping times τ with respect to a (right-continuous) filtration $(\mathcal{F}_t)_{t \geq 0}$ that makes X a strong Markov process under P_x for $x \in I\!R^d$. All stopping times considered throughout are assumed to be finite valued unless otherwise stated (upon recalling that extensions to infinite valued stopping times are both standard and straightforward). We will also consider the optimal stopping problem (2.1) with finite horizon obtained by imposing an upper bound $T > 0$ on τ . In this case we also need to account for the length of the remaining time so that (2.1) extends as follows

$$V(t,x) = \sup_{0 \leq \tau \leq T-t} \mathsf{E}_x \Big[e^{-\Lambda_\tau} G(X_\tau) + \int_0^\tau e^{-\Lambda_s} H(X_s)\, ds \Big] \tag{2.2}$$

for $t \in [0,T]$ and $x \in I\!R^d$. Note that this includes the case when the functions G and H are time dependent which can be formally obtained by setting $X_t^1 = t$ for $t \geq 0$. The functional Λ in (2.1) and (2.2) is defined by

$$\Lambda_t = \int_0^t \lambda(X_s)\, ds \tag{2.3}$$

where λ is a continuous function with values in $[0, \infty)$. The real-valued functions G and H are also assumed to be continuous. Under these hypotheses it is known (cf. [41] and [49]) that the first entry time of X into the (finely) closed set D where V equals G (the stopping set) is optimal in (2.1)/(2.2) provided that $G(X)$ and $H(X)$ satisfy mild integrability conditions. This is true for example if $\lambda > 0$ and both G and H are bounded but this sufficient condition can be considerably strengthened (see [41] and [49] for details). The (finely) open set where V is strictly larger than G (the continuation set) will be denoted by C. The (optimal stopping) boundary between the sets C and D will be denoted by ∂C. We will make use of and distinguish between the *first entry time* of X into D defined by

$$\tau_D = \inf \{ t \geq 0 \mid X_t \in D \} \tag{2.4}$$

and the *first hitting time* of X to D defined by

$$\sigma_D = \inf \{ t > 0 \mid X_t \in D \} \tag{2.5}$$

where D can also be replaced by any other measurable subset of $I\!R^d$ and an upper bound applies to admissible t in (2.4) and (2.5) when the horizon is finite as in (2.2). When the standard regularity hypotheses recalled above are satisfied, or any other sufficient conditions implying that τ_D is optimal in (2.1)/(2.2), we will say that the problem (2.1)/(2.2) is *well posed*. This will be a standing premise for the rest of the paper. Any additional hypotheses will always be invoked explicitly in the statements of the results below when needed.

2. *Strong Feller processes.* Recall that the process X is *strong Feller* if

$$x \mapsto \mathsf{E}_x\big[F(X_t)\big] \;\text{ is continuous} \tag{2.6}$$

for every real-valued (bounded) *measurable* function F with $t > 0$ given and fixed. Recall also that X is *Feller* if (2.6) holds for every real-valued (bounded) *continuous* function F. Recall finally that Feller processes are strong Markov. Strong Feller processes were introduced and initially studied by Girsanov [24]. All one-dimensional diffusions X in the sense of Itô and McKean [26] are known to be strong Feller processes because the transition density p of X with respect to its speed measure m (in the sense that $\mathsf{P}_x(X_t \in dy) = p(t; x, y)\, m(dy)$) can be chosen to be jointly continuous in all three arguments (cf. [26, p. 149]). Unique weak solutions to (non-degenerate) SDEs driven by a Wiener process in $I\!R^d$ are known to be not only strong Markov but also strong Feller processes (see e.g. [47, p. 170]). A time-space process such as $((t, W_t))_{t \geq 0}$ where W is a standard Wiener process is *not* a strong Feller process. Not all Lévy processes are strong Feller either. Hawkes [25, Theorem 2.2] showed that a Lévy process X is strong Feller if and only if $\mathsf{P}_x(X_t \in dy) \ll \ell(dy)$ for every $t > 0$ and $x \in I\!R^d$ where ℓ denotes Lebesgue measure on $I\!R^d$. Strong Feller property is important in relation to boundary point regularity. We will now present basic facts in this direction.

3. *Boundary point regularity.* There are four closely related concepts of boundary point regularity that we will address in the sequel. Throughout we let $b(c, r)$ denote the open ball in the Euclidean topology of $I\!R^d$ with centre at c and radius $r > 0$. By $\bar{C}$ we denote the closure of C and by D° we denote the interior of D. Recall that a real-valued function v is superharmonic on a set $A \subseteq I\!R^d$ relative to X if $\mathsf{E}_x[v(X_\tau)] \leq v(x)$ for all $x \in A$ and all (bounded) stopping times $\tau \leq \tau_{A^c}$ of X. A boundary point $z \in \partial C$ is said to be:

(2.7) *Probabilistically regular* (PR) if we have $\mathsf{P}_z(\sigma_D = 0) = 1$;

(2.8) *Green regular* (GR) if we have $\lim_{C \ni x \to z \in \partial C} \mathsf{P}_x(\tau_D \geq \varepsilon) = 0$ for each $\varepsilon > 0$;

(2.9) *Barrier regular* (BR) if there exists a superharmonic function $v > 0$ on $b(z,r) \cap C$ relative to X for some $r > 0$ such that $\lim_{C \ni x \to z \in \partial C} v(x) = 0$;

(2.10) *Dirichlet regular* (DR) if $\lim_{C \ni x \to z \in \partial C} \mathsf{E}_x[F(X_{\tau_D})] = F(z)$ for each real-valued (bounded) measurable function F on $b(z,r) \cap \bar{C}$ with $r > 0$ that is continuous at z.

Regularity of $z \in \partial C$ in definitions (2.7)-(2.10) refers to the set D. If we replace D in (2.7)-(2.10) by any measurable subset A of $I\!R^d$ then we speak about regularity of $z \in \partial C$ for the set A. By Blumenthal's 0-1 law (cf. [3, p. 30]) we know that the probability in (2.7) can only be either zero or one. The super(harmonic) function v in (2.9) is referred to as a *barrier* itself. The main example of a barrier is $v(x) = \mathsf{E}_x(\tau_D)$ for $x \in b(z,r) \cap C$ with $r > 0$ when $\lim_{C \ni x \to z \in \partial C} v(x) = 0$ holds (where τ_D could be replaced by $\tau_D \wedge 1$ to make it bounded).

It is well known (cf. [16, pp. 32-40]) that if X is strong Feller then

$$\text{PR} \Longleftrightarrow \text{GR} \Longleftrightarrow \text{BR}. \tag{2.11}$$

Moreover, if X is strong Feller and uniformly continuous on compacts in the sense that

$$\lim_{t \downarrow 0} \sup_{x \in K} \mathsf{P}_x \Big(\sup_{s \in [0,t]} |X_s - x| > \varepsilon \Big) = 0 \tag{2.12}$$

for each compact set K in $I\!R^d$ and each $\varepsilon > 0$ then

$$\text{PR} \Longleftrightarrow \text{GR} \Longleftrightarrow \text{BR} \Longleftrightarrow \text{DR} \tag{2.13}$$

where $|\cdot|$ denotes the Euclidean norm in $I\!R^d$. We will see in the proofs below that our main focus will be on the *Green regularity*. When the process X fails to be strong Feller however, then the first equivalence in (2.11)/(2.13) can break down generally, and we will then require probabilistic regularity for D° instead of D to gain the Green regularity. Further details in this direction will be presented in the next section.

We will close this subsection with a few historical details aimed at clarifying definitions (2.7)-(2.10) above. Note that many papers cited below contain *sufficient conditions* for boundary point regularity that are directly relevant for the main results in Sections 4 and 5 below.

Definition (2.7) embodies what probabilists understand under regularity. Definition (2.10) embodies what analysts understand under regularity. The implication (2.7) $\Rightarrow$ (2.10) was first proved by Doob [14] for a Wiener process and was then extended by Girsanov [24] to other strong Feller processes. The converse implication (2.10) $\Rightarrow$ (2.7) for strong Feller processes was derived by Krylov [33]. Definition (2.8) embodies a "hybrid" condition representing a mixture of (2.7) and (2.10) that makes it suitable for applications as we will see below. Definition (2.9) is often used to derive various sufficient conditions for regularity. Poincaré [43] used barriers to derive a sphere condition. Zaremba [53] replaced sphere by a cone (cf. [28, pp. 247-250]). Wiener [52] derived a necessary and sufficient condition for regularity using the capacity of a set (Wiener's test). These papers deal with the Laplace equation (when X is a Wiener process) and extensions to more general elliptic equations are normally not difficult (probabilistically this can be seen through time changes and comparison arguments). The same phenomenon does not hold for the heat equation (when X is a time-space Wiener process) and more general parabolic equations (see e.g. [17, Theorem 8.1] for a simple example). Petrovsky [42] derived sufficient conditions for regularity in the heat equation by considering boundaries as functions of time (Kolmogorov-Petrovsky's test). Necessary and sufficient conditions for regularity in the heat equation were announced by Landis [35]. An analogue of Wiener's test for the heat equation was derived in the papers by Lanconelli [34] and Evans & Gariepy [18] (see pp. 295-296 in the latter paper for related results and historical comments). We refer to the paper by Watson [51] and the references therein for subsequent analytic results and further developments. Boundary point regularity and continuity of the solution to the Dirichlet problem for standard Markov processes have been studied by Dembinski [13] using purely probabilistic methods (see also the references therein for further probabilistic papers on this topic).

4. *Stochastic flow regularity.* Stochastic processes whose sample paths are indexed by their initial points are referred to as *stochastic flows.* Motivated by needs in the proofs below we will assume that the standard Markov process X can be realised as a stochastic flow $(X_t^x)_{t \geq 0, x \in I\!R^d}$ on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$ in the sense that $\mathrm{Law}(X \,|\, \mathsf{P}_x) = \mathrm{Law}(X^x \,|\, \mathsf{P})$ where we set $X^x = (X_t^x)_{t \geq 0}$ for $x \in I\!R^d$.

Examples of stochastic flows include a standard Wiener flow $W = (W_t^x)_{t \geq 0, x \in I\!R}$ where $W_t^x = x + W_t$ (which extends to all Lévy processes analogously), an exponential Wiener flow $S = (S_t^x)_{t \geq 0, x \in I\!R}$ where $S_t^x = x \exp(\sigma W_t + (\mu - \sigma^2/2)t)$ for $\sigma > 0$ and $\mu \in I\!R$, and a reflecting Wiener flow $R = (R_t^x)_{t \geq 0, x \in I\!R}$ where $R_t^x = x \vee \sup_{0 \leq s \leq t} W_s - W_t$. Very often an explicit construction of the stochastic flow is not possible and then one usually aims to establish its existence satisfying some/further regularity properties. Among these we will need to consider *continuous*, *differentiable*, and *continuously differentiable* stochastic flows. For us in this paper it will mean that there exists a (universal) set $N \in \mathcal{F}$ satisfying $\mathsf{P}(N) = 0$ such that the mapping $x \mapsto X_t^x(\omega)$ is continuous, differentiable, or continuously differentiable on $I\!R^d$ for every $\omega \in \Omega \setminus N$ and each $t \geq 0$ given and fixed. The first spatial derivative of the stochastic-flow coordinate X^j with respect to x_i will be denoted by $\partial_i X_t^{j,x} := \partial_{x_i} X_t^{j,x} = \partial X_t^{j,x} / \partial x_i$ for $t \geq 0$ and $x = (x_1, \ldots, x_d) \in I\!R^d$ with $1 \leq i, j \leq d$. (The same notation will also be applied to deterministic functions throughout including their time derivatives whenever convenient.) Thus when the stochastic flow is continuously differentiable we know that $x \mapsto \partial_i X_t^{j,x}(\omega)$ is continuous on $I\!R^d$ for every $\omega \in \Omega \setminus N$ and each $t \geq 0$ where $P(N) = 0$ and $1 \leq i, j \leq d$. We will also assume that the (timewise) sample path regularity of X^j translates to the same sample path regularity of $\partial_i X^j$, i.e. if $t \mapsto X_t^{j,x}(\omega)$ is continuous or right-continuous with left limits, then $t \mapsto \partial_i X_t^{j,x}(\omega)$ is continuous or right-continuous with left limits for every $\omega \in \Omega \setminus N$ and each $x \in I\!R^d$ where $P(N) = 0$ and $1 \leq i, j \leq d$.

To obtain *sufficient conditions* for stochastic flow regularity, which are directly relevant for the main results in Sections 4 and 5 below, recall that a stochastic flow $X = (X_t^x)_{t \geq 0, x \in I\!R^d}$ may be viewed as a stochastic field $Z = (Z_z)_{z \in I\!R_+ \times I\!R^d}$, where we set $Z_z = X_t^x$ for $z = (t, x) \in I\!R_+ \times I\!R^d$, so that the results on sample path regularity of stochastic fields are applicable to stochastic flows. The earliest results of this kind for the existence of (Hölder) *continuous* modifications of stochastic processes (when the index set of a stochastic field is $I\!R_+$) were derived by Kolmogorov in 1934 (unpublished) and published subsequently by Slutsky [50] (see also [31, pp. 158-165] for extensions of these results to stochastic fields when the index set is $I\!R_+^n$ for $n \geq 1$). Sufficient conditions for the existence of *right-continuous* modifications of stochastic processes (with left limits) have been derived by Chentsov [6] and Cramér [8]. Sufficient conditions for the existence of *continuously differentiable* modifications of stochastic processes have been derived in the book by Cramér and Leadbetter [9, pp. 67-70]. All these conditions are of a Hölder-in-mean type involving either two-dimensional (for continuity) or three-dimensional (for right-continuity or differentiability) marginal laws of the process. Different sufficient conditions for the existence of continuously differentiable modifications of stochastic fields (indexed by $I\!R_+^n$ for $n \geq 1$) have been derived by Potthoff [44, Theorem 3.2] based on the ideas of Loève cited therein. These conditions require the existence of the first partial derivative of the original stochastic flow in the mean-square sense (thus again being of a Hölder-in-mean type however without specifying the admissible rate of convergence) combined with the existence of a continuous modification of the resulting partial derivative flow (which can be established at least formally using the extended Kolmogorov conditions for stochastic fields referred to above).

The preceding results give a variety of general sufficient conditions for the existence of a *regular* stochastic field and hence a *regular* stochastic flow as well. Entering into a more specific class of stochastic processes, it is well known that SDEs driven by semimartingales with differentiable coefficients having locally Lipschitz first partial derivatives generate continuously differentiable flows (cf. [45, Theorem 39, p. 305]). In particular, this is true for SDEs driven by a standard Wiener process or a more general Lévy process in $I\!R^d$. Each of these processes therefore satisfies the hypothesis on the existence of a continuously differentiable flow. To express the hypothesis in a compact form we will simply say that the process X can be realised as a continuously differentiable stochastic flow $(X_t^x)_{t\geq 0, x\in I\!R^d}$ in the space variable.

5. *Infinitesimal generator regularity.* We will assume in the sequel that the infinitesimal generator of X is given by

$$\begin{aligned}(2.14) \qquad I\!L_X F(x) = \frac{1}{2}\sum_{i,j=1}^d \sigma_{ij}(x)\frac{\partial^2 F}{\partial x_i \partial x_j}(x) + \sum_{i=1}^d \mu_i(x)\frac{\partial F}{\partial x_i}(x) - \lambda(x)F(x) \\ + \int_{I\!R^d\setminus\{0\}} \Big(F(y) - F(x) - \sum_{i=1}^d (y_i - x_i)\frac{\partial F}{\partial x_i}(x)\Big)\, \nu(x,dy)\end{aligned}$$

for any function $F : I\!R^d \to I\!R$ from its domain and $x \in I\!R^d$, where the matrix $(\sigma_{ij})_{i,j=1}^d$ with values in $I\!R^{d\times d}$ is symmetric and positive semi-definite (diffusion coefficient), the vector $(\mu_i)_{i=1}^d$ takes values in $I\!R^d$ (drift coefficient), λ takes values in $I\!R_+$ (killing coefficient), and $\nu(x,dy)$ is a non-negative measure on $I\!R^d\setminus\{0\}$ (the compensator of the measure of jumps of X). For more details we refer to [46, pp. 281-299] and [41, pp. 128-142]. The infinitesimal role of $I\!L_X$ is uniquely determined through its action on sufficiently regular (smooth) functions F that could also involve various boundary conditions (on curves or surfaces in $I\!R^d$) depending on the stochastic behaviour of the process X (on these curves or surfaces). It is well known that if F belongs to the domain of $I\!L_X$ then

$$(2.15) \qquad F(X_t) - F(X_0) - \int_0^t I\!L_X F(X_s)\, ds$$

is a (local) martingale. This is a single most useful consequence of the previous inclusion (if known) that we will need in the sequel. When X is a semimartingale then (2.15) with $I\!L_X$

from (2.14) can also be derived for sufficiently regular (smooth) functions F using stochastic calculus techniques (Itô's formula and its extensions). The importance of the infinitesimal generator (2.14) follows from the well-known fact that the optimal stopping problem (2.1) is equivalent to the free boundary problem

$$\mathbb{L}_X V = -H \text{ on } C \tag{2.16}$$

$$V = G \text{ on } \partial C \text{ (continuous fit)} \tag{2.17}$$

$$\frac{\partial V}{\partial x_i} = \frac{\partial G}{\partial x_i} \text{ on } \partial C \text{ for } 1 \leq i \leq d \text{ (smooth fit)} \tag{2.18}$$

where $\mathbb{L}_X G \leq -H$ on D, and the continuity condition (2.17) or (2.18) applies as a variational principle when the expectation in (2.1) with $\tau_{D'}$ in place of τ for $D' \neq D$ is discontinuous or has discontinuous first partial derivatives at $\partial C'$ as a function of the initial point $x \in \mathbb{R}^d$ respectively (for more details see [41, p. 49]). Continuity of the partial derivatives in (2.18) has been traditionally understood/derived in the directional sense as follows

$$\lim_{h \downarrow 0} \frac{\partial V}{\partial x_i}(x_1, \ldots, x_{i-1}, x_i \pm h, x_{i+1}, \ldots, x_d) = \lim_{h \downarrow 0} \frac{\partial G}{\partial x_i}(x_1, \ldots, x_{i-1}, x_i \mp h, x_{i+1}, \ldots, x_d) \tag{2.19}$$

upon assuming that $(x_1, \ldots, x_{i-1}, x_i \pm h, x_{i+1}, \ldots, x_d)$ belongs to C and $(x_1, \ldots, x_{i-1}, x_i \mp h, x_{i+1}, \ldots, x_d)$ belongs to D for $h > 0$ and $1 \leq i \leq d$. Our main aim in this paper is to derive the continuity of the partial derivatives in (2.18) globally at ∂C, i.e. we aim to show that if $x^n \in C$ converges to $x \in \partial C$ then $(\partial V/\partial x_i)(x^n)$ converges to $(\partial G/\partial x_i)(x)$ as $n \to \infty$ for $1 \leq i \leq d$. When combined with the interior regularity results for V on C, making it at least continuously differentiable (in the sense of classical derivatives), this fact will establish a global continuous differentiability of V on $\mathbb{R}^d$.

We will conclude this section with a few remarks on the interior regularity of V on C. It is well known that this can be achieved by considering the Dirichlet/Poisson problem $\mathbb{L}_X V = -H$ on a ball (elliptic case) or a rectangle (parabolic case) contained in C where the boundary values are determined by the value function V itself upon knowing/establishing that V is continuous (which normally presents no difficulty in specific examples). Since the boundary of a ball or a rectangle is known to be sufficiently regular we know that the Dirichlet/Poisson problem can be solved uniquely. For example, when $\nu \equiv 0$ in (2.14) it is known that (locally) Hölder coefficients in (2.14) yield a unique solution which is C^2 in the space variables and C^1 in the time variable (see [23, Theorem 6.13, p. 106] for the elliptic case and [19, Theorem 9, p. 69] for the parabolic case). This solution can then be identified with the value function V itself using the stochastic calculus or infinitesimal generator techniques as described above (see [41, p. 131] for further details) thus establishing the interior regularity of V on C as claimed. The central aim of the present paper is to establish the C^1 regularity of the value function V at the optimal stopping boundary ∂C that in turn is not accessible by these arguments.

3. Green regularity

In this section we present two sufficient conditions for the Green regularity of boundary points as defined in (2.8) above. The first condition is contained in the first equivalence of

(2.11) and we expose its proof for completeness and comparison (Lemma 1 & Corollary 2). The second condition (Lemma 4 & Corollary 5) has its origin in the facts that the mapping $x \mapsto \mathsf{E}_x(Z)$ is finely continuous if $Z \circ \theta_t \to Z$ as $t \downarrow 0$ where θ_t denotes the shift operator and the implication is applicable to $Z = \sigma_U$ when U is an open set in $I\!R^d$ (see [15, Corollaries 1 & 2, p. 123]). The two sufficient conditions applied to stochastic flows (Corollaries 3 & 6) will be used in the proofs of the main results in Sections 4 and 5 below.

Throughout this section we recall/assume that the (standard Markov) process $X = (X_t)_{t \ge 0}$ and the filtration $(\mathcal{F}_t)_{t \ge 0}$ (to which X is adapted) are *right-continuous* so that the first entry and hitting times of X to Borel (open and closed) sets are stopping times (cf. [3, Theorem 10.7, p. 54]). Recall that C denotes the continuation (open) set, $D = I\!R^d \setminus C$ denotes the stopping (closed) set, and ∂C denotes the boundary of the set C (see Section 2 above).

Lemma 1. *If X is strong Feller then*

$$x \mapsto \mathsf{P}_x(\sigma_D \ge \varepsilon) \text{ is upper semicontinuous on } I\!R^d \tag{3.1}$$

for each $\varepsilon > 0$ given and fixed.

Proof. Using that $\delta + \sigma_D \circ \theta_\delta \downarrow \sigma_D$ as $\delta \downarrow 0$, and letting $\varepsilon > 0$ be given and fixed, we find by the strong Markov property of X that

$$\begin{aligned} \mathsf{P}_x(\sigma_D \ge \varepsilon) &= \lim_{\delta \downarrow 0} \mathsf{P}_x(\delta + \sigma_D \circ \theta_\delta \ge \varepsilon) = \lim_{\delta \downarrow 0} \mathsf{E}_x\big[\mathsf{E}_x(I(\sigma_D \circ \theta_\delta \ge \varepsilon - \delta) \,|\, \mathcal{F}_\delta)\big] \\ &= \lim_{\delta \downarrow 0} \mathsf{E}_x\big[\mathsf{E}_{X_\delta}(I(\sigma_D \ge \varepsilon - \delta))\big] = \lim_{\delta \downarrow 0} \mathsf{E}_x\big[\mathsf{P}_{X_\delta}(\sigma_D \ge \varepsilon - \delta)\big] \\ &= \lim_{\delta \downarrow 0} \mathsf{E}_x\big[F_\delta(X_\delta)\big] = \lim_{\delta \downarrow 0} G_\delta(x) \end{aligned} \tag{3.2}$$

where $x \mapsto F_\delta(x) := \mathsf{P}_x(\sigma_D \ge \varepsilon - \delta)$ is measurable so that $x \mapsto G_\delta(x) := \mathsf{E}_x\big[F_\delta(X_\delta)\big]$ is continuous on $I\!R^d$ by the strong Feller property of X. Since moreover $\delta \mapsto G_\delta$ is decreasing on $(0, \infty)$ as $\delta \downarrow 0$, we see from (3.2) that (3.1) is satisfied as claimed. $\square$

Corollary 2. *If $x \in \partial C$ is probabilistically regular for D and X is strong Feller, then x is Green regular for D.*

Proof. Take any $x_n \in C$ converging to $x \in \partial C$ as $n \to \infty$. Then by (3.1) we get

$$0 \le \liminf_{n \to \infty} \mathsf{P}_{x_n}(\tau_D \ge \varepsilon) \le \limsup_{n \to \infty} \mathsf{P}_{x_n}(\tau_D \ge \varepsilon) \le \limsup_{n \to \infty} \mathsf{P}_{x_n}(\sigma_D \ge \varepsilon) \le \mathsf{P}_x(\sigma_D \ge \varepsilon) = 0 \tag{3.3}$$

for each $\varepsilon > 0$ given and fixed, where the final equality follows by probabilistic regularity of x for D. This shows that (2.8) is satisfied as claimed. $\square$

When the process X can be realised as a stochastic flow $(X_t^x)_{t \ge 0, x \in I\!R^d}$ we write

$$\tau_D^x = \inf\{\, t \ge 0 \mid X_t^x \in D \,\} \;\; \& \;\; \sigma_D^x = \inf\{\, t > 0 \mid X_t^x \in D \,\} \tag{3.4}$$

to denote the dependence of τ_D and σ_D on $x \in I\!R^d$. In this case we can reformulate the result of Corollary 2 as follows.

Corollary 3. *If* $x \in \partial C$ *is probabilistically regular for* D *and* X *is strong Feller, then* $\tau_D^{x_n} \to 0$ *in probability whenever* $x_n \in C$ *converges to* $x \in \partial C$ *as* $n \to \infty$ *.*

Proof. This is a direct consequence of the Green regularity established in Corollary 2. $\square$

When the process X fails to be strong Feller then the conclusions of Lemma 1, Corollary 2 and Corollary 3 can generally fail under probabilistic regularity of a point from ∂C for the set D. We now show that the conclusions remain valid if X can be realised as a stochastic flow that is continuous in the *space* variable and a point from ∂C is probabilistically regular for the *interior* D° of the set D.

Lemma 4. *If* X *can be realised as a stochastic flow such that*

$$x \mapsto X_t^x \text{ is continuous on } I\!R^d \tag{3.5}$$

almost surely for each $t \geq 0$ *given and fixed, then*

$$x \mapsto \mathsf{P}_x(\sigma_{D^\circ} \geq \varepsilon) \text{ is upper semicontinuous on } I\!R^d \tag{3.6}$$

for each $\varepsilon > 0$ *given and fixed.*

Proof. We first show that

$$x \mapsto \sigma_{D^\circ}^x \text{ is upper semicontinuous on } I\!R^d \tag{3.7}$$

almost surely. For this, take any $x_n \to x$ in $I\!R^d$ as $n \to \infty$. Denoting the exceptional set of P-measure zero in (3.5) by N_t, and setting $N := \cup_{t \in \mathbb{Q}_+} N_t$ which also is a set of P-measure zero, we know that (3.5) holds on $\Omega \backslash N$ for every $t \in \mathbb{Q}_+$. Let $\omega \in \Omega \backslash N$ be given and fixed. By definition of $\sigma_{D^\circ}(\omega)$ and right-continuity of $t \mapsto X_t^x(\omega)$ we know that for $\varepsilon > 0$ given and fixed, there exists $t_\varepsilon \in (\sigma_{D^\circ}^x(\omega), \sigma_{D^\circ}^x(\omega) + \varepsilon) \cap \mathbb{Q}_+$ such that $X_{t_\varepsilon}^x(\omega) \in D^\circ$. Because D° is open it follows that there exists $\delta_\varepsilon > 0$ such that $b(X_{t_\varepsilon}^x(\omega), \delta_\varepsilon) \subseteq D^\circ$. Since (3.5) holds on $\Omega \setminus N$ for $t_\varepsilon \in \mathbb{Q}_+$ we see that there exists $n_\varepsilon \geq 1$ such that $X_{t_\varepsilon}^{x_n}(\omega) \in b(X_{t_\varepsilon}^x(\omega), \delta_\varepsilon)$ for all $n \geq n_\varepsilon$. This shows that $\sigma_{D^\circ}^{x_n}(\omega) \leq t_\varepsilon$ for all $n \geq n_\varepsilon$ and hence we find that $\limsup_{n \to \infty} \sigma_{D^\circ}^{x_n}(\omega) \leq t_\varepsilon$. Letting $\varepsilon \downarrow 0$ we get $\limsup_{n \to \infty} \sigma_{D^\circ}^{x_n}(\omega) \leq \sigma_{D^\circ}^x(\omega)$ and this establishes (3.7) as claimed.

We next show that (3.6) holds. For this, take any $x_n \to x$ in $I\!R^d$ as $n \to \infty$ and set $A_n = \{\, \sigma_{D^\circ}^{x_n} \geq \varepsilon \,\}$ for $n \geq 1$. Then by Fatou's lemma for sets we find that

$$\begin{aligned} \limsup_{n \to \infty} \mathsf{P}_{x_n}(\sigma_{D^\circ} \geq \varepsilon) &= \limsup_{n \to \infty} \mathsf{P}(\sigma_{D^\circ}^{x_n} \geq \varepsilon) = \limsup_{n \to \infty} \mathsf{P}(A_n) \leq \mathsf{P}\big(\limsup_{n \to \infty} A_n\big) \\ &\leq \mathsf{P}(\sigma_{D^\circ}^x \geq \varepsilon) = \mathsf{P}_x(\sigma_{D^\circ} \geq \varepsilon) \end{aligned} \tag{3.8}$$

where the second inequality follows since $\omega \in \limsup_{n \to \infty} A_n$ if and only if $\omega \in A_{n_k}$ for $k \geq 1$, so that $\sigma_{D^\circ}^{x_{n_k}}(\omega) \geq \varepsilon$ for $k \geq 1$ and hence by (3.7) we get $\sigma_{D^\circ}^x(\omega) \geq \limsup_{n \to \infty} \sigma_{D^\circ}^{x_n}(\omega) \geq \limsup_{k \to \infty} \sigma_{D^\circ}^{x_{n_k}}(\omega) \geq \varepsilon$ implying the claim. This shows that (3.6) is satisfied as claimed. $\square$

Corollary 5. *If* $x \in \partial C$ *is probabilistically regular for* D° *and* X *can be realised as a stochastic flow such that* (3.5) *holds, then* x *is Green regular for* D° *(and thus* D *too).*

Proof. Take any $x_n \in C$ converging to $x \in \partial C$ as $n \to \infty$. Then similarly to the proof of Corollary 2, we find by (3.6) that (3.3) holds with D° in place of D for each $\varepsilon > 0$ given and fixed, where the final equality follows by probabilistic regularity of x for D°. This shows that (2.8) is satisfied with D° in place of D as claimed. □

Corollary 6. *If $x \in \partial C$ is probabilistically regular for D° and X can be realised as a stochastic flow such that* (3.5) *holds, then $\tau_{D^\circ}^{x_n} \to 0$ almost surely (and thus $\tau_D^{x_n} \to 0$ almost surely too) whenever $x_n \in C$ converges to $x \in \partial C$ as $n \to \infty$.*

Proof. This is a direct consequence of (3.7) upon noting that $\tau_{D^\circ}^{x_n} = \sigma_{D^\circ}^{x_n}$ for $n \ge 1$ and $\tau_{D^\circ}^{x} = \sigma_{D^\circ}^{x}$ since D° is open. □

According to [13] and the references therein, a point $z \in \partial C$ that is regular for D° is called a *stable* boundary point, and the boundary ∂C is said to be (*strongly*) *transversal* if $\sigma_D = \sigma_{D^\circ}$ almost surely with respect to P_x for all $x \in C$ (for all $x \in I\!R^d$). Note that the results of Corollary 5 and Corollary 6 can be rephrased in terms of stable boundary points. An important example of the strongly transversal boundary is obtained as follows.

Example 7. If $t \mapsto b(t)$ is (piecewise) *monotone* and (left/right) *continuous* on $I\!R_+$ and

$$D = \{ (t,x) \in I\!R_+ \times I\!R \mid x \ge b(t) \} \tag{3.9}$$

then for any regular (recurrent) Itô-McKean diffusion X we have $\sigma_D = \sigma_{D^\circ}$ almost surely with respect to P_x for every $x = b(t)$ with $t \ge 0$ (see the proof of Corollary 8 in [7]). Note that Corollary 5 in this case implies that probabilistic regularity of a boundary point $z = (t,x) \in \partial C$ implies its Green regularity despite the fact that the time-space process $((t,X_t))_{t \ge 0}$ is not strong Feller so that the (general) first equivalence in (2.11) is not applicable.

4. Continuity of the space derivative

In this section we show that probabilistic regularity of the optimal stopping boundary implies continuous *spatial* differentiability of the value function at the optimal stopping boundary whenever the process admits a continuously differentiable flow.

1. We first consider the case of infinite horizon in Theorem 8. This will be then extended to the case of finite horizon in Theorem 10 below. Similarly to (3.4) above we write $\Lambda_t^x = \int_0^t \lambda(X_s^x)\, ds$ to denote the dependence of Λ_t on $x \in I\!R^d$ for $t \ge 0$. We set $\delta_{i,j} = 1$ if $i = j$ and $\delta_{i,j} = 0$ for $i \ne j$ with $1 \le i,j \le d$.

Theorem 8. *Consider the optimal stopping problem* (2.1) *upon assuming that it is well posed in the sense that the stopping time τ_D from* (2.4) *is optimal. Assume that*

(4.1) V *is continuous on $I\!R^d$ and continuously differentiable on C;*

(4.2) G *is continuously differentiable on $I\!R^d$;*

(4.3) H *and λ are Lipschitz continuous on $I\!R^d$ in the sense that*

$$|H(x)-H(y)| \le K|x-y| \ \ \& \ \ |\lambda(x)-\lambda(y)| \le K|x-y|$$

for all $x, y \in I\!R^d$ with some constant $K > 0$ large enough.

Assume moreover that the process X can be realised as a continuously differentiable stochastic flow $(X_t^x)_{t\ge 0, x\in I\!R^d}$ in the space variable and that for $z \in \partial C$ given and fixed we have

$$\mathsf{E}\Big[\sup_{\alpha,\beta,\xi\in b(z,r)} e^{-\Lambda^{\beta}_{\tau_D^{\alpha}}} \big|\partial_j G(X^{\xi}_{\tau_D^{\alpha}})\partial_i X^{j,\xi}_{\tau_D^{\alpha}}\big| \Big] < \infty \tag{4.4}$$

$$\mathsf{E}\Big[\sup_{\alpha\in b(z,r)} \int_0^{\tau_D^{\alpha}} \sup_{\beta,\eta\in b(z,r)} e^{-\Lambda^{\beta}_t} \big|\partial_i X^{j,\eta}_t\big|\, dt \Big] < \infty \tag{4.5}$$

$$\mathsf{E}\Big[\sup_{\alpha,\beta,\gamma\in b(z,r)} \Big(e^{-\Lambda^{\beta}_{\tau_D^{\alpha}}} \big|G(X^{\gamma}_{\tau_D^{\alpha}})\big| \int_0^{\tau_D^{\alpha}} \sup_{\eta\in b(z,r)} \big|\partial_i X^{j,\eta}_t\big|\, dt \Big) \Big] < \infty \tag{4.6}$$

$$\mathsf{E}\Big[\sup_{\alpha\in b(z,r)} \int_0^{\tau_D^{\alpha}} \Big(\sup_{\beta,\gamma\in b(z,r)} e^{-\Lambda^{\beta}_t} \big|H(X^{\gamma}_t)\big| \int_0^t \sup_{\eta\in b(z,r)} \big|\partial_i X^{j,\eta}_s\big|\, ds \Big)\, dt \Big] < \infty \tag{4.7}$$

for some $r > 0$ with $\partial_i X^{j,z}_{0+} = \delta_{i,j}$ for $1 \le i, j \le d$. If X is strong Feller and z is probabilistically regular for D, or X is strong Markov and z is probabilistically regular for D°, then

$$V \text{ is continuously differentiable at } z \tag{4.8}$$

with $\partial_i V(z) = \partial_i G(z)$ for $1 \le i \le d$. If the hypotheses stated above hold at every $z \in \partial C$ then V is continuously differentiable on $I\!R^d$.

Proof. It will be clear from the proof below that the same arguments are applicable in any dimension $d \ge 1$ so that for ease of notation we will assume that $d = 1$ in the sequel.

(I): To illustrate the arguments in a clearer manner we first consider the special case when $\Lambda_t = 0$ for $t \ge 0$. Note that the conditions (4.6) and (4.7) are not needed in that case.

1. Take any $x_n \in C$ converging to $z \in \partial C$ as $n \to \infty$. Passing to a subsequence of $(x_n)_{n\ge 1}$ if needed there is no loss of generality in assuming that

$$\liminf_{n\to\infty} V_x(x_n) = \lim_{n\to\infty} \frac{V(x_n{+}\varepsilon_n) - V(x_n)}{\varepsilon_n} \tag{4.9}$$

for some $\varepsilon_n \downarrow 0$ as $n \to \infty$ (we write V_x to denote $\partial V/\partial x$ throughout). Let $\tau_n := \tau_D^{x_n}$ be the optimal stopping time for $V(x_n)$ when $n \ge 1$. Then by the mean value theorem and (4.3) we find that

$$\begin{aligned} V(x_n{+}\varepsilon_n) - V(x_n) &\ge \mathsf{E}\Big[G(X^{x_n+\varepsilon_n}_{\tau_n}) + \int_0^{\tau_n} H(X^{x_n+\varepsilon_n}_t)\, dt\Big] - \mathsf{E}\Big[G(X^{x_n}_{\tau_n}) + \int_0^{\tau_n} H(X^{x_n}_t)\, dt\Big] \\ &= \mathsf{E}\Big[G(X^{x_n+\varepsilon_n}_{\tau_n}) - G(X^{x_n}_{\tau_n})\Big] + \mathsf{E}\Big[\int_0^{\tau_n} \big(H(X^{x_n+\varepsilon_n}_t) - H(X^{x_n}_t)\big)\, dt\Big] \\ &\ge \mathsf{E}\Big[G_x(X^{\xi_n}_{\tau_n})\, \partial_x X^{\xi_n}_{\tau_n}\, \varepsilon_n\Big] - \mathsf{E}\Big[\int_0^{\tau_n} K\, |\partial_x X^{\eta_n(t)}_t|\, \varepsilon_n\, dt\Big] \end{aligned} \tag{4.10}$$

where ξ_n and $\eta_n(t)$ belong to $(x_n, x_n{+}\varepsilon_n)$ for $n \ge 1$. Dividing both sides by ε_n and letting $n \to \infty$ we find from (4.9)+(4.10) that

$$\liminf_{n\to\infty} V_x(x_n) \ge \lim_{n\to\infty} \mathsf{E}\Big[G_x(X^{\xi_n}_{\tau_n})\, \partial_x X^{\xi_n}_{\tau_n}\Big] - K \lim_{n\to\infty} \mathsf{E}\Big[\int_0^{\tau_n} |\partial_x X^{\eta_n(t)}_t|\, dt\Big] = G_x(z) \tag{4.11}$$

where in the final equality we use that $\tau_n \to 0$ almost surely as $n \to \infty$ by Green regularity of z for D as established in Corollary 3 and Corollary 6 above (in the former case one may need to pass to a subsequence of $(x_n)_{n \ge 1}$ which is sufficient for the present purposes) combined with the dominated convergence theorem which is applicable due to (4.4) and (4.5) respectively.

2. Similarly, there is no loss of generality in assuming that

$$\limsup_{n\to\infty} V_x(x_n) = \lim_{n\to\infty} \frac{V(x_n) - V(x_n - \varepsilon_n)}{\varepsilon_n} \tag{4.12}$$

for some $\varepsilon_n \downarrow 0$ as $n \to \infty$. By the mean value theorem and (4.3) we find that

$$\begin{aligned} V(x_n) - V(x_n - \varepsilon_n) &\le \mathsf{E}\Big[G(X_{\tau_n}^{x_n}) + \int_0^{\tau_n} H(X_t^{x_n})\, dt\Big] - \mathsf{E}\Big[G(X_{\tau_n}^{x_n-\varepsilon_n}) + \int_0^{\tau_n} H(X_t^{x_n-\varepsilon_n})\, dt\Big] \\ &= \mathsf{E}\Big[G(X_{\tau_n}^{x_n}) - G(X_{\tau_n}^{x_n-\varepsilon_n})\Big] + \mathsf{E}\Big[\int_0^{\tau_n} \big(H(X_t^{x_n}) - H(X_t^{x_n-\varepsilon_n})\big)\, dt\Big] \\ &\le \mathsf{E}\Big[G_x(X_{\tau_n}^{\xi_n})\, \partial_x X_{\tau_n}^{\xi_n}\, \varepsilon_n\Big] + \mathsf{E}\Big[\int_0^{\tau_n} K\, |\partial_x X_t^{\eta_n(t)}|\, \varepsilon_n\, dt\Big] \end{aligned} \tag{4.13}$$

where ξ_n and $\eta_n(t)$ belong to $(x_n - \varepsilon_n, x_n)$ for $n \ge 1$. Dividing both sides by ε_n and letting $n \to \infty$ we find from (4.12)+(4.13) that

$$\limsup_{n\to\infty} V_x(x_n) \le \lim_{n\to\infty} \mathsf{E}\Big[G_x(X_{\tau_n}^{\xi_n})\, \partial_x X_{\tau_n}^{\xi_n}\Big] + K \lim_{n\to\infty} \mathsf{E}\Big[\int_0^{\tau_n} |\partial_x X_t^{\eta_n(t)}|\, dt\Big] = G_x(z) \tag{4.14}$$

where in the final equality we use the same arguments as following (4.11) above. Combining (4.11) and (4.14) we see that $\lim_{n\to\infty} V_x(x_n) = G_x(z)$ and this completes the proof when $\Lambda_t = 0$ for $t \ge 0$.

(II): Next we consider the general case when $\Lambda_t \ne 0$ for $t \ge 0$. Note that the conditions (4.6) and (4.7) are needed in that case unless λ is constant for all $t \ge 0$. The proof in the general case can be carried out along the same lines as in the special case above and we only highlight the needed modifications throughout.

3. Taking any $x_n \in C$ converging to $z \in \partial C$ as $n \to \infty$ and arguing as in (4.9) above, we see that the right-hand side of the first inequality in (4.10) reads as follows

$$\begin{aligned} &\mathsf{E}\Big[e^{-\Lambda_{\tau_n}^{x_n+\varepsilon_n}} G(X_{\tau_n}^{x_n+\varepsilon_n}) - e^{-\Lambda_{\tau_n}^{x_n}} G(X_{\tau_n}^{x_n})\Big] \\ &\quad + \mathsf{E}\Big[\int_0^{\tau_n} \Big(e^{-\Lambda_t^{x_n+\varepsilon_n}} H(X_t^{x_n+\varepsilon_n}) - e^{-\Lambda_t^{x_n}} H(X_t^{x_n})\Big)\, dt\Big] \\ &= \mathsf{E}\Big[e^{-\Lambda_{\tau_n}^{x_n}} \Big(e^{\Lambda_{\tau_n}^{x_n} - \Lambda_{\tau_n}^{x_n+\varepsilon_n}} - 1\Big) G(X_{\tau_n}^{x_n+\varepsilon_n})\Big] + \mathsf{E}\Big[e^{-\Lambda_{\tau_n}^{x_n}} \Big(G(X_{\tau_n}^{x_n+\varepsilon_n}) - G(X_{\tau_n}^{x_n})\Big)\Big] \\ &\quad + \mathsf{E}\Big[\int_0^{\tau_n} e^{-\Lambda_t^{x_n}} \Big(e^{\Lambda_t^{x_n} - \Lambda_t^{x_n+\varepsilon_n}} - 1\Big) H(X_t^{x_n+\varepsilon_n})\, dt\Big] \\ &\quad + \mathsf{E}\Big[\int_0^{\tau_n} e^{-\Lambda_t^{x_n}} \Big(H(X_t^{x_n+\varepsilon_n}) - H(X_t^{x_n})\Big)\, dt\Big] \end{aligned} \tag{4.15}$$

for $n \geq 1$. The second expectation and the fourth expectation on the right-hand side of (4.15) can be handled in exactly the same way as the corresponding two expectations in (4.10), and this yields the conclusion of (4.11) above, i.e.

$$\liminf_{n\to\infty} V_x(x_n) \geq G_x(z) \tag{4.16}$$

provided that the liminf of the first expectation on the right-hand side of (4.15) divided by ε_n and the liminf of the third expectation on the right-hand side of (4.15) divided by ε_n are non-negative as $n \to \infty$. To see that both liminfs are non-negative, note that (4.3) and the mean value theorem imply that

$$\begin{aligned} \frac{e^{\Lambda^{x_n}_{\sigma_n} - \Lambda^{x_n+\varepsilon_n}_{\sigma_n}} - 1}{\varepsilon_n} &\geq \frac{e^{-\varepsilon_n K \int_0^{\sigma_n} |\partial_x X_s^{\eta_n(s)}|\, ds} - 1}{\varepsilon_n} \\ &= -K\Big(\int_0^{\sigma_n} |\partial_x X_s^{\eta_n(s)}|\, ds\Big)\, e^{-\zeta_n K \int_0^{\sigma_n} |\partial_x X_s^{\eta_n(s)}|\, ds} \end{aligned} \tag{4.17}$$

with σ_n equal to either τ_n (the first expectation) or $t \in [0, \tau_n]$ (the third expectation) where $\eta_n(s)$ belongs to $(x_n, x_n + \varepsilon_n)$ and ζ_n belongs to $(0, \varepsilon_n)$ for $n \geq 1$. Using then the same arguments as in (4.11) above with (4.6)+(4.7) in place of (4.4)+(4.5), we see that the inequality (4.17) yields the fact that the two liminfs are non-negative so that (4.16) holds as claimed.

4. Similarly, arguing as in (4.12) we see that the right-hand side of the first inequality in (4.13) reads as follows

$$\begin{aligned} &\mathsf{E}\Big[e^{-\Lambda^{x_n}_{\tau_n}} G(X^{x_n}_{\tau_n}) - e^{-\Lambda^{x_n-\varepsilon_n}_{\tau_n}} G(X^{x_n-\varepsilon_n}_{\tau_n})\Big] \\ &\quad + \mathsf{E}\Big[\int_0^{\tau_n} \Big(e^{-\Lambda^{x_n}_t} H(X^{x_n}_t) - e^{-\Lambda^{x_n-\varepsilon_n}_t} H(X^{x_n-\varepsilon_n}_t)\Big)\, dt\Big] \\ &= \mathsf{E}\Big[e^{-\Lambda^{x_n}_{\tau_n}}\Big(1 - e^{\Lambda^{x_n}_{\tau_n} - \Lambda^{x_n-\varepsilon_n}_{\tau_n}}\Big) G(X^{x_n}_{\tau_n})\Big] + \mathsf{E}\Big[e^{-\Lambda^{x_n-\varepsilon_n}_{\tau_n}}\Big(G(X^{x_n}_{\tau_n}) - G(X^{x_n-\varepsilon_n}_{\tau_n})\Big)\Big] \\ &\quad + \mathsf{E}\Big[\int_0^{\tau_n} e^{-\Lambda^{x_n}_t}\Big(1 - e^{\Lambda^{x_n}_t - \Lambda^{x_n-\varepsilon_n}_t}\Big) H(X^{x_n}_t)\, dt\Big] \\ &\quad + \mathsf{E}\Big[\int_0^{\tau_n} e^{-\Lambda^{x_n-\varepsilon_n}_t}\Big(H(X^{x_n}_t) - H(X^{x_n-\varepsilon_n}_t)\Big)\, dt\Big] \end{aligned} \tag{4.18}$$

for $n \geq 1$. The second expectation and the fourth expectation on the right-hand side of (4.18) can be handled in exactly the same way as the corresponding two expectations in (4.13) and this yields the conclusion of (4.14) above, i.e.

$$\limsup_{n\to\infty} V_x(x_n) \leq G_x(z) \tag{4.19}$$

provided that the limsup of the first expectation on the right-hand side of (4.18) divided by ε_n and the limsup of the third expectation on the right-hand side of (4.18) divided by ε_n are non-positive as $n \to \infty$. To see that both limsups are non-positive, note that (4.3) and the mean value theorem imply that

$$\frac{1 - e^{\Lambda^{x_n}_{\sigma_n} - \Lambda^{x_n-\varepsilon_n}_{\sigma_n}}}{\varepsilon_n} \leq \frac{1 - e^{-\varepsilon_n K \int_0^{\sigma_n} |\partial_x X_s^{\eta_n(s)}|\, ds}}{\varepsilon_n} \tag{4.20}$$

$$= K\big(\int_0^{\sigma_n} |\partial_x X_s^{\eta_n(s)}|\, ds\big)\, e^{-\zeta_n K \int_0^{\sigma_n} |\partial_x X_s^{\eta_n(s)}|\, ds}$$

with σ_n equal to either τ_n (the first expectation) or $t \in [0,\tau_n]$ (the third expectation) where $\eta_n(s)$ belongs to $(x_n{-}\varepsilon_n, x_n)$ and ζ_n belongs to $(0,\varepsilon_n)$ for $n \ge 1$. Using then the same arguments as in (4.14) above with (4.6)+(4.7) in place of (4.4)+(4.5), we see that the inequality (4.20) yields the fact that the two limsups are non-positive so that (4.19) holds as claimed. Combining (4.16) and (4.19) we see that $\lim_{n\rightarrow\infty} V_x(x_n) = G_x(z)$ and this completes the proof when $\Lambda_t \neq 0$ for $t \ge 0$. □

Remark 9. Note that the conditions (4.4)-(4.7) are used in the proof above as sufficient conditions for the dominated convergence theorem to establish the convergence relations (4.11) and (4.14) (when λ is zero) and their extensions (4.16) and (4.19) (when λ is not constant). (Recall from the proof that the conditions (4.6) and (4.7) are not needed when λ is constant.) These sufficient conditions, although applicable in a large number of examples, are not necessary in general and in some specific examples one can often exploit additional information (e.g. the geometric/analytic structure of the optimal stopping boundary) and derive the convergence relations without appealing to the dominated convergence theorem (see the proof of Theorem 3.1 in [37] for such an example). As it is exceedingly complicated to describe all possible ways that lead to relaxed forms of the sufficient conditions (4.4)-(4.7), we have stated them in their present form with a view that the structure of the proof above remains unchanged if these sufficient conditions are replaced by other/weaker ones. A similar remark applies to the condition (4.3). For instance, replacing the global Lipschitz continuity of H in (4.3) by a local Lipschitz continuity in the sense that

$$|H(x){-}H(y)| \le K_n|x{-}y| \tag{4.21}$$

for all $x,y \in b(z,R_n)$ with some constant $K_n > 0$ large enough where $R_n \rightarrow \infty$ as $n \rightarrow \infty$, it is seen from the proof above that the result of Theorem 8 (with $\lambda = 0$) remains valid if

$$\lim_{n\rightarrow\infty} K_n\, \mathsf{E}\Big[\int_0^{\bar{\tau}_n} |\partial_x X_t^{\eta_n}|\, dt\Big] = 0 \tag{4.22}$$

where $\bar{\tau}_n := \tau_n \wedge \inf\{\, t \ge 0 \mid X_t^{x_n+\varepsilon_n} \notin b(z,R_n)$ or $X_t^{x_n} \notin b(z,R_n)\,\}$ and $R_n > 0$ is chosen large enough so that

$$\mathsf{E}\Big[\int_0^{\bar{\tau}_n} \big|H(X_t^{x_n\pm\varepsilon_n}){-}H(X_t^{x_n})\big|\, dt\Big] \ge \mathsf{E}\Big[\int_0^{\tau_n} \big|H(X_t^{x_n\pm\varepsilon_n}){-}H(X_t^{x_n})\big|\, dt\Big] - \varepsilon_n\delta \tag{4.23}$$

for all $n \ge 1$ with $\delta > 0$ given and fixed. Similarly, the global Lipschitz continuity of λ in (4.3) can be replaced by a local Lipschitz continuity and we will omit further details. Finally, the proof above shows that it is sufficient to have continuous differentiability of the flow near the optimal stopping boundary only.

2. The optimal stopping problem (2.1) considered in Theorem 8 has infinite horizon. The arguments used in the proof carry over to the optimal stopping problem (2.2) with finite horizon as long as continuous *spatial* differentiability of the value function is considered. We formally present this extension in the next theorem. Continuous *temporal* differentiability of the value function requires different arguments and will be considered in the next section.

Recall that the optimal stopping problem (2.2) includes the case when the functions G and H are time dependent which can be formally obtained by setting $X^1_t = t$ for $t \geq 0$. Thus the process X in this case is given by $X_t = (t, X^2_t, \ldots, X^d_t)$ for $t \geq 0$. The continuation set is given by $C = \{\, (t,x) \in [0,T] \times I\!R^{d-1} \mid V(t,x) > G(t,x) \,\}$ and the stopping set is given by $D = \{\, (t,x) \in [0,T] \times I\!R^{d-1} \mid V(t,x) = G(t,x) \,\}$. Note that the process $C := X^1$ can always be realised as a stochastic flow by setting $C^t_s = t+s$ for $t \geq 0$ and $s \geq 0$. Hence when $(X^2, \ldots, X^d)$ can be realised as a stochastic flow in the space variable x from $I\!R^{d-1}$ we will denote the entire flow by $(X^{t,x}_s)$ for $s \geq 0$ and $(t,x) \in [0,T] \times I\!R^{d-1}$. Note that $X^{t,x}_0 = (t, x_2, \ldots, x_d)$ for $t \in [0,T]$ and $x = (x_2, \ldots, x_d) \in I\!R^{d-1}$.

Theorem 10. *Consider the optimal stopping problem* (2.2) *upon assuming that it is well posed in the sense that the stopping time* τ_D *from* (2.4) *is optimal. Assume that*

(4.24) V *is continuous on* $[0,T] \times I\!R^{d-1}$ *and continuously differentiable on* C;

(4.25) G *is continuously differentiable on* $[0,T] \times I\!R^{d-1}$;

(4.26) $x \mapsto H(t,x)$ *and* $x \mapsto \lambda(t,x)$ *are Lipschitz continuous on* $I\!R^{d-1}$ *in the sense that*

$$|H(t,x) - H(t,y)| \leq K|x-y| \;\;\&\;\; |\lambda(t,x) - \lambda(t,y)| \leq K|x-y|$$

for every $t \in [0,T]$ *and all* $x, y \in I\!R^{d-1}$ *with some constant* $K > 0$ *large enough.*

Assume moreover that the process X *can be realised as a continuously differentiable stochastic flow* $(X^{t,x}_s)$ *in the space variable for* $s \geq 0$ *and* $(t,x) \in [0,T] \times I\!R^{d-1}$ *and that for* $z \in \partial C$ *given and fixed the conditions* (4.4)-(4.7) *are satisfied for some* $r > 0$ *with* $\partial_i X^{j,z}_{0+} = \delta_{i,j}$ *for* $2 \leq i \leq d$ *and* $1 \leq j \leq d$. *If* z *is probabilistically regular for* D° *then*

$$\partial_2 V, \ldots, \partial_d V \;\; \textit{exist and are continuous at } z \tag{4.27}$$

with $\partial_i V(z) = \partial_i G(z)$ *for* $2 \leq i \leq d$. *If the hypotheses stated above hold at every* $z \in \partial C$ *then* $\partial_2 V, \ldots, \partial_d V$ *exist and are continuous on* $[0,T] \times I\!R^{d-1}$.

Proof. This can be established using exactly the same arguments as in the proof of Theorem 8 upon noting that adding ε_n to any but the first (time) coordinate of the process X does not alter the remaining time horizon. $\square$

Remark 11. Note that the comments on the sufficient conditions from Theorem 8 made in Remark 9 above extend to the corresponding sufficient conditions in Theorem 10 and we will omit further details in this direction.

3. The result and proof of Theorems 8 and 10 extend to the case when the gain function G in the optimal stopping problem (2.1)/(2.2) is *not smooth* away from the optimal stopping boundary ∂C. Instead of formulating a general theorem of this kind, which would be overly technical and rather difficult to read, we will illustrate key arguments of such extensions through an important example next. A different method of proof is based on extensions of the Itô-Tanaka formula dealing with singularities of G on curves and surfaces (cf. [38] and [39]) and this will be presented in the next section.

Example 12 (Continuity of the space derivative in the American put). Consider the optimal stopping problem

$$V(t,x) = \sup_{0\le\tau\le T-t} \mathsf{E}\Big[e^{-r\tau}\big(K - X_\tau^x\big)^+\Big] \tag{4.28}$$

where $(t,x)\in[0,T]\times(0,\infty)$, $r>0$, $K>0$ and the supremum is taken over stopping times τ of X solving the stochastic differential equation

$$dX_t = rX_t\,dt + \sigma X_t\,dB_t \tag{4.29}$$

with $X_0 = x$ where $\sigma>0$ and B is a standard Brownian motion (see [41, Section 25] for further details). Horizon in the optimal stopping problem (4.28) is finite so that the setting belongs to Theorem 10 above. Since the gain function $G(x) := (K-x)^+$ for $x>0$ is not differentiable at K we see that the condition (4.25) fails and hence we cannot conclude that

$$V_x \text{ is continuous on } \partial C \tag{4.30}$$

using Theorem 10 (we write V_x to denote $\partial V/\partial x$ throughout). We will now show however that the method of proof of Theorems 8 and 10 extends to cover the case of the non-differentiable gain function $G(x) = (K-x)^+$ for $x>0$. This will also serve as an illustration of how similar other cases of non-smooth gain functions G in the optimal stopping problem (2.1)/(2.2) can be handled. The derivation of (4.30) will be divided in three steps as follows.

1. Well-known arguments show that the optimal stopping time in (4.28) equals $\tau_D^{t,x} = \inf\{\, s\in[0,T-t] \mid X_s^x \le b(t+s)\,\}$ where the optimal stopping boundary $t\mapsto b(t)$ is increasing on $[0,T]$ with $0<b(0)<b(T)=K$ (see [41, Subsection 25.2]). If a point $z=(t,b(t))\in\partial C$ is given and fixed, then by the increase of b combined with the law of iterated logarithm for standard Brownian motion (cf. [28, p. 112]) we see that z is probabilistically regular for D° (formally this could also be derived from probabilistic regularity of z for D combined with the fact of Example 7 above). Since X can be realised as a continuous stochastic flow $x\mapsto xX_t^1$ on $(0,\infty)$, where we set $X_t^1 = \exp(\sigma B_t + (r-\sigma^2/2)\,t)$ for $t\ge 0$, it follows by Corollary 5 that z is Green regular for D° . Taking any sequence $(t_n,x_n)\in C$ converging to z as $n\to\infty$, it follows therefore by Corollary 6 that $\tau_D^{t_n,x_n}\to 0$ almost surely as $n\to\infty$. Note that the latter Green regularity has been obtained without appeal to a strong Feller property which fails for the time-space process $((t,X_t))_{0\le t\le T}$ in this case.

2. We next connect to the first part of the proof of Theorems 8 and 10. Passing to a subsequence of $((t_n,x_n))_{n\ge1}$ if needed there is no loss of generality in assuming that

$$\liminf_{n\to\infty} V_x(t_n,x_n) = \lim_{n\to\infty} \frac{V(t_n,x_n+\varepsilon_n)-V(t_n,x_n)}{\varepsilon_n} \tag{4.31}$$

for some $\varepsilon_n\downarrow 0$ as $n\to\infty$. Let $\tau_n := \tau_D^{t_n,x_n}$ denote the optimal stopping time for $V(t_n,x_n)$ when $n\ge 1$. Then using that $K>x_nX_{\tau_n}^1$ if and only if $\tau_n < T-t_n$ we find that

$$\begin{aligned} V(t_n,x_n+\varepsilon_n)-V(t_n,x_n) &\ge \mathsf{E}\Big[e^{-r\tau_n}\big(K-(x_n+\varepsilon_n)X_{\tau_n}^1\big)^+\Big] - \mathsf{E}\Big[e^{-r\tau_n}\big(K-x_nX_{\tau_n}^1\big)^+\Big] \\ &\ge \mathsf{E}\Big[\Big(e^{-r\tau_n}\big(K-(x_n+\varepsilon_n)X_{\tau_n}^1\big) - e^{-r\tau_n}\big(K-x_nX_{\tau_n}^1\big)\Big)I(\tau_n < T-t_n)\Big] \end{aligned} \tag{4.32}$$

$$= \mathsf{E}\Big[e^{-r\tau_n}(-\varepsilon_n)X^1_{\tau_n} I(\tau_n < T-t_n)\Big]$$

for $n \geq 1$. Dividing both sided by ε_n and letting $n \to \infty$ we find from (4.31)+(4.32) that

$$\liminf_{n\to\infty} V_x(t_n, x_n) \geq -\lim_{n\to\infty} \mathsf{E}\Big[e^{-r\tau_n} X^1_{\tau_n} I(\tau_n < T-t_n)\Big] = -1 \tag{4.33}$$

where in the last equality we use that $\tau_n \to 0$ almost surely as $n \to \infty$ combined with the dominated convergence theorem due to $\mathsf{E}(\sup_{0\leq t\leq T} X^1_t) < \infty$.

3. We finally connect to the second part of the proof of Theorems 8 and 10. Similarly, there is no loss of generality in assuming that

$$\limsup_{n\to\infty} V_x(t_n, x_n) = \lim_{n\to\infty} \frac{V(t_n, x_n) - V(t_n, x_n - \varepsilon_n)}{\varepsilon_n} \tag{4.34}$$

for some $\varepsilon_n \downarrow 0$ as $n \to \infty$. Then using the same arguments as in (4.32) we find that

$$\begin{aligned} V(t_n, x_n) - V(t_n, x_n - \varepsilon_n) &\leq \mathsf{E}\Big[e^{-r\tau_n}\big(K - x_n X^1_{\tau_n}\big)^+\Big] - \mathsf{E}\Big[e^{-r\tau_n}\big(K - (x_n - \varepsilon_n)X^1_{\tau_n}\big)^+\Big] \\ &\leq \mathsf{E}\Big[\Big(e^{-r\tau_n}\big(K - x_n X^1_{\tau_n}\big) - e^{-r\tau_n}\big(K - (x_n - \varepsilon_n)X^1_{\tau_n}\big)\Big) I(\tau_n < T-t_n)\Big] \\ &= \mathsf{E}\Big[e^{-r\tau_n}(-\varepsilon_n)X^1_{\tau_n} I(\tau_n < T-t_n)\Big] \end{aligned} \tag{4.35}$$

for $n \geq 1$. Dividing both sided by ε_n and letting $n \to \infty$ we find from (4.34)+(4.35) that

$$\limsup_{n\to\infty} V_x(t_n, x_n) \leq -\lim_{n\to\infty} \mathsf{E}\Big[e^{-r\tau_n} X^1_{\tau_n} I(\tau_n < T-t_n)\Big] = -1 \tag{4.36}$$

where in the last equality we use the same arguments as in (4.33) above. Combining (4.33) and (4.36) we see that $\lim_{n\to\infty} V_x(t_n, x_n) = G_x(z) = -1$ and this completes the proof of (4.30).

5. Continuity of the time derivative

In this section we show that probabilistic regularity of the optimal stopping boundary implies continuous temporal differentiability of the value function at the optimal stopping boundary whenever the process admits a continuous flow. We assume throughout that the process is given by $X_t = (t, X^2_t, \ldots, X^d_t)$ for $t \geq 0$ as discussed prior to Theorem 10 above.

1. We first consider the case of infinite horizon in Theorem 13. This will be then extended to the case of finite horizon in Theorem 15 below.

Theorem 13. *Consider the optimal stopping problem* (2.1) *upon assuming that it is well posed in the sense that the stopping time* τ_D *from* (2.4) *is optimal. Assume that*

(5.1) V *is continuous on* $\mathbb{R}_+ \times \mathbb{R}^{d-1}$ *and continuously differentiable on* C ;

(5.2) G *is continuously differentiable on* $\mathbb{R}_+ \times \mathbb{R}^{d-1}$;

(5.3) $t \mapsto H(t, x)$ *and* $t \mapsto \lambda(t, x)$ *are Lipschitz continuous on* $\mathbb{R}_+$ *in the sense that*

$$|H(t,x)-H(s,x)| \le K|t-s| \;\; \& \;\; |\lambda(t,x)-\lambda(s,x)| \le K|t-s|$$

for all $t,s \in I\!R_+$ *and every* $x \in I\!R^{d-1}$ *with some constant* $K>0$ *large enough.*

Assume moreover that the process X *can be realised as a continuous stochastic flow* $(X_s^{t,x})$ *in the space variable for* $s \ge 0$ *and* $(t,x) \in I\!R_+ \times I\!R^{d-1}$ *and that for* $z \in \partial C$ *given and fixed the following conditions are satisfied*

$$\mathsf{E}\Big[\sup_{\alpha,\beta,\xi\in b(z,r)} e^{-\Lambda^\beta_{\tau_D^\alpha}} \big|\partial_t G(X^\xi_{\tau_D^\alpha})\big|\Big] < \infty \tag{5.4}$$

$$\mathsf{E}\Big[\sup_{\alpha\in b(z,r)} \int_0^{\tau_D^\alpha} \sup_{\beta\in b(z,r)} e^{-\Lambda^\beta_t}\, dt\Big] < \infty \tag{5.5}$$

$$\mathsf{E}\Big[\Big(\sup_{\alpha,\beta,\gamma\in b(z,r)} e^{-\Lambda^\beta_{\tau_D^\alpha}} \big|G(X^\gamma_{\tau_D^\alpha})\big|\, \tau_D^\alpha\Big)\Big] < \infty \tag{5.6}$$

$$\mathsf{E}\Big[\sup_{\alpha\in b(z,r)} \int_0^{\tau_D^\alpha} \Big(\sup_{\beta,\gamma\in b(z,r)} e^{-\Lambda^\beta_t} \big|H(X^\gamma_t)\big|\, t\Big)\, dt\Big] < \infty \tag{5.7}$$

for some $r>0$ *. If* z *is probabilistically regular for* D° *then*

$$\partial_t V \text{ exists and is continuous at } z \tag{5.8}$$

with $\partial_t V(z) = \partial_t G(z)$ *. If the hypotheses stated above hold at every* $z \in \partial C$ *then* $\partial_t V$ *exists and is continuous on* $I\!R_+ \times I\!R^{d-1}$ *.*

Proof. Due to $X_t = (t, X_t^2, \ldots, X_t^d)$ for $t \ge 0$ as assumed throughout we see that the setting of Theorem 13 reduces to the setting of Theorem 8. All the claims therefore follow by applying Theorem 8 upon noting that $\partial_t X_t^{1,z} = 1$ and $\partial_t X_t^{i,z} = 0$ for $2 \le i \le d$ with $t \ge 0$ and $z \in I\!R_+ \times I\!R^{d-1}$ so that the sufficient conditions (4.4)-(4.7) in Theorem 8 transform to the sufficient conditions (5.4)-(5.7) stated above. $\square$

Remark 14. Note that the comments on the sufficient conditions from Theorem 8 made in Remark 9 above extend to the corresponding sufficient conditions in Theorem 13 and we will omit further details in this direction.

2. The optimal stopping problem considered in Theorem 13 has infinite horizon and the arguments used in the proof are analogous to the arguments used in the proofs of Theorems 8 and 10 above. Continuous temporal differentiability of the value function on finite horizon requires different arguments and will be considered in the next theorem. A key difficulty in the previous approach is that adding ε_n to the first (time) coordinate of the process X (see (4.10) above) alters the remaining time horizon so that the stopping time which is *optimal* for $V(t_n, x_n)$ is *no longer* admissible for $V(t_n+\varepsilon_n, x_n)$ with $n \ge 1$. To overcome this difficulty we will apply a Taylor expansion of the second order (Itô's formula) instead of the first order as in the proofs of Theorems 8 and 10 above.

Theorem 15. *Consider the optimal stopping problem* (2.2) *upon assuming that it is well posed in the sense that the stopping time* τ_D *from* (2.4) *is optimal. Assume that*

$$V \text{ is continuous on } [0,T]\times I\!R^{d-1} \text{ and continuously differentiable on } C\,; \tag{5.9}$$

(5.10) $(t,x) \mapsto G(t,x)$ *is once continuously differentiable with respect to* t *and twice continuously differentiable with respect to* x *on* $[0,T]\times I\!R^{d-1}$;

(5.11) $t \mapsto \tilde{H}(t,x) := (G_t + I\!L_X G + H)(t,x)$ *and* $t \mapsto \lambda(t,x)$ *are Lipschitz continuous on* $[0,T]$ *in the sense that*

$$|\tilde{H}(t,x) - \tilde{H}(s,x)| \le K|t-s| \;\;\&\;\; |\lambda(t,x) - \lambda(s,x)| \le K|t-s|$$

for all $t,s \in [0,T]$ *and every* $x \in I\!R^{d-1}$ *with some constant* $K > 0$ *large enough.*

Assume moreover that the process X *can be realised as a continuous stochastic flow* $(X_s^{t,x})$ *in the space variable for* $s \in [0,T-t]$ *and* $(t,x) \in [0,T] \times I\!R^{d-1}$ *and that for* $z \in \partial C$ *given and fixed the following conditions are satisfied*

$$\mathsf{E}\Big[e^{-\Lambda_\sigma^{t,x}} G(t+\sigma, X_\sigma^x)\Big] = G(t,x) + \mathsf{E}\Big[\int_0^\sigma e^{-\Lambda_s^{t,x}} (G_t + I\!L_X G)(t+s, X_s^x)\, ds\Big] \tag{5.12}$$

$$\mathsf{E}\Big[\sup_{(t,x)\in b(z,\varepsilon)} \sup_{T-t-\varepsilon \le s \le T-t} e^{-\Lambda_s^{t,x}} \big|\tilde{H}(t+s, X_s^x)\big|\Big] < \infty \tag{5.13}$$

for all stopping times σ *of* X *with values in* $[0,T-t]$ *and all* $(t,x) \in b(z,\varepsilon)$ *with some* $\varepsilon > 0$. *If* z *is probabilistically regular for* D° *then*

$$\partial_t V \;\text{ exists and is continuous at } z \tag{5.14}$$

with $\partial_t V(z) = \partial_t G(z)$. *If the hypotheses stated above hold at every* $z \in \partial C$ *then* $\partial_t V$ *is continuous on* $[0,T]\times I\!R^{d-1}$.

Proof. It will be clear from the proof below that the same arguments are applicable in any dimension $d \ge 1$ so that for ease of notation we will assume that $d = 1$ in the sequel.

(I): To illustrate the arguments in a clearer manner we first consider the special case when $\Lambda_t = 0$ for $t \ge 0$.

1. Take any $(t_n, x_n) \in C$ converging to $z \in \partial C$ as $n \to \infty$. Passing to a subsequence of $((t_n, x_n))_{n\ge 1}$ if needed there is no loss of generality in assuming that

$$\liminf_{n\to\infty} V_t(t_n, x_n) = \lim_{n\to\infty} \frac{V(t_n+\varepsilon_n, x_n) - V(t_n, x_n)}{\varepsilon_n} \tag{5.15}$$

for some $\varepsilon_n \downarrow 0$ as $n \to \infty$ (we write V_t to denote $\partial V/\partial t$ throughout). Let $\tau_n := \tau_D^{t_n,x_n}$ denote the optimal stopping time for $V(t_n, x_n)$ and set $\hat{\tau}_n := \tau_n \wedge (T - t_n - \varepsilon_n)$ for $n \ge 1$. Then by (5.11) and (5.12) we find that

$$\begin{aligned} &V(t_n+\varepsilon_n, x_n) - V(t_n, x_n) \\ &\ge G(t_n+\varepsilon_n, x_n) + \mathsf{E}\Big[\int_0^{\hat{\tau}_n} \big(G_t + I\!L_X G + H\big)(t_n+\varepsilon_n+s, X_s^{x_n})\, ds\Big] \\ &\quad - G(t_n, x_n) - \mathsf{E}\Big[\int_0^{\tau_n} \big(G_t + I\!L_X G + H\big)(t_n+s, X_s^{x_n})\, ds\Big] \end{aligned} \tag{5.16}$$

$$
\begin{aligned}
&= G(t_n+\varepsilon_n, x_n) - G(t_n, x_n) + \mathsf{E}\Big[\int_0^{\hat{\tau}_n} \big(\tilde{H}(t_n+\varepsilon_n+s, X_s^{x_n}) - \tilde{H}(t_n+s, X_s^{x_n}) \big)\, ds \Big] \\
&\quad - \mathsf{E}\Big[\int_{\hat{\tau}_n}^{\tau_n} \tilde{H}(t_n+s, X_s^{x_n})\, ds \Big] \\
&\geq G(t_n+\varepsilon_n, x_n) - G(t_n, x_n) - K\, \varepsilon_n\, \mathsf{E}(\tau_n) \\
&\quad - \mathsf{E}\Big[\sup_{T-t_n-\varepsilon_n \leq s \leq T-t_n} |\tilde{H}(t_n+s, X_s^{x_n})|\ \varepsilon_n\ I(T-t_n-\varepsilon_n < \tau_n \leq T-t_n) \Big]
\end{aligned}
$$

for $n \geq 1$. Dividing both sides by ε_n and letting $n \to \infty$ we find from (5.15) and (5.16) that

$$
\liminf_{n\to\infty} V_t(t_n, x_n) \geq G_t(z) \tag{5.17}
$$

where we use that $\tau_n \to 0$ almost surely as $n \to \infty$ by probabilistic regularity of z for D° and Corollary 6 above combined with the dominated convergence theorem which is applicable due to (5.13) above.

2. Similarly, there is no loss of generality in assuming that

$$
\limsup_{n\to\infty} V_t(t_n, x_n) = \lim_{n\to\infty} \frac{V(t_n, x_n) - V(t_n-\varepsilon_n, x_n)}{\varepsilon_n} \tag{5.18}
$$

for some $\varepsilon_n \downarrow 0$ as $n \to \infty$. By (5.11) and (5.12) we find that

$$
\begin{aligned}
V(t_n, x_n) - V(t_n-\varepsilon_n, x_n) &\leq G(t_n, x_n) - G(t_n-\varepsilon_n, x_n) \\
&\quad + \mathsf{E}\Big[\int_0^{\tau_n} \big(\tilde{H}(t_n+s, X_s^{x_n}) - \tilde{H}(t_n-\varepsilon_n+s, X_s^{x_n}) \big)\, ds \Big] \\
&\leq G(t_n, x_n) - G(t_n-\varepsilon_n, x_n) + K\, \varepsilon_n\, \mathsf{E}(\tau_n)
\end{aligned} \tag{5.19}
$$

for $n \geq 1$. Dividing both sides by ε_n and letting $n \to \infty$ we find from (5.18) and (5.19) that

$$
\limsup_{n\to\infty} V_t(t_n, x_n) \leq G_t(z) \tag{5.20}
$$

where we use the same arguments as following (5.17). Combining (5.17) and (5.20) we see that $\lim_{n\to\infty} V_t(t_n, x_n) = G_t(z)$ and this completes the proof when $\Lambda_t = 0$ for $t \geq 0$.

(II): Next we consider the general case when $\Lambda_t \neq 0$ for $t \geq 0$. The proof in the general case can be carried out along the same lines as in the special case above and we only highlight the needed modifications throughout.

3. Taking any $(t_n, x_n) \in C$ converging to $z \in \partial C$ as $n \to \infty$ and arguing as in (5.15) above, we see that the right-hand side of the first inequality in (5.16) reads as follows

$$
\begin{aligned}
&G(t_n+\varepsilon_n, x_n) - G(t_n, x_n) \\
&\quad + \mathsf{E}\Big[\int_0^{\hat{\tau}_n} \Big(e^{-\Lambda_s^{t_n+\varepsilon_n, x_n}}\, \tilde{H}(t_n+\varepsilon_n+s, X_s^{x_n}) - e^{-\Lambda_s^{t_n, x_n}}\, \tilde{H}(t_n+s, X_s^{x_n}) \Big)\, ds \Big] \\
&\quad - \mathsf{E}\Big[\int_{\hat{\tau}_n}^{\tau_n} e^{-\Lambda_s^{t_n, x_n}}\, \tilde{H}(t_n+s, X_s^{x_n})\, ds \Big]
\end{aligned} \tag{5.21}
$$

$$
\begin{aligned}
&= G(t_n{+}\varepsilon_n, x_n) - G(t_n, x_n) \\
&\quad + \mathsf{E}\Big[\int_0^{\hat{\tau}_n} e^{-\Lambda_s^{t_n,x_n}} \Big(e^{\Lambda_s^{t_n,x_n} - \Lambda_s^{t_n+\varepsilon_n,x_n}} - 1 \Big) \tilde{H}(t_n{+}\varepsilon_n{+}s, X_s^{x_n})\, ds \Big] \\
&\quad + \mathsf{E}\Big[\int_0^{\hat{\tau}_n} \Big(e^{-\Lambda_s^{t_n,x_n}} \Big(\tilde{H}(t_n{+}\varepsilon_n{+}s, X_s^{x_n}) - \tilde{H}(t_n{+}s, X_s^{x_n}) \Big)\, ds \Big] \\
&\quad - \mathsf{E}\Big[\int_{\hat{\tau}_n}^{\tau_n} e^{-\Lambda_s^{t_n,x_n}}\, \tilde{H}(t_n{+}s, X_s^{x_n})\, ds \Big]
\end{aligned}
$$

for $n \geq 1$. The second and third expectation on the right-hand side of (5.21) can be handled in exactly the same way as the corresponding expectations in (5.16), and this yields the conclusion of (5.17), provided that the liminf of the first expectation on the right-hand side of (5.21) divided by ε_n is non-negative as $n \to \infty$. To see that the liminf is non-negative, note that (5.11) and the mean value theorem imply that

$$
\frac{e^{\Lambda_s^{t_n,x_n} - \Lambda_s^{t_n+\varepsilon_n,x_n}} - 1}{\varepsilon_n} \geq \frac{e^{-\varepsilon_n K s} - 1}{\varepsilon_n} = -Ks\, e^{-\zeta_n K s} \tag{5.22}
$$

where ζ_n belongs to $(0, \varepsilon_n)$ for $n \geq 1$. Using then the same arguments as in (5.17) above, we see that the inequality (5.22) yields the fact that the liminf is non-negative so that (5.17) holds in the general case when $\Lambda_t \neq 0$ for $t \geq 0$ as well.

4. Similarly, arguing as in (5.18) we see that the right-hand side of the first inequality in (5.19) reads as follows

$$
\begin{aligned}
&G(t_n, x_n) - G(t_n{-}\varepsilon_n, x_n) \\
&\quad + \mathsf{E}\Big[\int_0^{\tau_n} \Big(e^{-\Lambda_s^{t_n,x_n}}\, \tilde{H}(t_n{+}s, X_s^{x_n}) - e^{-\Lambda_s^{t_n-\varepsilon_n,x_n}}\, \tilde{H}(t_n{-}\varepsilon_n{+}s, X_s^{x_n}) \Big)\, ds \Big] \\
&= G(t_n, x_n) - G(t_n{-}\varepsilon_n, x_n) \\
&\quad + \mathsf{E}\Big[\int_0^{\tau_n} e^{-\Lambda_s^{t_n,x_n}} \Big(1 - e^{\Lambda_s^{t_n,x_n} - \Lambda_s^{t_n-\varepsilon_n,x_n}} \Big) \tilde{H}(t_n{+}s, X_s^{x_n})\, ds \Big] \\
&\quad + \mathsf{E}\Big[\int_0^{\tau_n} \Big(e^{-\Lambda_s^{t_n-\varepsilon_n,x_n}} \Big(\tilde{H}(t_n{+}s, X_s^{x_n}) - \tilde{H}(t_n{-}\varepsilon_n{+}s, X_s^{x_n}) \Big)\, ds \Big]
\end{aligned} \tag{5.23}
$$

for $n \geq 1$.

The second expectation on the right-hand side of (5.23) can be handled in exactly the same way as the corresponding expectation in (5.19), and this yields the conclusion of (5.20), provided that the limsup of the first expectation on the right-hand side of (5.23) divided by ε_n is non-positive as $n \to \infty$. To see that the limsup is non-positive, note that (5.11) and the mean value theorem imply that

$$
\frac{1 - e^{\Lambda_s^{t_n,x_n} - \Lambda_s^{t_n-\varepsilon_n,x_n}}}{\varepsilon_n} \leq \frac{1 - e^{-\varepsilon_n K s}}{\varepsilon_n} = Ks\, e^{-\zeta_n K s} \tag{5.24}
$$

where ζ_n belongs to $(0, \varepsilon_n)$ for $n \geq 1$. Using then the same arguments as in (5.20) above, we see that the inequality (5.24) yields the fact that the limsup is non-positive so that (5.20) holds in the general case when $\Lambda_t \neq 0$ for $t \geq 0$ as well. Combining the conclusions of (5.17) and (5.20) we see that $\lim_{n\to\infty} V_t(t_n, x_n) = G_t(z)$ and this completes the proof. $\square$

Remark 16. Note that the comments on the sufficient conditions from Theorem 8 made in Remark 9 above extend to the corresponding sufficient conditions in Theorem 15 and we will omit further details in this direction. Note also that the proof of (5.20) above could also be accomplished by means of the mean value theorem (as in the proof of Theorems 8 and 10) without appeal to the identity (5.12).

3. The result and proof of Theorem 13 and Theorem 15 extend to the case when the gain function G in the optimal stopping problem (2.1)/(2.2) is *not smooth* away from the optimal stopping boundary ∂C. Instead of formulating a general theorem of this kind, which would be overly technical and rather difficult to read, we will illustrate key arguments of such extensions through an important example that was already considered in Example 12 above for the space derivative. The method of proof to be presented below is different from the method of proof applied in Example 12 above.

Example 17 (Continuity of the time derivative in the American put). Consider the optimal stopping problem (4.28) above where X solves (4.29). Horizon in the optimal stopping problem (4.28) is finite so that the setting belongs to Theorem 15 above. Since the gain function $G(x) = (K-x)^+$ for $x > 0$ is not differentiable at K we see that the condition (5.10) fails and hence we cannot conclude that

$$V_t \text{ is continuous on } \partial C \tag{5.25}$$

using Theorem 15 (we write V_t to denote $\partial V/\partial t$ throughout). We will now show however that the method of proof of Theorem 15 extends to cover the case of the non-differentiable gain function $G(x) = (K-x)^+$ for $x > 0$. This will also serve as an illustration of how similar other cases of non-smooth gain functions G in the optimal stopping problem (2.1)/(2.2) can be handled. The derivation of (5.25) will be divided in three steps as follows.

1. We first recall the facts about the optimal stopping problem (4.28) stated in the first step of the proof of (4.30) above. In particular, taking any sequence $(t_n, x_n) \in C$ converging to $z = (t, b(t)) \in \partial C$ we know that $\tau_D^{t_n,x_n} \to 0$ almost surely as $n \to \infty$. Moreover, applying the Itô-Tanaka formula, we find using (4.29) that

$$\begin{aligned} e^{-rt}(K-X_t)^+ = (K-x)^+ - \int_0^t re^{-rs} K\, I(X_s<K)\, ds - \int_0^t e^{-rs} \sigma X_s\, I(X_s<K)\, dB_s \\ + \int_0^t \tfrac{1}{2} e^{-rs}\, d\ell_s^K(X) \end{aligned} \tag{5.26}$$

for $t \in [0,T]$ where $\ell^K(X)$ is the local time process of X defined by

$$\ell_t^K(X) = \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_0^t I(K-\varepsilon < X_s < K+\varepsilon)\, d\langle X, X\rangle_s \tag{5.27}$$

where the convergence takes place in probability and the quadratic variation process $\langle X, X\rangle$ of X is given by $\langle X, X\rangle_t = \int_0^t \sigma^2 X_s^2\, ds$ for $t \in [0,T]$. It is easily verified that the third term on the right-hand side in (5.26) defines a continuous martingale for $t \in [0,T]$. Hence by the optional sampling theorem we find that the Bolza formulated optimal stopping problem (4.28) can be Lagrange reformulated (see [41, p. 141] for the terminology) as follows

$$
\begin{aligned}
\tilde{V}(t,x) &:= V(t,x) - (K-x)^+ \\
&= \sup_{0 \le \tau \le T-t} \mathsf{E}\Big[\int_0^\tau \tfrac{1}{2} e^{-rs}\, d\ell_s^K(X^x) - \int_0^\tau r e^{-rs} K I(X_s^x < K)\, ds \Big]
\end{aligned}
\tag{5.28}
$$

for $t \in [0,T]$ and $x > 0$. Thus the optimal stopping problems (4.28) and (5.28) are equivalent and a stopping time is optimal in (4.28) if and only if it is optimal in (5.28).

2. We next connect to the first part of the proof of Theorem 15. Passing to a subsequence of $((t_n,x_n))_{n \ge 1}$ if needed there is no loss of generality in assuming that

$$
\liminf_{n\to\infty} V_t(t_n,x_n) = \lim_{n\to\infty} \frac{V(t_n{+}\varepsilon_n,x_n) - V(t_n,x_n)}{\varepsilon_n}
\tag{5.29}
$$

for some $\varepsilon_n \downarrow 0$ as $n \to \infty$. Let $\tau_n := \tau_D^{t_n,x_n}$ be the optimal stopping time for $V(t_n,x_n)$ and thus $\tilde{V}(t_n,x_n)$ as well. Set $\hat{\tau}_n := \tau_n \wedge (T{-}t_n{-}\varepsilon_n)$ for $n \ge 1$. We then have

$$
\begin{aligned}
V(t_n{+}\varepsilon_n,x_n) - V(t_n,x_n) &= \tilde{V}(t_n{+}\varepsilon_n,x_n) - \tilde{V}(t_n,x_n) \\
&\ge \mathsf{E}\Big[\int_0^{\hat{\tau}_n} \tfrac{1}{2} e^{-rs}\, d\ell_s^K(X^{x_n}) - \int_0^{\hat{\tau}_n} r e^{-rs} K I(X_s^{x_n} < K)\, ds \Big] \\
&\quad - \mathsf{E}\Big[\int_0^{\tau_n} \tfrac{1}{2} e^{-rs}\, d\ell_s^K(X^{x_n}) - \int_0^{\tau_n} r e^{-rs} K I(X_s^{x_n} < K)\, ds \Big] \\
&\ge -\mathsf{E}\Big[\int_{\hat{\tau}_n}^{\tau_n} \tfrac{1}{2} e^{-rs}\, d\ell_s^K(X^{x_n})\, I(T{-}t_n{-}\varepsilon_n < \tau_n \le T{-}t_n) \Big] \\
&\ge -\frac{1}{2}\, e^{-r(T-t_n-\varepsilon_n)}\, \mathsf{E}\big[\ell_{T-t_n}^K(X^{x_n}) - \ell_{T-t_n-\varepsilon_n}^K(X^{x_n}) \big]
\end{aligned}
\tag{5.30}
$$

for all $n \ge 1$. By (5.27) and Fatou's lemma we find that

$$
\begin{aligned}
&\mathsf{E}\big[\ell_{T-t_n}^K(X^{x_n}) - \ell_{T-t_n-\varepsilon_n}^K(X^{x_n}) \big] \\
&\quad = \mathsf{E}\Big[\lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_{T-t_n-\varepsilon_n}^{T-t_n} I(K{-}\varepsilon < X_s^{x_n} < K{+}\varepsilon)\, \sigma^2 (X_s^{x_n})^2\, ds \Big] \\
&\quad \le \sigma^2 x_n^2 \liminf_{\varepsilon \downarrow 0} \int_{T-t_n-\varepsilon_n}^{T-t_n} \frac{1}{2\varepsilon}\, \mathsf{E}\big[I\big(\tfrac{K-\varepsilon}{x_n} < X_s^1 < \tfrac{K+\varepsilon}{x_n}\big) (X_s^1)^2 \big]\, ds \\
&\quad = \sigma^2 x_n^2 \liminf_{\varepsilon \downarrow 0} \int_{T-t_n-\varepsilon_n}^{T-t_n} \Big(\frac{1}{2\varepsilon} \int_{\frac{K-\varepsilon}{x_n}}^{\frac{K+\varepsilon}{x_n}} x^2 f_{X_s^1}(x)\, dx \Big)\, ds \\
&\quad = \sigma^2 K^2 \int_{T-t_n-\varepsilon_n}^{T-t_n} f_{X_s^1}\big(\tfrac{K}{x_n}\big)\, ds
\end{aligned}
\tag{5.31}
$$

for all $n \ge 1$ where $f_{X_s^1}$ denotes the density function of X_s^1 for $s > 0$ and in the last equality we use the dominated convergence theorem. Using the scaling property $B_s \sim \sqrt{s}\, B_1$

it is easily verified that $f_{X^1_s}$ is given by

$$f_{X^1_s}(x) = \frac{1}{\sigma x\sqrt{s}}\,\varphi\Big(\frac{\log(x)-(r-\sigma^2/2)s}{\sigma\sqrt{s}}\Big) \tag{5.32}$$

for $x>0$ and $s>0$ where φ denotes the standard normal density function given by $\varphi(x)=(1/\sqrt{2\pi})\,e^{-x^2/2}$ for $x\in I\!R$. Inserting (5.32) into (5.31) we find that

$$\mathsf{E}\big[\ell^K_{T-t_n}(X^{x_n})-\ell^K_{T-t_n-\varepsilon_n}(X^{x_n})\big] \le c\,\varepsilon_n \tag{5.33}$$

for all $n\ge n_0$ with some $n_0\ge 1$ large enough, where the constant $c=c(T-t)$ is given by

$$c=\sigma K^2\,\sup\,\frac{e^{-y}}{\sqrt{s}}\,\varphi\Big(\frac{y-(r-\sigma^2/2)s}{\sigma\sqrt{s}}\Big) \tag{5.34}$$

with the supremum being taken over all $s\in[(T-t)/2,2(T-t)]$ and $y\in I\!R$ (upon substituting $y=\log(x)$ in (5.32) above). Making use of (5.33) in (5.31) we obtain

$$V(t_n+\varepsilon_n,x_n)-V(t_n,x_n)\ge -c\,\varepsilon_n \tag{5.35}$$

for all $n\ge n_0$. Note that we can formally replace x_n in (5.35) by x because the constant c depends only on $T-t>0$ and the resulting inequality holds uniformly over all $x>0$.

Having (5.35) we modify the optimal stopping time τ_n by setting $\tau_n^\delta:=\tau_n\wedge\delta$ where $\delta>0$ is any (small) number such that $t_n+\varepsilon_n+\delta\le T$ for all $n\ge n_1$ where $n_1\ge n_0$ is sufficiently large. (Note that this is possible since $t<T$ with $t_n\to t$ and $\varepsilon_n\downarrow 0$ as $n\to\infty$.) Since $(t,x)\mapsto e^{-rt}V(t,x)$ is superharmonic on $[0,T]\times(0,\infty)$ and harmonic on C, we find that

$$\begin{aligned}(5.36)\qquad V(t_n+\varepsilon_n,x_n)-V(t_n,x_n) &\ge \mathsf{E}\Big[e^{-r\tau_n^\delta}\Big(V(t_n+\varepsilon_n+\tau_n^\delta,X^{x_n}_{\tau_n^\delta})-V(t_n+\tau_n^\delta,X^{x_n}_{\tau_n^\delta})\Big)\Big]\\ &= \mathsf{E}\Big[e^{-r\tau_n}\Big(V(t_n+\varepsilon_n+\tau_n,X^{x_n}_{\tau_n})-(K-X^{x_n}_{\tau_n})^+\Big)I(\tau_n\le\delta)\Big]\\ &\quad+\mathsf{E}\Big[e^{-r\delta}\Big(V(t_n+\varepsilon_n+\delta,X^{x_n}_\delta)-V(t_n+\delta,X^{x_n}_\delta)\Big)I(\tau_n>\delta)\Big]\\ &\ge -c\,\varepsilon_n\mathsf{P}(\tau_n>\delta)\end{aligned}$$

for all $n\ge n_1$ where in the final inequality we use (5.35) applied to $(t_n+\delta,x)$ in place of (t_n,x_n) for $n\ge 1$ and holding uniformly over all $x>0$. Dividing both sides in (5.36) by ε_n we find from (5.29) that

$$\liminf_{n\to\infty}\,V_t(t_n,x_n)\ge 0=G_t(z) \tag{5.37}$$

where we use that $\tau_n\to 0$ almost surely so that $\mathsf{P}(\tau_n>\delta)\to 0$ as $n\to\infty$.

3. We finally connect to the second part of the proof of Theorem 15. Similarly, there is no loss of generality in assuming that

$$\limsup_{n\to\infty}\,V_t(t_n,x_n)=\lim_{n\to\infty}\frac{V(t_n,x_n)-V(t_n-\varepsilon_n,x_n)}{\varepsilon_n} \tag{5.38}$$

for some $\varepsilon_n \downarrow 0$ as $n \to \infty$. We then have

$$
\begin{aligned}
(5.39) \qquad V(t_n, x_n) - V(t_n - \varepsilon_n, x_n) &= \tilde{V}(t_n, x_n) - \tilde{V}(t_n - \varepsilon_n, x_n) \\
&\le \mathsf{E}\Big[\int_0^{\tau_n} \tfrac{1}{2} e^{-rs}\, d\ell_s^K(X^{x_n}) - \int_0^{\tau_n} r e^{-rs} K I(X_s^{x_n} < K)\, ds \Big] \\
&\quad - \mathsf{E}\Big[\int_0^{\tau_n} \tfrac{1}{2} e^{-rs}\, d\ell_s^K(X^{x_n}) - \int_0^{\tau_n} r e^{-rs} K I(X_s^{x_n} < K)\, ds \Big] \\
&= 0
\end{aligned}
$$

for $n \ge 1$. Note that this inequality also follows from (4.28) from where we see directly that $t \mapsto V(t,x)$ is decreasing on $[0,T]$ for $x > 0$. Dividing both sides in (5.39) by ε_n we find from (5.38) and (5.39) that

$$
(5.40) \qquad \limsup_{n\to\infty} V_t(t_n, x_n) \le 0 = G_t(z)\, .
$$

Combining (5.37) and (5.40) we see that $\lim_{n\to\infty} V_t(t_n, x_n) = 0 = G_t(z)$ so that (5.25) holds as claimed and the proof is complete. □

Remark 18. Note that the method of proof presented in Example 17 first derives Lipschitz continuity of $t \mapsto V(t,x)$ uniformly over all x and then 'lifts' this continuity to C^1 regularity of $t \mapsto V(t,x)$ at $z \in \partial C$ using the superharmonic property of $(t,x) \mapsto e^{-rt}V(t,x)$ on $[0,T] \times (0,\infty)$. To our knowledge this 'lifting' method is applied in Example 17 for the first time in the literature. In addition to yielding the first known probabilistic proof of (5.25) in the American put problem, it is also clear from the arguments used in Example 17 that the 'lifting' method is applicable to a large class of diffusion/Markov processes in optimal stopping and free boundary problems with non-smooth gain functions.

Acknowledgements. T. De Angelis gratefully acknowledges partial support by EPSRC Grant EP/R021201/1 while working on the paper.

References

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

[2] Blanchet, A. Dolbeault, J. *and* Monneau, R. (2006). On the continuity of the time derivative of the solution to the parabolic obstacle problem with variable coefficients. *J. Math. Pures Appl. (9)* 85 (371–414).

[3] Blumenthal, R. M. *and* Getoor, R. K. (1968). *Markov Processes and Potential Theory.* Academic Press.

[4] Caffarelli, L. A. (1977). The regularity of free boundaries in higher dimensions. *Acta Math.* 139 (155–184).

[5] Caffarelli, L. A. *and* Salsa, S. (2005). *A Geometric Approach to Free Boundary Problems.* Graduate Studies in Mathematics 68, American Mathematical Society.

[6] Chentsov, N. N. (1956). Weak convergence of stochastic processes whose trajectories have no discontinuities of the second kind and the "heuristic" approach to the Kolmogorov-Smirnov tests. *Theory Probab. Appl.* 1 (140–144).

[7] Cox, A. M. G. *and* Peskir, G. (2015). Embedding laws in diffusions by functions of time. *Ann. Probab.* 43 (2481–2510).

[8] Cramér, H. (1966). On stochastic processes whose trajectories have no discontinuities of the second kind. *Ann. Mat. Pura Appl.* 71 (85–92).

[9] Cramér, H. *and* Leadbetter, M. R. (1967). *Stationary and Related Stochastic Processes. Sample Function Properties and Their Applications.* John Wiley & Sons.

[10] De Angelis, T. (2018). From optimal stopping boundaries to Rost's reversed barriers and the Skorokhod embedding. *Ann. Inst. Henri Poincaré Probab. Statist.* 54 (1098–1133).

[11] De Angelis, T. *and* Ekström, E. (2017). The dividend problem with a finite horizon. *Ann. Appl. Probab.* 27 (3525–3546).

[12] De Angelis, T. Gensbittel, F. *and* Villeneuve, S. (2017). A Dynkin game on assets with incomplete information on the return. To appear in *Math. Oper. Res.*

[13] Dembinski, V. (1984). Regular points, transversal sets, and Dirichlet problem for standard processes. *Z. Wahrsch. Verw. Gebiete* 66 (507–527).

[14] Doob, J. L. (1954). Semimartingales and subharmonic functions. *Trans. Amer. Math. Soc.* 77 (86–121).

[15] Dynkin, E. B. (1965). *Markov Processes.* Vol I. Springer.

[16] Dynkin, E. B. (1965). *Markov Processes.* Vol II. Springer.

[17] Effros, E. G. *and* Kazdan, J. L. (1970). Applications of Choquet simplexes to elliptic and parabolic boundary value problems. *J. Differential Equations* 8 (95–134).

[18] Evans, L. C. *and* Gariepy, R. F. (1982). Wiener's criterion for the heat equation. *Arch. Rational Mech. Anal.* 78 (293–314).

[19] Friedman, A. (1964). *Partial Differential Equations of Parabolic Type.* Prentice-Hall.

[20] Friedman, A. (1975). Parabolic variational inequalities in one space dimension and smoothness of the free boundary. *J. Funct. Anal.* 18 (151–176).

[21] Friedman, A. (1982). *Variational Principles and Free-Boundary Problems.* John Wiley & Sons.

[22] Gevrey, M. (1913). Sur les équations aux dérivées partielles du type parabolique. *J. Math. Pures Appl.* 9 (305-476).

[23] Gilbarg, D. *and* Trudinger, N. S. (2001). *Elliptic Partial Differential Equations of the Second Order.* Springer.

[24] Girsanov, I. V. (1960). Strong Feller processes. I. General properties. *Theory Probab. Appl.* 5 (5–24).

[25] Hawkes, J. (1979). Potential theory of Lévy processes. *Proc. London Math. Soc.* 38 (335–352).

[26] Itô, K. *and* McKean, H. P. Jr. (1965). *Diffuson Processes and their Sample Paths.* Springer.

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

[28] Karatzas, I. *and* Shreve, S. E. (1991). *Brownian Motion and Stochastic Calculus.* Springer.

[29] Kellogg, O. D. (1912). Harmonic functions and Green's integral. *Trans. Amer. Math. Soc.* 13 (109–132).

[30] Kellogg, O. D. (1931). On the derivatives of harmonic functions on the boundary. *Trans. Amer. Math. Soc.* 33 (486–510).

[31] Khoshnevisan, D. (2002). *Multiparameter Processes. An Introduction to Random Fields.* Springer.

[32] König, M. (2011). On J. Schauder's method to solve elliptic differential equations. *J. Fixed Point Theory Appl.* 9 (135–196).

[33] Krylov, N. V. (1966). On regular boundary points for Markov processes. Theory Probab. Appl. 11 (609–614).

[34] Lanconelli, E. (1973). Sul problema di Dirichlet per l'equazione del calore. *Ann. Mat. Pura Appl.* 97 (83–114).

[35] Landis, E. M. (1969). Necessary and sufficient conditions for regularity of a boundary point in the Dirichlet problem for the heat-conduction equation. *Soviet Math. Dokl.* 10 (380–384).

[36] Øksendal, B. (1990). The high contact principle in optimal stopping and stochastic waves. *Sem. Stoch. Process.* (San Diego 1989), *Progr. Probab.* 18, Birkhäuser (177–192).

[37] Peskir, G. (1998). Optimal stopping of the maximum process: The maximality principle. *Ann. Probab.* 26 (1614–1640).

[38] Peskir, G. (2005). A change-of-variable formula with local time on curves. *J. Theoret. Probab.* 18 (499–535).

[39] Peskir, G. (2007). A change-of-variable formula with local time on surfaces. *Sém. de Probab.* XL, *Lecture Notes in Math.* 1899, Springer (69–96).

[40] Peskir, G. (2019). Continuity of the optimal stopping boundary for two-dimensional diffusions. *Ann. Appl. Probab.* 29 (505–530).

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

[42] Petrowsky, I. (1935). Zur ersten Randwertaufgabe der Wärmeleitungsgleichung. *Compos. Math.* 1 (383–419).

[43] Poincaré, H. (1899). *Théorie du Potential Newtonien.* Paris, Carré & Naud.

[44] Potthoff, J. (2010). Sample properties of random fields III: Differentiability. *Commun. Stoch. Anal.* 4 (335–353).

[45] Protter, P. E. (2005). *Stochastic Integration and Differential Equations.* Springer.

[46] Revuz, D. *and* Yor, M. (1999). *Continuous Martingales and Brownian Motion.* Springer.

[47] Rogers, L. C. G. *and* Williams, D. (2000). *Diffusions, Markov Processes and Martingales: Itô Calculus* (Vol 2). Cambridge University Press.

[48] Schauder, J. (1934). Über lineare elliptische Differentialgleichungen zweiter Ordnung. *Math. Z.* 38 (257–282).

[49] Shiryaev, A. N. (1978). *Optimal Stopping Rules.* Springer.

[50] Slutsky, E. (1937). Alcune proposizioni sulla teoria della funzioni aleatorie. *Giorn. Ist. Ital. Attuari* 8 (193–199).

[51] Watson, N. A. (2014). Regularity of boundary points in the Dirichlet problem for the heat equation. *Bull. Aust. Math. Soc.* 90 (476–485).

[52] Wiener, N. (1924). The Dirichlet problem. *J. Math. Phys.* 3 (127–146).

[53] Zaremba, S. (1909). Sur le principe du minimum. *Bull. Int. Acad. Sci. Cracovie* (197–264).

Tiziano De Angelis
School of Mathematics
University of Leeds
Leeds LS2 9JT
United Kingdom
`t.deangelis@leeds.ac.uk`

Goran Peskir
Department of Mathematics
The University of Manchester
Oxford Road
Manchester M13 9PL
United Kingdom
`goran@maths.man.ac.uk`