ON THE COMPENSATOR IN THE DOOB-MEYER DECOMPOSITION OF THE SNELL ENVELOPE*

SAUL D. JACKA† AND DOMINYKAS NORGILAS†

Abstract. Let G be a semimartingale, and S its Snell envelope. Under the assumption that $G \in \mathcal{H}^1$, we show that the finite-variation part of S is absolutely continuous with respect to the decreasing part of the finite-variation part of G. In the Markovian setting, this enables us to identify sufficient conditions for the value function of the optimal stopping problem to belong to the domain of the extended (martingale) generator of the underlying Markov process. We then show that the *dual* of the optimal stopping problem is a stochastic control problem for a controlled Markov process, and the optimal control is characterised by a function belonging to the domain of the martingale generator. Finally, we give an application to the smooth pasting condition.

Key words. Doob-Meyer decomposition, optimal stopping, Snell envelope, stochastic control, martingale duality, smooth pasting

AMS subject classifications. 60G40, 60G44, 60J25, 60G07, 93E20

1. Introduction. Given a (gains) process $G = (G_t)_{t\geq 0}$, living on the usual filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, the classical optimal stopping problem is to find a maximal reward $v_0 = \sup_{\tau\geq 0} \mathbb{E}[G_\tau]$, where the supremum is taken over all $\mathbb{F}$ - stopping times. In order to compute v_0, we consider, the value process $v_t = \operatorname{ess\,sup}_{\tau\geq t} \mathbb{E}[G_\tau|\mathcal{F}_\sigma]$, $t \geq 0$. It is, or should be, well-known (see, for example, El Karoui [16], Karatzas and Shreve [31]) that under suitable integrability and regularity conditions on the process G, the Snell envelope of G, denoted by $S = (S_t)_{t\geq 0}$, is the minimal supermartingale which dominates G and aggregates the value process at each $\mathbb{F}$-stopping time $\sigma \geq 0$, so that $S_\sigma = v_\sigma$ almost surely. Moreover, $\tau_\sigma := \inf\{r \geq \sigma : S_r = G_r\}$ is the minimal optimal stopping time, so, in particular, $S_\sigma = v_\sigma = \mathbb{E}[G_{\tau_\sigma}|\mathcal{F}_\sigma]$ almost surely. A successful construction of the process S leads, therefore, to the solution of the initial optimal stopping problem.

In the Markovian setting the gains process takes the form $G = g(X)$, where $g(\cdot)$ is some payoff function applied to an underlying Markov process X. Under very general conditions, the Snell envelope is then characterised as the least super-mean-valued function $V(\cdot)$ that majorizes $g(\cdot)$. A standard technique to find the value function $V(\cdot)$ is to solve the corresponding obstacle (free-boundary) problem. For an exposition of the general theory of optimal stopping in both settings we also refer to Peskir and Shiryaev [39].

The main aim of this paper is to answer the following canonical question of interest:

Question. *When does the value function $V(\cdot)$ belong to the domain of the extended (martingale) generator of the underlying Markov process X?*

Very surprisingly, given how long general optimal stopping problems have been studied (see Snell [49]), we have been unable to find any general results about this.

As the title suggests, we tackle the question by considering the optimal stopping

*Submitted to the editors 14/04/2018.

Funding: This work was funded by the EPSRC under the grant EP/P00377X/1, the Alan Turing Institute under the EPSRC grant EP/N510129/1 and the EPSRC Doctoral Training Partnerships grant EP/M508184/1.

†Department of Statistics, University of Warwick, Coventry CV4 7AL, United Kingdom (s.d.jacka@warwick.ac.uk, d.norgilas@warwick.ac.uk).

problem in a more general (semimartingale) setting first. If a gains process G is sufficiently integrable, then S is of class (D) and thus uniquely decomposes into the difference of a uniformly integrable martingale, say M, and a predictable, increasing process, say A, of integrable variation. From the general theory of optimal stopping it can be shown that $\bar{\tau}_\sigma := \inf\{r \geq \sigma : A_r > 0\}$ is the maximal optimal stopping time, while the stopped process $S^{\bar{\tau}_\sigma} = (S_{t\wedge\bar{\tau}_\sigma})_{t\geq 0}$ is a martingale. Now suppose that G is a semimartingale itself. Then its finite variation part can be further decomposed into the sum of increasing and decreasing processes that are, as random measures, mutually singular. Off the support of the decreasing one, G is (locally) a submartingale, and thus in this case it is suboptimal to stop, and we again expect S to be (locally) a martingale. This also suggests that A increases only if the decreasing component of the finite variation part of G decreases. In particular, we prove the following fundamental result (see Theorem 3.6):

the finite-variation process in the Doob-Meyer decomposition of S is absolutely continuous with respect to the decreasing part of the corresponding finite-variation process in the decomposition of G.

This being a very natural conjecture, it is not surprising that some variants of it have already been considered. As a helpful referee pointed out to us, several versions of Theorem 3.6 were established in the literature on reflected BSDEs under various assumptions on the gains process, see El Karoui et. al. [17] (G is a continuous semimartingale), Crepéy and Matoussi [9] (G is a càdlàg quasi-martingale), Hamadéne and Ouknine [23] (G is a limiting process of a sequence of sufficiently regular semimartingales). We note that these results (except Hamadéne and Ouknine [23], where the assumed regularity of G is exploited) are proved essentially by using (or appropriately extending) the related (but different) result established in Jacka [27]. There, under the assumption that S and G are both continuous and sufficiently integrable semimartingales, the author shows that a local time of $S - G$ at zero is absolutely continuous with respect to the decreasing part of the finite-variation process in the decomposition of G. Our proof of Theorem 3.6 relies on the classical methods establishing the Doob-Meyer decomposition of a supermartingale.

The first part of section 3 is devoted to the groundwork necessary to establish Theorem 3.6. It turns out that an answer to the motivating question of this paper then follows naturally. In particular, in the second part of section 3, in Theorem 3.18, we show that, under very general assumptions on the underlying Markov process X, if the payoff function $g(\cdot)$ belongs to the domain of the martingale generator of X, so does the value function $V(\cdot)$ of the optimal stopping problem.

In section 4 we discuss some applications. First, we consider a dual approach to optimal stopping problems due to Davis and Karatzas [10] (see also Rogers [43], and Haugh and Kogan [24]). In particular, from the absolute continuity result announced above, it follows that the dual is a stochastic control problem *for a controlled Markov process*, which opens the doors to the application of all the available theory related to such problems (see Fleming and Soner [19]). Secondly, if the value function of the optimal stoping problem belongs to the domain of the martingale generator, under a few additional (but general) assumptions, we also show that the celebrated *smooth fit* principle holds for (killed) one-dimensional diffusions.

2. Preliminaries.

2.1. General framework. Fix a time horizon $T \in (0, \infty]$. Let G be an adapted, càdlàg gains process on $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{0 \le t \le T}, \mathbb{P})$, where $\mathbb{F}$ is a right-continuous and complete filtration. We suppose that $\mathcal{F}_0$ is trivial. In the case $T = \infty$, we interpret $\mathcal{F}_\infty = \sigma\Big(\cup_{0 \le t < \infty} \mathcal{F}_t \Big)$ and $G_\infty = \liminf_{t \to \infty} G_t$. For two $\mathbb{F}$-stopping times σ_1, σ_1 with $\sigma_1 \le \sigma_2$ $\mathbb{P}$-a.s., by $\mathcal{T}_{\sigma_1, \sigma_2}$ we denote the set of all $\mathbb{F}$-stopping times τ such that $\mathbb{P}(\sigma_1 \le \tau \le \sigma_2) = 1$. We will assume that the following condition is satisfied:

$$\mathbb{E}\Big[\sup_{0 \le t \le T} |G_t| \Big] < \infty, \tag{2.1}$$

and let

$\bar{\mathbb{G}}$ be the space of all adapted, càdlàg processes such that (2.1) holds.

The *optimal stopping problem* is to compute the maximal expected reward

$$v_0 := \sup_{\tau \in \mathcal{T}_{0,T}} \mathbb{E}[G_\tau].$$

Remark 2.1. First note that by (2.1), $\mathbb{E}[G_\tau] < \infty$ for all $\tau \in \mathcal{T}_{0,T}$, and thus v_0 is finite. Moreover, most of the general results regarding optimal stopping problems are proved under the assumption that G is a non-negative (hence the *gains*) process. However, under (2.1), $N = (N_t)_{0 \le t \le T}$ given by $N_t = \mathbb{E}[\sup_{0 \le s \le T} |G_s| | \mathcal{F}_t]$ is a uniformly integrable martingale, while $\hat{G} := N + G$ defines a non-negative process (even if G is allowed to take negative values). Then

$$\hat{v}_0 := \sup_{\tau \in \mathcal{T}_{0,T}} \mathbb{E}[N_\tau + G_\tau] = \mathbb{E}\Big[\sup_{0 \le t \le T} |G_t| \Big] + \sup_{\tau \in \mathcal{T}_{0,T}} \mathbb{E}[G_\tau],$$

and finding $\hat{v}_0$ is the same as finding v_0. Hence we may, and shall, assume without loss of generality that $G \ge 0$.

The key to our study is provided by the family $\{v_\sigma\}_{\sigma \in \mathcal{T}_{0,T}}$ of random variables

$$v_\sigma := \operatorname*{ess\,sup}_{\tau \in \mathcal{T}_{\sigma,T}} \mathbb{E}[G_\tau | \mathcal{F}_\sigma], \quad \sigma \in \mathcal{T}_{0,T}. \tag{2.2}$$

Note that, since each deterministic time $t \in [0, T]$ is also a stopping time, (2.2) defines an adapted *value* process $(v_t)_{0 \le t \le T}$. For $\sigma \in \mathcal{T}_{0,T}$, it is tempting to regard v_σ as the process $(v_t)_{0 \le t \le T}$ evaluated at the stopping time σ. It turns out that there is indeed a modification $(S_t)_{0 \le t \le T}$ of the process $(v_t)_{0 \le t \le T}$ that aggregates the family $\{v_\sigma\}_{\sigma \in \mathcal{T}_{0,T}}$ at each stopping time σ (see Theorem D.7 in Karatzas and Shreve [31]). This process S is the Snell envelope of G.

THEOREM 2.2 (Characterisation of S). *Let $G \in \bar{\mathbb{G}}$. The Snell envelope process S of G satisfies*

$$S_\sigma = \operatorname*{ess\,sup}_{\tau \in \mathcal{T}_{\sigma,T}} \mathbb{E}[G_\tau | \mathcal{F}_\sigma], \quad \mathbb{P} - a.s., \sigma \in \mathcal{T}_{0,T}. \tag{2.3}$$

Moreover, S is the minimal càdlàg supermartingale that dominates G.

For the proof of Theorem 2.2 under slightly more general assumptions on the gains process G consult Appendix I in Dellacherie and Meyer [12] or Proposition 2.26 in El Karoui [16].

If $G \in \bar{\mathbb{G}}$, it is clear that G is a uniformly integrable process. In particular, it is also of class (D), i.e. the family of random variables $\{G_\tau \mathbb{1}_{\{\tau<\infty\}} : \tau \text{ is a stopping time}\}$ is uniformly integrable. On the other hand, a right-continuous adapted process Z belongs to the class (D) if there exists a uniformly integrable martingale $\hat{N}$, such that, for all $t \in [0,T]$, $|Z_t| \leq \hat{N}_t$ $\mathbb{P}$-a.s. (see e.g. Dellacherie and Meyer [12], Appendix I and references therein). In our case, by (2.3) and using the conditional version of Jensen's inequality, for $t \in [0,T]$, we have

$$|S_t| \leq \mathbb{E}\Big[\sup_{0\leq s\leq T}|G_s|\Big|\mathcal{F}_t\Big] := N_t \quad \mathbb{P}\text{-a.s.}$$

But, since $G \in \bar{\mathbb{G}}$, N is a uniformly integrable martingale, which proves the following

Lemma 2.3. *Suppose $G \in \bar{\mathbb{G}}$. Then S is of class (D).*

Let $\mathcal{M}_0$ denote the set of right-continuous martingales started at zero. Let $\mathcal{M}_{0,loc}$ and $\mathcal{M}_{0,UI}$ denote the spaces of local and uniformly integrable martingales (started at zero), respectively. Similarly, the adapted processes of finite and integrable variation will be denoted by FV and IV, respectively.

It is well-known that a right-continuous (local) supermartingale P has a unique decomposition $P = B - I$ where $B \in \mathcal{M}_{0,loc}$ and I is an increasing (FV) process which is predictable. This can be regarded as the general Doob-Meyer decomposition of a supermartingale. Specialising to class (D) supermartingales we have a stronger result (this is a consequence of, for example, Protter [40] Theorem 16, p.116 and Theorem 11, p.112):

Theorem 2.4 (Doob-Meyer decomposition). *Let $G \in \bar{\mathbb{G}}$. Then the Snell envelope process S admits a unique decomposition*

$$S = M^* - A, \tag{2.4}$$

where $M^* \in \mathcal{M}_{0,UI}$, and A is a predictable, increasing IV process.

Remark 2.5. It is normal to assume that the process A in the Doob-Meyer decomposition of S is started at zero. The duality result alluded to in the introduction is one reason why we do not do so here.

An immediate consequence of Theorem 2.4 is that S is a semimartingale. In addition, we also assume that G is a semimartingale with the following decomposition:

$$G = N + D, \tag{2.5}$$

where $N \in \mathcal{M}_{0,loc}$ and D is a FV process. Unfortunately, the decomposition (2.5) is not, in general, unique. On the other hand, uniqueness *is* obtained by requiring the FV term to also be predictable, at the cost of restricting only to locally integrable processes. If there exists a decomposition of a semimartingale X with a predictable FV process, then we say that X is *special*. For a special semimartingale we always choose to work with its *canonical* decomposition (so that a FV process is predictable). Let

$$\mathbb{G} \text{ be the space of semimartingales in } \bar{\mathbb{G}}.$$

LEMMA 2.6. *Suppose $G \in \mathbb{G}$. Then G is a special semimartingale.*

See Theorems 36 and 37 (p.132) in Protter [40] for the proof.

The following lemma provides a further decomposition of a semimartingale (see Proposition 3.3 (p.27) in Jacod and Shiryaev [28]). In particular, the FV term of a special semimartingale can be uniquely (up to initial values) decomposed in a predictable way, into the difference of two increasing, mutually singular FV processes.

LEMMA 2.7. *Suppose that K is a càdlàg, adapted process such that $K \in FV$. Then there exists a unique pair (K^+, K^-) of adapted increasing processes such that $K - K_0 = K^+ - K^-$ and $\int |dK_s| = K^+ + K^-$. Moreover, if K is predictable, then K^+, K^- and $\int |dK_s|$ are also predictable.*

2.2. Markovian setting.

The Markov process. Let $(E, \mathcal{E})$ be a metrizable Lusin space endowed with the σ-field of Borel subsets of E. Let $X = (\Omega, \mathcal{G}, \mathcal{G}_t, X_t, \theta_t, \mathbb{P}_x : x \in E, t \in \mathbb{R}_+)$ be a Markov process taking values in $(E, \mathcal{E})$. We assume that a sample space Ω is such that the usual semi-group of shift operators $(\theta_t)_{t \geq 0}$ is well-defined (which is the case, for example, if $\Omega = E^{[0,\infty)}$ is the canonical path space). If the corresponding semigroup of X, (P_t), is the primary object of study, then we say that X is a realisation of a Markov semigroup (P_t). In the case of (P_t) being sub-Markovian, i.e. $P_t 1_E \leq 1_E$, we extend it to a Markovian semigroup over $E^\Delta = E \cup \{\Delta\}$, where Δ is a coffin-state. We also denote by $\mathcal{C}(X) = (\Omega, \mathcal{F}, \mathcal{F}_t, X_t, \theta_t, \mathbb{P}_x : x \in E, t \in \mathbb{R}_+)$ the *canonical realisation* associated with X, defined on Ω with the filtration $(\mathcal{F}_t)$ deduced from $\mathcal{F}_t^0 = \sigma(X_s : s \leq t)$ by standard regularisation procedures (completeness and right-continuity).

In this paper our standing assumption is that the underlying Markov process X is a *right process* (consult Getoor [20], Sharpe [46] for the general theory). Essentially, right processes are the processes satisfying Meyer's regularity hypotheses (*hypothèses droites*) HD1 and HD2. If a given Markov semigroup (P_t) satisfies HD1 and μ is an arbitrary probability measure on $(E, \mathcal{E})$, then there exists a homogeneous E-valued Markov process X with transition semigroup (P_t) and initial law μ. Moreover, a realisation of such (P_t) is right-continuous (Sharpe [46], Theorem 2.7). Under the second fundamental hypothesis, HD2, $t \to f(X_t)$ is right-continuous for every α-excessive function f. Recall, for $\alpha > 0$, a universally measurable function $f : E \to \mathbb{R}$ is α-super-median if $e^{-\alpha t} P_t f \leq f$ for all $t \geq 0$, and α-excessive if it is α-super-median and $e^{-\alpha t} P_t f \to f$ as $t \to 0$. If (P_t) satisfies HD1 and HD2 then the corresponding realisation X is strong Markov (Getoor [20], Theorem 9.4 and Blumenthal and Getoor [7], Theorem 8.11).

Remark 2.8. One has the following inclusions among classes of Markov processes:

$$\text{(Feller)} \subset \text{(Hunt)} \subset \text{(right)}$$

Let $\mathcal{L}$ be a given extended infinitesimal (martingale) generator of X with a domain $\mathbb{D}(\mathcal{L})$, i.e. we say a Borel function $f : E \to \mathbb{R}$ belongs to $\mathbb{D}(\mathcal{L})$ if there exists a Borel function $h : E \to \mathbb{R}$, such that $\int_0^t |h(X_s)| ds < \infty$, $\forall t \geq 0$, $\mathbb{P}_x$-a.s. for each x and the process $M^f = (M_t^f)_{t \geq 0}$, given by

$$M_t^f := f(X_t) - f(x) - \int_0^t h(X_s) ds, \quad t \geq 0,\ x \in E, \tag{2.6}$$

is a local martingale under each $\mathbb{P}_x$ (see Revuz and Yor [42] p.285), and then we write $h = \mathcal{L} f$.

Remark 2.9. Note that if $A \in \mathcal{E}$ and $\mathbb{P}_x(\lambda(\{t: \ X_t \in A\} = 0) = 1$ for each $x \in E$, where λ is Lebesgue measure, then h may be altered on A without affecting the validity of (2.6), so that, in general, the map $f \to h$ is not unique. This is why we refer to *a* martingale generator.

Optimal stopping problem. Let $X = (\Omega, \mathcal{G}, \mathcal{G}_t, X_t, \theta_t, \mathbb{P}_x : x \in E, t \in \mathbb{R}_+)$ be a right process. Given a function $g : E \to \mathbb{R}$ (with $g(\Delta) = 0$), $\alpha \geq 0$ and $T \in \mathbb{R}_+ \cup \{\infty\}$ define a corresponding gains process G^α (we simply write G if $\alpha = 0$) by $G^\alpha_t = e^{-\alpha t} g(X_t)$ for $t \in [0, T]$. In the case of $T = \infty$, we make a convention that $G^\alpha_\infty = \liminf_{t\to\infty} G^\alpha_t$. Let $\mathcal{E}^e, \mathcal{E}^u$ be the σ-algebras on E generated by excessive functions and universally measurable sets, respectively (recall that $\mathcal{E} \subset \mathcal{E}^e \subset \mathcal{E}^u$). We write

$$g \in \mathcal{Y}, \text{ given that } g(\cdot) \text{ is } \mathcal{E}^e\text{-measurable and } G^\alpha \text{ is of class (D).}$$

For a filtration $(\hat{\mathcal{G}}_t)$, and $(\hat{\mathcal{G}}_t)$ - stopping times σ_1 and σ_2, with $\mathbb{P}_x[0 \leq \sigma_1 \leq \sigma_2 \leq T] = 1$, $x \in E$, let $\mathcal{T}_{\sigma_1,\sigma_2}(\hat{\mathcal{G}})$ be the set of $(\hat{\mathcal{G}}_t)$ - stopping times τ with $\mathbb{P}_x[\sigma_1 \leq \tau \leq \sigma_2] = 1$. Consider the following optimal stopping problem:

$$V(x) = \sup_{\tau \in \mathcal{T}_{0,T}(\mathcal{G})} \mathbb{E}_x[e^{-\alpha\tau} g(X_\tau)], \quad x \in E.$$

By convention we set $V(\Delta) = g(\Delta)$. The following result is due to El Karoui et al. [18].

THEOREM 2.10. *Let $X = (\Omega, \mathcal{G}, \mathcal{G}_t, X_t, \theta_t, \mathbb{P}_x : x \in E, t \in \mathbb{R}_+)$ be a right process with canonical filtration $(\mathcal{F}_t)$. If $g \in \mathcal{Y}$, then*

$$V(x) = \sup_{\tau \in \mathcal{T}_{0,T}(\mathcal{F})} \mathbb{E}_x[e^{-\alpha\tau} g(X_\tau)], \quad x \in E,$$

and $(e^{-\alpha t} V(X_t))$ is a Snell envelope of G^α, i.e. for all $x \in E$ and $\tau \in \mathcal{T}_{0,T}(\mathcal{F})$

$$e^{-\alpha\tau} V(X_\tau) = \operatorname*{ess\,sup}_{\sigma \in \mathcal{T}_{\tau,T}(\mathcal{F})} \mathbb{E}_x[G^\alpha_\sigma | \mathcal{F}_\tau] \quad \mathbb{P}_x\text{-a.s.}$$

The first important consequence of the theorem is that we can (and will) work with the canonical realisation $\mathcal{C}(X)$. The second one provides a crucial link between the Snell envelope process in the general setting and the value function in the Markovian framework.

Remark 2.11. The restriction to gains processes of the form $G = g(X)$ (or G^α if $\alpha > 0$) is *much less restrictive than might appear*. Given that we work on the canonical path space with θ being the usual shift operator, we can expand the state-space of X by appending an adapted functional F, taking values in the space $(E', \mathcal{E}')$, with the property that

$$\text{(2.7)} \qquad \{F_{t+s} \in A\} \in \sigma(F_s) \cup \sigma(\theta_s \circ X_u : \ 0 \leq u \leq t), \quad \text{for all } A \in \mathcal{E}'.$$

This allows us to deal with time-dependent problems, running rewards and other path-functionals of the underlying Markov process.

LEMMA 2.12. *Suppose X is a canonical Markov process taking values in the space $(E, \mathcal{E})$ where E is a locally compact, countably based Hausdorff space and $\mathcal{E}$ is its Borel σ-algebra. Suppose also that F is a path functional of X satisfying (2.7) and taking values in the space $(E', \mathcal{E}')$ where E' is a locally compact, countably based Hausdorff* *space with Borel σ-algebra $\mathcal{E}'$, then, defining $Y = (X, F)$, Y is still Markovian. If X is a strong Markov process and F is right-continuous, then Y is strong Markov. If X is a Feller process and F is right-continuous , then Y is strong Markov, has a càdlàg modification and the completion of the natural filtration of X, $\mathbb{F}$, is right-continuous and quasi-left continuous, and thus Y is a right process.*

Example 2.13. If X is a one-dimensional Brownian motion, then Y, defined by

$$Y_t = \left(X_t, L_t^0, \sup_{0\le s\le t} X_s, \int_0^t \exp(-\int_0^s \alpha(X_u)du) f(X_s)ds \right), \quad t \ge 0,$$

where L^0 is the local time of X at 0, is a Feller process on the filtration of X.

3. Main results.

In this section we retain the notation of subsection 2.1 and subsection 2.2.

3.1. General framework.

The assumption that $G \in \mathbb{G}$ (i.e. G is a semimartingale with integrable supremum and $G = N + D$ is its canonical decomposition), neither ensures that $N \in \mathcal{M}_0$, nor that D is an IV process, the latter, it turns out, being sufficient for the main result of this section to hold. In order to prove Theorem 3.6 we will need a stronger integrability condition on G.

For any adapted càdlàg process H, define

$$H^* = \sup_{0\le t\le T} |H_t| \tag{3.1}$$

and

$$||H||_{\mathcal{S}^p} = ||H^*||_{L^p} := \mathbb{E}\big[|H^*|^p\big]^{1/p}, \quad 1 \le p \le \infty. \tag{3.2}$$

Remark 3.1. Note that $\bar{\mathbb{G}} = \mathcal{S}^1$, so that under the current conditions we have that $G \in \mathcal{S}^1$.

For a special semimartingale X with canonical decomposition $X = \bar{B} + \bar{I}$, where $\bar{B} \in \mathcal{M}_{0,loc}$ and $\bar{I}$ is a predictable FV process (with $I_0 = X_0$), define the $\mathcal{H}^p$ norm, for $1 \le p \le \infty$, by

$$||X||_{\mathcal{H}^p} = ||\bar{B}||_{\mathcal{S}^p} + \Big|\Big| \int_0^T |d\bar{I}_s| \Big|\Big|_{L^p} + ||I_0||_{L^p}, \tag{3.3}$$

and, as usual, write $X \in \mathcal{H}^p$ if $||X||_{\mathcal{H}^p} < \infty$.

Remark 3.2. A more standard definition of the $\mathcal{H}^p$ norm is with $||\bar{B}||_{\mathcal{S}^p}$ replaced by $||[\bar{B}, \bar{B}]_T^{1/2}||_{L^p}$. However, the Burkholder-Davis-Gundy inequalities (see Protter [40], Theorem 48 and references therein) imply the equivalence of these norms.

The following lemma follows from the fact that $\bar{I}^* \le \int_0^T |d\bar{I}_s| + |I_0|$, $\mathbb{P}$−a.s:

LEMMA 3.3. *On the space of special semimartingales, the $\mathcal{H}^p$ norm is stronger than $\mathcal{S}^p$ for $1 \le p < \infty$, i.e. convergence in $\mathcal{H}^p$ implies convergence in $\mathcal{S}^p$.*

In general, it is challenging to check whether a given process belongs to $\mathcal{H}^1$, and thus the assumption that $G \in \mathcal{H}^1$ might be too stringent. On the other hand, under the assumptions in the Markov setting (see subsection 3.2), we will have that G is *locally* in $\mathcal{H}^1$. Recall that a semimartingale X belongs to $\mathcal{H}^p_{loc}$, for $1 \le p \le \infty$, if there exists a sequence of stopping times $\{\sigma_n\}_{n\in\mathbb{N}}$, increasing to infinity almost surely, such that for each $n \geq 1$, the stopped process X^{σ_n} belongs to $\mathcal{H}^p$. Hence, the main assumption in this section is the following:

ASSUMPTION 3.4. *G is a semimartingale in both $\mathcal{S}^1$ and $\mathcal{H}^1_{loc}$.*

Remark 3.5. Given that $G \in \mathcal{H}^1$, Lemma 3.3 implies that Assumption 3.4 is satisfied, and thus all the results of subsection 2.1 hold. Moreover, we then have a canonical decomposition of G

$$G = N + D, \tag{3.4}$$

with $N \in \mathcal{M}_{0,UI}$ and a predictable IV process D. On the other hand, under Assumption 3.4, the decomposition (3.4) still holds, however, N and D are only locally uniformly integrable martingale (started at zero) and the process of integrable variation, respectively, i.e. $G^{\sigma_n} \in \mathcal{M}_{0,UI}$ and I^{σ_n} is a process of IV, where $\{\sigma_n\}_{n\geq 1}$ is a localising sequence.

We finally arrive to the main result of this section:

THEOREM 3.6. *Suppose Assumption 3.4 holds. Let A be a predictable, increasing IV process in the decomposition of the Snell envelope S, as in Theorem 2.4. Let D^- (D^+) denote the decreasing (increasing) components of D, as in Lemma 2.7. Then A is, as a measure, absolutely continuous with respect to D^- almost surely on $[0,T]$, and μ, defined by*

$$\mu_t := \frac{dA_t}{dD^-_t}, \quad 0 \leq t \leq T,$$

has a version that satisfies $0 \leq \mu_t \leq 1$ almost surely.

Remark 3.7. As is usual in semimartingale calculus, we treat a process of bounded variation and its corresponding Lebesgue-Stiltjes signed measure as synonymous.

The proof of Theorem 3.6 is based on the discrete-time approximation of the predictable FV processes in the decompositions of S (2.4) and G (2.5). In particular, let $\mathcal{P}_n = \{0 = t^n_0 < t^n_1 < t^n_2 < ... < t^n_{k_n} = T\}$, $n = 1, 2, ...$, be an increasing sequence of partitions of $[0,T]$ with $\max_{1\leq k\leq k_n} t^n_k - t^n_{k-1} \to 0$ as $n \to \infty$. Note that here $T < \infty$ is fixexd, but arbitrary. Let $S^n_t = S_{t^n_k}$ if $t^n_k \leq t < t^n_{k+1}$ and $S^n_T = S_T$ define the discretizations of S, and set

$$A^n_t = 0 \quad \text{if } 0 \leq t < t^n_1,$$

$$A^n_t = \sum_{j=1}^{k} \mathbb{E}[S_{t^n_{j-1}} - S_{t^n_j} | \mathcal{F}_{t^n_{j-1}}] \quad \text{if } t^n_k \leq t < t^n_{k+1},\ k = 1, 2, ..., k_n - 1,$$

$$A^n_T = \sum_{j=1}^{k_n} \mathbb{E}[S_{t^n_{j-1}} - S_{t^n_j} | \mathcal{F}_{t^n_{j-1}}].$$

If S is *regular* in the sense that for every stopping time τ and nondecreasing sequence $(\tau_n)_{n\in\mathbb{N}}$ of stopping times with $\tau = \lim_{n\to\infty} \tau_n$, we have $\lim_{n\to\infty} \mathbb{E}[S_{\tau_n}] = \mathbb{E}[S_\tau]$, or equivalently, if A is continuous, Doléans [14] showed that $A^n_t \to A_t$ uniformly in L^1 as $n \to \infty$ (see also Rogers and Williams [44], VI.31, Theorem 31.2). Hence, given that S is regular, we can extract a subsequence $\{A^{n_l}_t\}$, such that $\lim_{l\to\infty} A^{n_l}_t = A_t$ a.s. On the other hand, it is enough for G to be regular:

LEMMA 3.8. *Suppose $G \in \bar{\mathbb{G}}$ is a regular gains process. Then so is its Snell envelope process S.*

See Appendix A for the proof.

Remark 3.9. If it is not known that G is regular, Kobylanski and Quenez [32], in a slightly more general setting, showed that S is still regular, *provided* that G is upper semicontinuous in expectation along stopping times, i.e. for all $\tau \in \mathcal{T}^{0,T}$ and for all sequences of stopping times $(\tau_n)_{n\geq 1}$ such that $\tau_n \uparrow \tau$, we have

$$\mathbb{E}[G_\tau] \geq \limsup_{n\to\infty} \mathbb{E}[G_{\tau_n}].$$

The case where S is not regular is more subtle. In his classical paper Rao [41] utilised the Dunford-Pettis compactness criterion and showed that, in general, $A^n_t \to A_t$ only *weakly* in L^1 as $n \to \infty$ (a sequence $(X_n)_{n\in\mathbb{N}}$ of random variables in L^1 converges weakly in L^1 to X if for every bounded random variable Y we have that $\mathbb{E}[X_n Y] \to \mathbb{E}[XY]$ as $n \to \infty$).

Recall that *weak* convergence in L^1 does not imply convergence in probability, and therefore, we cannot immediately deduce an almost sure convergence along a subsequence. However, it turns out that by modifying the sequence of approximating random variables, the required convergence can be achieved. This has been done in recent improvements of the Doob-Meyer decomposition (see Jakubowski [29] and Beiglböck et al. [4]. Also, Siorpaes [48] showed that there is a subsequence that works for all $(t,\omega) \in [0,T]\times\Omega$ simultaneously). In particular, Jakubowski proceeds as Rao, but then uses Komlós's theorem [34] and proves the following (Jakubowski [29], Theorem 3 and Remark 1):

THEOREM 3.10. *There exists a subsequence $\{n_l\}$ such that for $t \in \cup_{n=1}^\infty \mathcal{P}_n$ and as $L \to \infty$*

$$\frac{1}{L}\Big(\sum_{l=1}^{L} A^{n_l}_t\Big) \to A_t, \quad \textit{a.s. and in } L^1. \tag{3.5}$$

In particular, in any subsequence we can find a further subsequence such that (3.5) *holds.*

Proof of Theorem 3.6. Let $(\sigma_n)_{n\geq 1}$ be a localising sequence for G such that, for each $n \geq 1$, $G^{\sigma_n} = (G_{t\wedge\sigma_n})_{0\leq t\leq T}$ is in $\mathcal{H}^1$. Similarly, set $S^{\sigma_n} = (S_{t\wedge\sigma_n})_{0\leq t\leq T}$ for a fixed $n \geq 1$. We need to prove that

$$0 \leq A^{\sigma_n}_t - A^{\sigma_n}_s \leq (D^-)^{\sigma_n}_t - (D^-)^{\sigma_n}_s \text{ a.s.,} \tag{3.6}$$

since then, as $\sigma_n \uparrow \infty$ almost surely, as $n \to \infty$, and by uniqueness of A and D^-, the result follows. In particular, since A is increasing, the first inequality in (3.6) is immediate, and thus we only need to prove the second one.

After localisation we assume that $G \in \mathcal{H}$. For any $0 \leq t \leq T$ and $0 \leq \epsilon \leq T - t$ we have that

$$\begin{aligned}
\mathbb{E}[S_{t+\epsilon}|\mathcal{F}_t] &= \mathbb{E}\Big[\operatorname*{ess\,sup}_{\tau\in\mathcal{T}_{t+\epsilon,T}} \mathbb{E}[G_\tau|\mathcal{F}_{t+\epsilon}]\Big|\mathcal{F}_t\Big] \\
&\geq \mathbb{E}\Big[\mathbb{E}[G_\tau|\mathcal{F}_{t+\epsilon}]\Big|\mathcal{F}_t\Big] \\
&= \mathbb{E}[G_\tau|\mathcal{F}_t] \text{ a.s.,}
\end{aligned}$$

where $\tau \in \mathcal{T}_{t+\epsilon,T}$ is arbitrary. Therefore

$$\mathbb{E}[S_{t+\epsilon}|\mathcal{F}_t] \geq \operatorname*{ess\,sup}_{\tau\in\mathcal{T}_{t+\epsilon,T}} \mathbb{E}[G_\tau|\mathcal{F}_t] \text{ a.s.} \tag{3.7}$$

Then by (2.3) and using (3.7) together with the properties of the *essential supremum* (see also Lemma A.1 in the Appendix A) we obtain

$$\begin{aligned}
\mathbb{E}[S_t - S_{t+\epsilon}|\mathcal{F}_t] &\leq \operatorname*{ess\,sup}_{\tau\in\mathcal{T}_{t,T}} \mathbb{E}[G_\tau|\mathcal{F}_t] - \operatorname*{ess\,sup}_{\tau\in\mathcal{T}_{t+\epsilon,T}} \mathbb{E}[G_\tau|\mathcal{F}_t] \\
&\leq \operatorname*{ess\,sup}_{\tau\in\mathcal{T}_{t,T}} \mathbb{E}[G_\tau - G_{\tau\vee(t+\epsilon)}|\mathcal{F}_t] \\
&= \operatorname*{ess\,sup}_{\tau\in\mathcal{T}_{t,t+\epsilon}} \mathbb{E}[G_\tau - G_{\tau\vee(t+\epsilon)}|\mathcal{F}_t] \qquad (3.8)\\
&= \operatorname*{ess\,sup}_{\tau\in\mathcal{T}_{t,t+\epsilon}} \mathbb{E}[G_\tau - G_{t+\epsilon}|\mathcal{F}_t] \text{ a.s.}
\end{aligned}$$

The first equality in (3.8) follows by noting that $\mathcal{T}_{t+\epsilon,T} \subset \mathcal{T}_{t,T}$, and that for any $\tau \in \mathcal{T}_{t+\epsilon,T}$ the term inside the expectation vanishes. Using the decomposition of G and by observing that, for all $\tau \in \mathcal{T}_{t,t+\epsilon}$, $(D^+_\tau - D^+_{t+\epsilon}) \leq 0$, while N is a uniformly integrable martingale, we obtain

$$\begin{aligned}
\mathbb{E}[S_t - S_{t+\epsilon}|\mathcal{F}_t] &\leq \operatorname*{ess\,sup}_{\tau\in\mathcal{T}_{t,t+\epsilon}} \mathbb{E}[D^-_{t+\epsilon} - D^-_\tau|\mathcal{F}_t] \\
&= \mathbb{E}[D^-_{t+\epsilon} - D^-_t|\mathcal{F}_t] \text{ a.s.} \qquad (3.9)
\end{aligned}$$

Finally, for $0 \leq s < t \leq T$, applying Theorem 3.10 to A together with (3.9) gives

$$\begin{aligned}
A_t - A_s &= \lim_{L\to\infty} \frac{1}{L}\Big(\sum_{l=1}^{L}\sum_{j=k'}^{k} \mathbb{E}[S_{t^{n_l}_{j-1}} - S_{t^{n_l}_j}|\mathcal{F}_{t^{n_l}_{j-1}}]\Big) \\
&\leq \lim_{L\to\infty} \frac{1}{L}\Big(\sum_{l=1}^{L}\sum_{j=k'}^{k} \mathbb{E}[D^-_{t^{n_l}_j} - D^-_{t^{n_l}_{j-1}}|\mathcal{F}_{t^{n_l}_{j-1}}]\Big) \text{ a.s.,} \qquad (3.10)
\end{aligned}$$

where $k' \leq k$ are such that $t^{n_l}_{k'} \leq s < t^{n_l}_{k'+1}$ and $t^{n_l}_k \leq t < t^{n_l}_{k+1}$. Note that D^- is also the predictable, increasing IV process in the Doob-Meyer decomposition of the class (D) supermartingale $(G - D^+)$. Therefore we can approximate it in the same way as A, so that $D^-_t - D^-_s$ is the almost sure limit along, possibly, a further subsequence $\{n_{l_k}\}$ of $\{n_l\}$, of the right hand side of (3.10). ∎

We finish this section with a lemma that gives an easy test as to whether the given process belongs to $\mathcal{H}^1_{loc}$ (consult Appendix A for the proof).

Lemma 3.11. *Let $X \in \mathbb{G}$ with a canonical decomposition $X = L + K$, where $L \in \mathcal{M}_{0,loc}$ and K is a predictable FV process. If the jumps of K are uniformly bounded by some finite constant $c > 0$, then $X \in \mathcal{H}^1_{loc}$.*

3.2. Markovian setting. In the rest of the section (and the paper) we consider the following optimal stopping problem:

$$V(x) = \sup_{\tau\in\mathcal{T}^{0,T}} \mathbb{E}_x[g(X_\tau)], \quad x \in E, \tag{3.11}$$

for a measurable function $g : E \to \mathbb{R}$ and a Markov process X satisfying the following set of assumptions:

ASSUMPTION 3.12. *X is a right process.*

ASSUMPTION 3.13. $\sup_{0\le t\le T}|g(X_t)|\in L^1(\mathbb{P}_x)$, $x\in E$.

ASSUMPTION 3.14. *$g\in\mathbb{D}(\mathcal{L})$, i.e. $g(\cdot)$ belongs to the domain of a martingale generator of X.*

Remark 3.15. Lemma 2.12 tells us that if X is Feller and F is an adapted path-functional of the form given in (2.7) then (a modification of) (X,F) satisfies Assumption 3.12.

Example 3.16. Let $X=(X_t)_{t\ge 0}$ be a Markov process and let $\mathbb{D}(\hat{\mathcal{L}})$ be the domain of a classical infinitesimal generator of X, i.e. the set of measurable functions $f:E\to\mathbb{R}$, such that $\lim_{t\to 0}(\mathbb{E}_x[f(X_t)]-f(x))/t$ exists. Then $\mathbb{D}(\hat{\mathcal{L}})\subset\mathbb{D}(\mathcal{L})$. In particular,

1. if $X=(X_t)_{t\ge 0}$ is a solution of an SDE driven by a Brownian motion in $\mathbb{R}^d$, then $C_b^2(\mathbb{R}^d,\mathbb{R})\subset\mathbb{D}(\hat{\mathcal{L}})$;
2. if the state space E is finite (so that X is a continuous time Markov chain), then any measurable and bounded $f:E\to\mathbb{R}$ belongs to $\mathbb{D}(\hat{\mathcal{L}})$
3. if X is a Lévy process on $\mathbb{R}^d$ with finite variance increments then $C_b^2(\mathbb{R}^d,\mathbb{R})\subset\mathbb{D}(\hat{\mathcal{L}})$

Note that the gains process is of the form $G=g(X)$, while by Theorem 2.10, the corresponding Snell envelope is given by

$$S_t^T:=\begin{cases}V(X_t):t<T,\\ g(X_T):t\ge T.\end{cases}$$

In a similar fashion to that in the general setting, Assumption 3.13 ensures the class (D) property for the gains and Snell envelope processes. Moreover, under Assumption 3.14,

$$g(X_t)=g(x)+M_t^g+\int_0^t\mathcal{L}g(X_s)ds,\quad 0\le t\le T,\,x\in E,\tag{3.12}$$

and the FV process in the semimartingale decomposition of $G=g(X)$ is absolutely continuous with respect to Lebesgue measure, and therefore predictable, so that (3.12) is a canonical semimartingale decomposition of $G=g(X)$. Then, by Assumption 3.13, and using Lemma 3.11, we also deduce that $g(X)\in\mathcal{H}^1_{loc}$.

Remark 3.17. When $T<\infty$, the optimal stopping problem, in general, is time-inhomogeneous, and we need to replace the process X_t by the process $Z_t=(t,X_t)$, $t\in[0,T]$, so that (3.11) reads

$$\tilde{V}(t,x)=\sup_{\tau\in\mathcal{T}_{0,T-t}}\mathbb{E}_{t,x}[\tilde{g}(t+\tau,X_{t+\tau})],\quad x\in E,\tag{3.13}$$

where $\tilde{g}:[0,T]\times E\to\mathbb{R}$ is a new payoff function (consult Peskir and Shiryaev [39] for examples). In this case, Assumption 3.14 should be replaced by a requirement that there exists a measurable function $\tilde{h}:[0,T]\times E\to\mathbb{R}$ such that $M_t^{\tilde{g}}:=\tilde{g}(Z_t)-\tilde{g}(0,x)-\int_0^t\tilde{h}(Z_s)ds$ defines a local martingale.

The crucial result of this section is the following:

THEOREM 3.18. *Suppose Assumptions 3.12,3.13 and 3.14 hold. Then $V\in\mathbb{D}(\mathcal{L})$.*

Proof. In order to be consistent with the notation in the general framework, let

$$D_t := g(X_0) + \int_0^t \mathcal{L}g(X_s)ds, \quad 0 \le t \le T.$$

Recall Lemma 2.7. Then D^+ and D^- are explicitly given (up to initial values) by

$$D_t^+ := \int_0^t \mathcal{L}g(X_s)^+ ds,$$
$$D_t^- := \int_0^t \mathcal{L}g(X_s)^- ds.$$

In particular, D^- is, as a measure, absolutely continuous with respect to Lebesgue measure. By applying Theorem 3.6, we deduce that

$$V(X_t) = V(x) + M_t^* - \int_0^t \mu_s \mathcal{L}g(X_s)^- ds, \quad 0 \le t \le T,\ x \in \mathbb{R}, \tag{3.14}$$

where μ is a non-negative Radon-Nikodym derivative with $0 \le \mu_s \le 1$. Then we also have that $\int_0^t |\mu_s \mathcal{L}g(X_s)^-| ds < \infty$, for every $0 \le t \le T$.

In order to finish the proof we are left to show that there exists a suitable measurable function $\lambda : E \to \mathbb{R}$ such that $A_t = \int_0^t \mu_s \mathcal{L}g(X_s)^- ds = \int_0^t \lambda(X_s)ds$ a.s., for all $t \in [0,T]$. For this, recall that a process Z (on $(\Omega, \mathcal{G}, \mathcal{G}_t, X_t, \theta_t, \mathbb{P}_x : x \in E, t \in \mathbb{R}_+)$ or just on $\mathcal{C}(X)$) is *additive* if $Z_0 = 0$ a.s. and $Z_{t+s} = Z_t + Z_s \circ \theta_t$ a.s., for all $s,t \in [0,T]$. Then, for any measurable function $f : E \to \mathbb{R}$, $Z_t^f = f(X_t) - f(x)$ defines an additive process. (Çinlar et al. [8] gives necessary and sufficient conditions for Z^f to be a semimartingale.) More importantly, if Z^f is a semimartingale, then the martingale and FV processes in the decomposition of Z^f are also additive, see Theorem 3.18 in Çinlar et al. [8].

Finally, we have that $A_t = \int_0^t \mu_s \mathcal{L}g(X_s)^- ds$, $t \in [0,T]$, is an increasing additive process such that $dA_t \ll dt$. Set $K_t = \liminf_{s \downarrow 0, s \in \mathbb{Q}} (A_{t+s} - A_t)/s$ and $\beta(x) = \mathbb{E}_x[K_0]$, $x \in E$. Then by Proposition 3.56 in Çinlar et al. [8], we have that, for $t \in [0,T]$, $A_t = \int_0^t \beta(X_s)ds$ $\mathbb{P}_x$-a.s. for each $x \in E$. ∎

Remark 3.19. In some specific examples it is possible to relax Assumption 3.14. Let $\mathcal{S} := \{x \in E : V(x) = g(x)\}$ be the stopping region. It is well-known that $S = V(X)$ is a martingale on the go region $\mathcal{S}^c$, i.e. M^c given by

$$M_t^c \overset{def}{=} \int_0^t 1_{(X_{s-} \in \mathcal{S}^c)} dS_s$$

is a martingale (see Lemma A.2). This implies that $\int_0^t 1_{(X_{s-} \in \mathcal{S}^c)} dA_s = 0$, and therefore we note that in order for $V \in \mathbb{D}(\mathcal{L})$, we need D to be absolutely continuous with respect to Lebesgue measure λ only on the stopping region i.e. that $\int_0^\cdot 1_{(X_{s-} \in \mathcal{S})} dD_s \ll \lambda$. For example, let $E = \mathbb{R}$, fix $K \in \mathbb{R}_+$ and consider $g(\cdot)$ given by $g(x) = (K - x)^+$, $x \in E$. We can easily show, under very weak conditions, that $\mathcal{S} \subset [0, K]$ and so we need only have that $\int_0^\cdot 1_{(X_{s-} < K)} dD_s$ is absolutely continuous.

4. Applications: duality, smooth fit. In this section we retain the setting of subsection 3.2.

4.1. Duality. Let $x \in E$ be fixed. As before, let $\mathcal{M}^x_{0,UI}$ denote all the right-continuous uniformly integrable càdlàg martingales (started at zero) on the filtered space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}_x)$, $x \in E$. The main result of Rogers [43] in the Markovian setting reads:

THEOREM 4.1. *Suppose Assumption 3.12 and 3.13 hold. Then*

$$V(x) = \sup_{\tau \in \mathcal{T}^{0,T}} \mathbb{E}_x[G_\tau] = \inf_{M \in \mathcal{M}^x_{0,UI}} \mathbb{E}_x\Big[\sup_{0 \le t \le T}\Big(G_t - M_t\Big)\Big], \quad x \in E. \tag{4.1}$$

We call the right hand side of (4.1) the *dual* of the optimal stopping problem. In particular, the right hand side of (4.1) is a "generalised stochastic control problem of Girsanov type", where a controller is allowed to choose a martingale from $\mathcal{M}^x_{0,UI}$, $x \in E$. Note that an optimal martingale for the dual is M^*, the martingale appearing in the Doob-Meyer decomposition of S, while any other martingale in $\mathcal{M}^x_{0,UI}$ gives an upper bound of $V(x)$. We already showed that $M^* = M^V$, which means that, when solving the dual problem, one can search only over martingales of the form M^f, for $f \in \mathbb{D}(\mathcal{L})$, or equivalently over the functions $f \in \mathbb{D}(\mathcal{L})$. We can further define $\mathcal{D}_{\mathcal{M}_{0,UI}} \subset \mathbb{D}(\mathcal{L})$ by

$$\mathcal{D}_{\mathcal{M}_{0,UI}} := \{f \in \mathbb{D}(\mathcal{L}) : f \ge g, f \text{ is superharmonic}, M^f \in \mathcal{M}_{0,UI}\}.$$

To conclude that $V \in \mathcal{D}_{\mathcal{M}_{0,UI}}$ we need to show that V is superharmonic, i.e. for all stopping times $\sigma \in \mathcal{T}^{0,T}$ and all $x \in E$, $\mathbb{E}_x[V(X_\sigma)] \le V(x)$. But this follows immediately from the Optional Sampling theorem, since $S = V(X)$ is a uniformly integrable supermartingale. Hence, as expected, we can restrict our search for the best minimising martingale to the set $\mathcal{D}_{\mathcal{M}_{0,UI}}$.

THEOREM 4.2. *Suppose that $G = g(X)$ and the assumptions of Theorem 3.18 hold. Let $\mathcal{D}_{\mathcal{M}_{0,UI}}$ be the set of admissible controls. Then the dual problem, i.e. the right hand side of (4.1), is a stochastic control problem for a controlled Markov process (X, Y^f, Z^f), $f \in \mathcal{D}_{\mathcal{M}_{0,UI}}$ (defined by (4.2) and (4.3)), with a value function $\hat{V}$ given by (4.4)*

Proof. For any $f \in \mathcal{D}_{\mathcal{M}^x_{0,UI}}$, $x \in E$ and $y, z \in \mathbb{R}$, define processes Y^f and Z^f via

$$Y^f_t := y + \int_0^t \mathcal{L}f(X_s)ds, \quad 0 \le t \le T, \tag{4.2}$$

$$Z^f_{s,t} := \sup_{s \le r \le t}\Big(f(x) + g(X_r) - f(X_r) + Y^f_r\Big), \quad 0 \le s \le t \le T, \tag{4.3}$$

and to allow arbitrary starting positions, set $Z^f_t = Z^f_{0,t} \vee z$, for $z \ge g(x) + y$. Note that, for any $f \in \mathbb{D}(\mathcal{L})$, Y^f is an additive functional of X. Lemma 2.12 implies that if $f \in \mathcal{D}_{\mathcal{M}_{0,UI}}$ then (X, Y^f, Z^f) is a Markov process.

Define $\hat{V} : E \times \mathbb{R}^2 \to \mathbb{R}$ by

$$\hat{V}(x, y, z) = \inf_{f \in \mathcal{D}_{\mathcal{M}^x_{0,UI}}} \mathbb{E}_{x,y,z}[Z^f_T], \quad (x, y, z) \in E \times \mathbb{R} \times \mathbb{R}. \tag{4.4}$$

It is clear that this is a stochastic control problem for the controlled Markov process (X, Y^f, Z^f), where the admissible controls are functions in $\mathcal{D}_{\mathcal{M}_{0,UI}}$. Moreover, since $V \in \mathcal{D}_{\mathcal{M}_{0,UI}}$, by virtue of Theorem 4.1, and adjusting initial conditions as necessary, we have

$$V(x) = \hat{V}(x, 0, g(x)) = \mathbb{E}_{x,0,g(x)}[Z^V_T], \quad x \in E.$$

□

4.2. Some remarks on the *smooth pasting* condition. We will now discuss the implications of Theorem 3.18 for the smoothness of the value function $V(\cdot)$ of the optimal stopping problem given in (3.11).

Remark 4.3. While in Theorem 4.4 (resp. Theorem 4.9) we essentially recover (a small improvement of) Theorem 2.3 in Peskir [37] (resp. Theorem 2.3 in Samee [45]), the novelty is that we prove the results by means of stochastic calculus, as opposed to the analytic approach in [37] (resp. [45]).

In addition to Assumption 3.13 and Assumption 3.14, we now assume that X is a one-dimensional diffusion in the Itô-McKean [26] sense, so that X is a strong Markov process with continuous sample paths. We also assume that the state space $E \subset \mathbb{R}$ is an interval with endpoints $-\infty \leq a \leq b \leq +\infty$. Nnote that the diffusion assumption implies Assumption 3.12. Finally, we assume that X is *regular*: for any $x, y \in \text{int}(E)$, $\mathbb{P}_x[\tau_y < \infty] > 0$, where $\tau_y = \min\{t \geq 0 : X_t = y\}$. Let $\alpha \geq 0$ be fixed; α corresponds to a killing rate of the sample paths of X.

The case without killing: $\alpha = 0$. Let $s(\cdot)$ denote a scale function of X, i.e. a continuous, strictly increasing function on E such that for l, r, $x \in E$, with $a \leq l < x < r \leq b$, we have

$$\mathbb{P}_x(\tau_r < \tau_l) = \frac{s(x) - s(l)}{s(r) - s(l)}, \tag{4.5}$$

see Revuz and Yor [42], Proposition 3.2 (p.301) for the proof of existence and properties of such a function.

From (4.5), using regularity of X and that $V(X)$ is a supermartingale of class (D) we have that $V(\cdot)$ is s-concave:

$$V(x) \geq V(l)\frac{s(r) - s(x)}{s(r) - s(l)} + V(r)\frac{s(x) - s(l)}{s(r) - s(l)}, \quad x \in [l, r]. \tag{4.6}$$

THEOREM 4.4. *Suppose the assumptions of Theorem 3.18 are satisfied, so that $V \in \mathbb{D}(\mathcal{L})$. Further assume that X is a regular, strong Markov process with continuous sample paths. Let $Y = s(X)$, where $s(\cdot)$ is a scale function of X.*

1. *Assume that for each $y \in [s(a), s(b)]$, the local time of Y at y, L^y, is singular with respect to Lebesgue measure. Then, if $s \in C^1$, $V(\cdot)$, given by (3.11), belongs to C^1.*
2. *Assume that $([Y, Y]_t)_{t \geq 0}$ is, as a measure, absolutely continuous with respect to Lebesgue measure. If $s'(\cdot)$ is absolutely continuous, then $V \in C^1$ and $V'(\cdot)$ is also absolutely continuous.*

Remark 4.5. If $\mathcal{G}$ is the filtration of a Brownian motion, B, then $Y = s(X)$ is a stochastic integral with respect to B (a consequence of martingale representation):

$$Y_t = Y_0 + \int_0^t \sigma_s dB_s. \tag{4.7}$$

Moreover, Proposition 3.56 in Çinlar et al. [8] ensures that $\sigma_t = \sigma(Y_t)$ for a suitably measurable function σ and

$$[Y, Y]_t = \int_0^t \sigma^2(Y_s) ds.$$

In this case, both, the singularity of the local time of Y and absolute continuity of $[Y, Y]$ (with respect to Lebesgue measure), are inherited from those of Brownian motion. On the other hand, if X is a regular diffusion (not necessarily a solution to an SDE driven by a Brownian motion), absolute continuity of $[Y, Y]$ still holds, if the speed measure of X is absolutely continuous (with respect to Lebesgue measure).

Proof. Note that $Y = s(X)$ is a Markov process, and let $\mathcal{K}$ denote its martingale generator. Moreover, $V(x) = W(s(x))$ (see Lemma 4.7 and the following remark), where, on the interval $[s(a), s(b)]$, $W(\cdot)$ is the smallest nonnegative concave majorant of the function $\hat{g}(y) = g \circ s^{-1}(y)$. Then, since $V \in \mathbb{D}(\mathcal{L})$,

$$V(X_t) = V(x) + M_t^V + \int_0^t \mathcal{L}V(X_u)du, \quad 0 \le t \le T,$$

and thus

$$W(Y_t) = W(y) + M_t^V + \int_0^t (\mathcal{L}V) \circ s^{-1}(Y_u)du, \quad 0 \le t \le T.$$

Therefore, $W \in \mathbb{D}(\mathcal{K})$, since

$$W(Y_t) = W(y) + M_t^V + \int_0^t \mathcal{K}W(Y_u)du, \tag{4.8}$$

for $y \in [s(a), s(b)]$, $0 \le t \le T$, with $\mathcal{K}W = \mathcal{L}V \circ s^{-1} \le 0$.

On the other hand, using the generalised Itô formula for concave/convex functions (see e.g. Revuz and Yor [42], Theorem 1.5 p.223) we have

$$W(Y_t) = W(y) + \int_0^t W_+^{'}(Y_u)dY_u - \int_{s(a)}^{s(b)} L_t^z \nu(dz),$$

for $y \in [s(a), s(b)]$, $0 \le t \le T$, where L_t^z is the local time of Y_t at z, and ν is a non-negative σ-finite measure corresponding to the second derivative of $-W$ in the sense of distributions. Then, by the uniqueness of the decomposition of a special semimartingale, we have that, for $t \in [0, T]$,

$$-\int_0^t \mathcal{K}W(Y_u)du = \int_{s(a)}^{s(b)} L_t^z \nu(dz) \quad \text{a.s.} \tag{4.9}$$

We prove the first claim by contradiction. Suppose that $\nu(\{z_0\}) > 0$ for some $z_0 \in (s(a), s(b))$. Then, using (4.9) we have that

$$-\int_0^t \mathcal{K}W(Y_u)du = L_t^{z_0}\nu(\{z_0\}) + \int_{s(a)}^{s(b)} \mathbb{1}_{\{z \neq z_0\}} L_t^z \nu(dz) \quad \text{a.s.} \tag{4.10}$$

Since $L_t^{z_0}$ is positive with positive probability and, by assumption, L^y, $y \in [s(a), s(b)]$, is singular with respect to Lebesgue measure, the process on the right hand side of (4.10) is not absolutely continuous with respect to Lebesgue measure, which contradicts absolute continuity of the left hand side. Therefore, $\nu(\{z_0\}) = 0$, and since z_0 was arbitrary, we have that ν does not charge points. It follows that $W \in C^1$. Since $s \in C^1$ by assumption, we conclude that $V \in C^1$.

We now prove the second claim. By assumption, $[Y,Y]$ is absolutely continuous with respect to Lebesgue measure (on the time axis). Invoking Proposition 3.56 in Çinlar et al. [8] again, we have that

$$[Y,Y]_t = \int_0^t \sigma^2(Y_u)du$$

(as in Remark 4.5). A time-change argument allows us to conclude that Y is a time-change of a BM and that we may neglect the set $\{t : \sigma^2(Y_t) = 0\}$ in the representation (4.8). Thus

$$W(Y_t) = W(Y_0) + \int_0^t 1_{N^c}(Y_u)dM_u^V + \int_0^t 1_{N^c}(Y_u)\mathcal{K}W(Y_u)du$$

where N is the zero set of σ. Then, using the occupation time formula (see, for example, Revuz and Yor [42], Theorem 1.5 p.223) we have that

$$-\int_0^t \mathcal{K}W(Y_u)du = \int_0^t f(Y_u)d[Y,Y]_u = \int_{s(b)}^{s(b)} f(z)L_t^z dz \quad \text{a.s.},$$

where $f : [s(a), s(b)] \to \mathbb{R}$ is given by $f : y \mapsto -\frac{\mathcal{K}W}{\sigma^2}1_{N^c}(y)$. Now observe that, for $0 \le r \le t \le T$, $\eta([r,t]) := \int_{s(a)}^{s(b)} f(z)\Big(L_t^z - L_r^z\Big)dz$ and $\pi([r,t]) := \int_{s(a)}^{s(b)} \Big(L_t^z - L_r^z\Big)\nu(dz)$ define measures on the time axis, which, by virtue of (4.9), are equal (and thus both are absolutely continuous with respect to Lebesgue measure). Now define $T^{\underline{l},\bar{l}} := \{t : Y_t \in [\underline{l}, \bar{l}]\}$, $s(a) \le \underline{l} \le \bar{l} \le s(b)$. Then the restrictions of η and π to $T^{\underline{l},\bar{l}}$, $\eta|_{T^{\underline{l},\bar{l}}}$ and $\pi|_{T^{\underline{l},\bar{l}}}$, are also equal. Moreover, since Y is a local martingale, it is also a semimartingale. Therefore, for every $0 \le t \le T$, L_t^z is carried by the set $\{t : Y_t = z\}$ (see Protter [40], Theorem 69 p.217). Hence, for each $t \in [0,T]$,

$$\eta|_{T^{\underline{l},\bar{l}}}([0,t]) = \int_{\underline{l}}^{\bar{l}} L_t^z f(z)dz = \int_{\underline{l}}^{\bar{l}} L_t^z \nu(dz) = \pi|_{T^{\underline{l},\bar{l}}}([0,t]), \tag{4.11}$$

and, since $\underline{l}$ and $\bar{l}$ are arbitrary, the left and right hand sides of (4.11) define measures on $[s(a), s(b)] \subseteq \mathbb{R}$, which are equal. It follows that ν is absolutely continuous with respect to Lebesgue measure on $[s(a), s(b)]$ and $f(z)dz = \nu(dz)$. This proves that $W \in C^1$ and $W'(\cdot)$ is absolutely continuous on $[s(a), s(b)]$ with Radon-Nikodym derivative f. Since the product and composition of absolutely continuous functions are absolutely continuous, we conclude that $V'(\cdot)$ is absolutely continuous (since $s'(\cdot)$ is, by assumption). ∎

Remark 4.6. We note that for a smooth fit principle to hold, it is not necessary that $s \in C^1$. Given that all the other conditions of Theorem 4.4 hold, it is sufficient that $s(\cdot)$ is differentiable at the boundary of the continuation region. On the other hand, if $g \in \mathbb{D}(\mathcal{L})$, $V \in C^1$, even if $g \notin C^1$.

Moreover, since $V = g$ on the stopping region, Theorem 4.4 tells us that $g \in C^1$ on the interior of the stopping region. However, the question whether this stems already from the assumption that $g \in \mathbb{D}(\mathcal{L})$ is more subtle. For example, if $g \in \mathbb{D}(\mathcal{L})$ and g is a difference of two convex functions, then by the generalised Itô formula and the local time argument (similarly to the proof of Theorem 4.4) we could conclude that $g \in C^1$ on the whole state space E.

Case with killing: $\alpha > 0$. We now generalise the results of the Theorem 4.4 in the presence of a non-trivial killing rate. Consider the following optimal stopping problem

$$V(x) = \sup_{\tau \in \mathcal{T}^{0,T}} \mathbb{E}_x[e^{-\alpha\tau} g(X_\tau)], \quad x \in E. \tag{4.12}$$

Note that, since $\alpha > 0$, using the regularity of X together with the supermartingale property of $V(X)$ we have that

$$V(x) \geq V(l)\mathbb{E}_x[e^{-\alpha\tau_l} 1_{\tau_l < \tau_r}] + V(r)\mathbb{E}_x[e^{-\alpha\tau_r} 1_{\tau_r < \tau_l}], \quad x \in [l, r] \subseteq E. \tag{4.13}$$

Define increasing and decreasing functions $\psi, \phi : E \to \mathbb{R}$, respectively, by

$$\psi(x) = \begin{cases} \mathbb{E}_x[e^{-\alpha\tau_c}], & \text{if } x \leq c \\ 1/\mathbb{E}_c[e^{-\alpha\tau_x}], & \text{if } x > c \end{cases} \quad \phi(x) = \begin{cases} 1/\mathbb{E}_c[e^{-\alpha\tau_x}], & \text{if } x \leq c \\ \mathbb{E}_x[e^{-\alpha\tau_c}], & \text{if } x > c \end{cases} \tag{4.14}$$

where $c \in E$ is arbitrary. Then, $(\Psi_t)_{0 \leq t \leq T}$ and $(\Phi_t)_{0 \leq t \leq T}$, given by

$$\Psi_t = e^{-\alpha t}\psi(X_t), \quad \Phi_t = e^{-\alpha t}\phi(X_t), \quad 0 \leq t \leq T,$$

respectively, are local martingales (and also supermartingales, since ψ, ϕ are non-negative); see Dynkin [15] and Itô and McKean [26].

Let $p_1, p_2 : [l, r] \to [0, 1]$ (where $[l, r] \subseteq E$) be given by

$$p_1(x) = \mathbb{E}_x[e^{-\alpha\tau_l} 1_{\tau_l < \tau_r}], \quad p_2(x) = \mathbb{E}_x[e^{-\alpha\tau_r} 1_{\tau_r < \tau_l}].$$

Continuity of paths of X implies that $p_i(\cdot), i = 1, 2$, are both continuous (the proof of continuity of the scale function in (4.5) can be adapted for a killed process). In terms of the functions $\psi(\cdot)$, $\phi(\cdot)$ of (4.14), using appropriate boundary conditions, one calculates

$$p_1(x) = \frac{\psi(x)\phi(r) - \psi(r)\phi(x)}{\psi(l)\phi(r) - \psi(r)\phi(l)}, \quad p_2(x) = \frac{\psi(l)\phi(x) - \psi(x)\phi(l)}{\psi(l)\phi(r) - \psi(r)\phi(l)}, \quad x \in [l, r]. \tag{4.15}$$

Let $\tilde{s} : E \to \mathbb{R}_+$ be the continuous increasing function defined by $\tilde{s}(x) = \psi(x)/\phi(x)$. Substituting (4.15) into (4.13) and then dividing both sides by $\phi(x)$ we get

$$\frac{V(x)}{\phi(x)} \geq \frac{V(l)}{\phi(l)} \cdot \frac{\tilde{s}(r) - \tilde{s}(x)}{\tilde{s}(r) - \tilde{s}(l)} + \frac{V(r)}{\phi(r)} \cdot \frac{\tilde{s}(x) - \tilde{s}(l)}{\tilde{s}(r) - \tilde{s}(l)}, \quad x \in [l, r] \subseteq E,$$

so that $V(\cdot)/\phi(\cdot)$ is $\tilde{s}$-concave.

Recall that (4.13) essentially follows from $V(\cdot)$ being α-superharmonic, so that it satisfies $\mathbb{E}_x[e^{-\alpha\tau} V(X_\tau)] \leq V(x)$ for $x \in E$ and any stopping time τ. Since Φ and Ψ are local martingales, it follows that the converse is also true, i.e. given a measurable function $f : E \to \mathbb{R}$, $f(\cdot)/\phi(\cdot)$ is $\tilde{s}$-concave if and only if $f(\cdot)$ is α-superharmonic (Dayanik and Karatzas [11], Proposition 4.1). This shows that a value function $V(\cdot)$ is the minimal majorant of $g(\cdot)$ such that $V(\cdot)/\phi(\cdot)$ is $\tilde{s}$-concave.

LEMMA 4.7. *Suppose $[l, r] \subseteq E$ and let $W(\cdot)$ be the smallest nonnegative concave majorant of $\tilde{g} := (g/\phi) \circ \tilde{s}^{-1}$ on $[\tilde{s}(l), \tilde{s}(r)]$, where $\tilde{s}^{-1}$ is the inverse of $\tilde{s}$. Then $V(x) = \phi(x)W(\tilde{s}(x))$ on $[l, r]$.*

Proof. Define $\hat{V}(x) = \phi(x)W(\tilde{s}(x))$ on $[l, r]$. Then, trivially, $\hat{V}(\cdot)$ majorizes $g(\cdot)$ and $\hat{V}(\cdot)/\phi(\cdot)$ is $\tilde{s}$-concave. Therefore $V(x) \leq \hat{V}(x)$ on $[l, r]$.

On the other hand, let $\hat{W}(y) = (V/\phi)(\tilde{s}^{-1}(y))$ on $[\tilde{s}(l), \tilde{s}(r)]$. Since $V(x) \geq g(x)$ and $(V/\phi)(\cdot)$ is $\tilde{s}$-concave on $[l, r]$, $\hat{W}(\cdot)$ is concave and majorizes $(g/\phi) \circ \tilde{s}^{-1}(\cdot)$ on $[\tilde{s}(l), \tilde{s}(r)]$. Hence, $W(y) \leq \hat{W}(y)$ on $[\tilde{s}(l), \tilde{s}(r)]$.

Finally, $(V/\phi)(x) \leq (\hat{V}/\phi)(x) = W(\tilde{s}(x)) \leq \hat{W}(\tilde{s}(x)) = (V/\phi)(x)$ on $[l, r]$. ∎

Remark 4.8. When $\alpha = 0$, let $(\psi, \phi) = (s, 1)$. Then Lemma 4.7 is just Proposition 4.3. in Dayanik and Karatzas [11].

With the help of Lemma 4.7 and using parallel arguments to those in the proof of Theorem 4.4 we can formulate sufficient conditions for V to be in C^1 and have absolutely continuous derivative.

THEOREM 4.9. *Suppose the assumptions of Theorem 3.18 are satisfied, so that $V \in \mathbb{D}(\mathcal{L})$. Further assume that X is a regular Markov process with continuous sample paths. Let $\psi(\cdot), \phi(\cdot)$ be as in (4.14) and consider the process $Y = \tilde{s}(X)$.*

1. *Assume that, for each $y \in [\tilde{s}(a), \tilde{s}(b)]$, the local time of Y at $y \in [\tilde{s}(a), \tilde{s}(b)]$, $\hat{L}^y$, is singular with respect to Lebesgue measure. Then if $\psi, \phi \in C^1$, $V(\cdot)$, given by (4.12), belongs to C^1.*
2. *Assume that $[Y, Y]$ is, as a measure, absolutely continuous with respect to Lebesgue measure. If $\psi'(\cdot), \phi'(\cdot)$ are both absolutely continuous, then $V'(\cdot)$ is aslo absolutely continuous.*

Proof. First note that Y is not necessarily a local martingale, while ΦY is. Indeed, $\Phi Y = \Psi$. Hence

$$(N_t)_{0 \leq t \leq T} := \Big(\int_0^t \Phi_t dY_t + [\Phi, Y]_t \Big)_{0 \leq t \leq T}$$

is the difference of two local martingales, and thus is a local martingale itself. Using the generalised Itô formula for concave/convex functions, we have

$$\Phi_t W(Y_t) = \Phi_0 W(y) + \int_0^t W(Y_s) d\Phi_s + \int_0^t W'_+(Y_s) dN_s - \int_{\tilde{s}(l)}^{\tilde{s}(r)} \Phi_t \hat{L}_t^z \nu(dz), \tag{4.16}$$

for $y \in [\tilde{s}(l), \tilde{s}(r)]$, $0 \leq t \leq T$, where $\hat{L}_t^z$ is the local time of Y_t at z, and ν is a non-negative σ-finite measure corresponding to the derivative W'' in the sense of distributions.

On the other hand, if $g \in \mathbb{D}(\mathcal{L})$, then $V \in \mathbb{D}(\mathcal{L})$. Therefore,

$$e^{-\alpha t} V(X_t) = V(x) + \int_0^t e^{-\alpha s} dM_s^V + \int_0^t e^{-\alpha s} \{\mathcal{L} - \alpha\} V(X_s) ds, \quad 0 \leq t \leq T. \tag{4.17}$$

Then, similarly to before, from the uniqueness of the decomposition of the Snell envelope, we have that the martingale and FV terms in (4.16) and (4.17) coincide. Hence, for $t \in [0, T]$,

$$\int_{\tilde{s}(l)}^{\tilde{s}(r)} e^{-\alpha t} \phi(X_t) \hat{L}_t^z \nu(dz) = -\int_0^t e^{-\alpha s} \{\mathcal{L} - \alpha\} V(X_s) ds \quad \text{a.s.}$$

Using the same arguments as in the proof of Theorem 4.4 we can show that both statements of this theorem hold. The details are left to the reader. ∎

Acknowledgments. We are grateful to two anonymous referees and Prof. Goran Peskir for useful comments and suggestions.

REFERENCES

[1] L. Andersen and M. Broadie, *Primal-dual simulation algorithm for pricing multidimensional American options*, Management Science, 50(9):1222–1234, (2004).

[2] S. Ankirchner, M. Klein, and T. Kruse, *A verification theorem for optimal stopping problems with expectation constraints*, Appl. Math. Optim., 1–33, (2015).

[3] D. Assaf, L. Goldstein, and E. Samuel-Cahn, *Ratio prophet inequalities when the mortal has several choices*, Ann. Appl. Probab., 12(3):972–984, (2002).

[4] M. Beiglboeck, W. Schachermayer, and B. Veliyev, *A short proof of the Doob-Meyer theorem*, Stochastic Process. Appl., 122(4):1204–1209, (2012).

[5] D. Belomestny, C. Bender, and J. Schoenmakers, *True upper bounds for Bermudan products via non-nested Monte Carlo*, Math. Finance, 19(1):53–71, (2009).

[6] D. Belomestny, *Solving optimal stopping problems via empirical dual optimization*, Ann. Appl. Probab., 23(5):1988–2019, (2013).

[7] R. M. Blumenthal and R. K. Getoor, *Markov processes and potential theory*, Courier Corporation, (2007).

[8] E. Çinlar, J. Jacod, P. Protter and M. J. Sharpe, *Semimartingales and Markov processes*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, 54(2):161-219, (1980).

[9] S. Crepéy and A. Matoussi, *Reflected and doubly reflected BSDEs with jumps: a priori estimates and comparison*, Ann. Appl. Probab., 18(5):2041-2069, (2008).

[10] M. Davis and I. Karatzas, *A deterministic approach to optimal stopping*, Probability, Statistics and Optimisation (ed. FP Kelly). NewYork Chichester: John Wiley & Sons Ltd, 455–466, (1994).

[11] S. Dayanik and I. Karatzas, *On the optimal stopping problem for one-dimensional diffusions*, Stochastic Process. Appl., 107(2):173–212, (2003).

[12] C. Dellacherie and P.-A. Meyer, *Probabilities and potential. B, volume 72 of North-Holland Mathematics Studies*, (1982).

[13] V. V. Desai, V. F. Farias, and C. C. Moallemi, *Pathwise optimization for optimal stopping problems*, Management Science, 58(12):2292–2308, (2012).

[14] C. Doléans, *Existence du processus croissant naturel associé à un potentiel de la classe (D)*, Probab. Theory Related Fields, 9(4):309–314, (1968).

[15] E. B. Dynkin, *Markov processes*, vol. 2, Springer, (1965).

[16] N. El Karoui, *Les aspects probabilistes du contrôle stochastique*, Éc. Été Probab. St.-Flour IX-1979, 73-238. Springer, Berlin, Heidelberg, (1981).

[17] N. El Karoui, C. Kapoudjian, E. Pardoux, S. Peng, and M. C. Quenez, *Reflected solutions of backward SDE's, and related obstacle problems for PDE's*, Ann. Probab., 25(2):702–737, (1997).

[18] N. El Karoui, J. P. Lepeltier, and A. Millet, *A probabilistic approach to the reduite in optimal stopping*, Probab. Math. Statist., 13(1):97–121, (1992).

[19] W. H. Fleming and H. M. Soner, *Controlled Markov processes and viscosity solutions*, Vol. 25, Springer Science & Business Media, (2006).

[20] R. K. Getoor, *Markov processes: Ray processes and right processes*, Springer, Vol. 440, (2006).

[21] P. Glasserman, B. Yu, *Number of paths versus number of basis functions in American option pricing*, Ann. Appl. Probab., 14(4):2090–2119, (2004).

[22] I. Gyöngy and D. Šiška, *On randomized stopping*, Bernoulli, 14(2):352–361, (2008).

[23] S. Hamadène and Y. Ouknine *Reflected backward stochastic differential equation with jumps and random obstacle*, Electron. J. Probab., 8, (1993).

[24] M. B. Haugh and L. Kogan, *Pricing American options: a duality approach*, Oper. Res., 52(2):258–270, (2004).

[25] T. Hill, *Prophet inequalities and order selection in optimal stopping problems*, Proc. Amer. Math. Soc., 88(1):131–137, (1983).

[26] K. Itô and H. P. McKean, Jr., *Diffusion Processes and Their Sample Paths*, Grundlehren der Mathematischen Wissenschaften, 125, (1965).

[27] S. D. Jacka, *Local times, optimal stopping and semimartingales*, Ann. Probab., 21(1):329–339, (1993).

[28] J. Jacod and A. N. Shiryaev, *Limit theorems for stochastic processes*, Vol. 288, Springer Science & Business Media, (2013).

[29] A. Jakubowski, *An almost sure approximation for the predictable process in the Doob-Meyer decomposition theorem*, In Séminaire de Probabilités XXXVIII, Springer, Berlin, Heidelberg, 158–164, (2005).

[30] O. Kallenberg, *Foundations of modern probability*, Springer Science & Business Media, (2006).

[31] I. Karatzas, S. E. Shreve, *Methods of mathematical finance*, Vol 39, New York: Springer, (1998).
[32] M. Kobylanski, M.-C. Quenez, et al., *Optimal stopping time problem in a general framework*, Electron. J. Probab., 17(72):1–28, (2012).
[33] A. Kolodko and J. Schoenmakers, *Upper bounds for Bermudan style derivatives*, Monte Carlo Methods and Appl., 10(3-4):331–343, (2004).
[34] J. Komlós, *A generalization of a problem of Steinhaus*, Acta Math. Hungar., 18(1-2):217–229, (1967).
[35] N. V. Krylov, *Controlled diffusion processes*, Vol. 14, Springer Science & Business Media, (2008).
[36] C. W. Miller, *Nonlinear pde approach to time-inconsistent optimal stopping*, SIAM J. Control Optim., 55(1):557–573, (2017).
[37] G. Peskir *Principle of smooth fit and diffusions with angles*, Stochastics, 79(3-4):293–302, (2007).
[38] G. Peskir *A duality principle for the Legendre transform*, J. Convex Anal., 19(3):609–630, (2012).
[39] G. Peskir and A. N. Shiryaev, *Optimal stopping and free-boundary problems*, Birkhäuser Basel, (2006).
[40] P. E. Protter, *Stochastic integration and differential equations*, Springer, (2005).
[41] K. M. Rao, *On decomposition theorems of Meyer*, Math. Scand., 24(1):66–78, (1969).
[42] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, Vol. 293, Springer Science & Business Media, (2013).
[43] L. C. G. Rogers, *Monte Carlo valuation of American options*, Math. Finance, 12(3):271–286, (2002).
[44] L. C. G. Rogers and D. Williams, *Diffusions, Markov processes and martingales: Volume 2, Itô Calculus*, Vol. 2, Cambridge university press, (2000).
[45] F. Samee, *On the principle of smooth fit for killed diffusions*, Electron. Commun. Probab., 15:89–98, (2010).
[46] M. Sharpe, *General theory of Markov processes*, Vol. 133, Academic press, (1988).
[47] A. N. Shiryaev, *Optimal stopping rules*, Vol. 8, Springer Science & Business Media, (2007).
[48] P. Siorpaes, *On a dyadic approximation of predictable processes of finite variation*, Electron. Commun. Probab., 19(22):1–12, (2014).
[49] J. L. Snell, *Applications of martingale system theorems*, Trans. Amer. Math. Soc., 73(2):293–312 (1952).

Appendix A.

Lemma A.1. *For each $0 \le t \le T$, the family of random variables $\{\mathbb{E}[G_\tau|\mathcal{F}_t] : \tau \in \mathcal{T}_{t,T}\}$ is directed upwards, i.e. for any σ_1, $\sigma_2 \in \mathcal{T}_{t,T}$, there exists $\sigma_3 \in \mathcal{T}_{t,T}$, such that*

$$\mathbb{E}[G_{\sigma_1}|\mathcal{F}_t] \vee \mathbb{E}[G_{\sigma_1}|\mathcal{F}_t] \le \mathbb{E}[G_{\sigma_3}|\mathcal{F}_t], \ \ a.s.$$

Proof. Fix $t \in [0, T]$. Suppose σ_1, $\sigma_2 \in \mathcal{T}_{t,T}$ and define $A := \{\mathbb{E}[G_{\sigma_1}|\mathcal{F}_t] \ge \mathbb{E}[G_{\sigma_2}|\mathcal{F}_t]\}$. Let $\sigma_3 := \sigma_1 \mathbb{1}_A + \sigma_2 \mathbb{1}_{A^c}$. Note that $\sigma_3 \in \mathcal{T}_{t,T}$. Using $\mathcal{F}_t$-measurability of A, we have

$$\begin{aligned}\mathbb{E}[G_{\sigma_3}|\mathcal{F}_t] &= \mathbb{1}_A \mathbb{E}[G_{\sigma_1}|\mathcal{F}_t] + \mathbb{1}_{A^c}\mathbb{E}[G_{\sigma_2}|\mathcal{F}_t] \\ &= \mathbb{E}[G_{\sigma_1}|\mathcal{F}_t] \vee \mathbb{E}[G_{\sigma_2}|\mathcal{F}_t] \text{ a.s.,}\end{aligned}$$

which proves the claim. ∎

Lemma A.2. *Let $G \in \bar{\mathbb{G}}$ and S be its Snell envelope with decomposition $S = M^* - A$. For $0 \le t \le T$ and $\epsilon > 0$, define*

$$K_t^\epsilon = \inf\{s \ge t : G_s \ge S_s - \epsilon\}. \tag{A.1}$$

Then $A_{K_t^\epsilon} = A_t$ a.s. and the processes $(A_{K_t^\epsilon})$ and A are indistinguishable.

Proof. From the directed upwards property (Lemma A.1) we know that $\mathbb{E}[S_t] = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}[G_\tau]$. Then for a sequence $(\tau_n)_{n\in\mathbb{N}}$ of stopping times in $\mathcal{T}_{t,T}$, such that

$\lim_{n\to\infty} \mathbb{E}[G_{\tau_n}] = \mathbb{E}[S_t]$, we have

$$\mathbb{E}[G_{\tau_n}] \le \mathbb{E}[S_{\tau_n}] = \mathbb{E}[M^*_{\tau_n} - A_{\tau_n}] = \mathbb{E}[S_t] - \mathbb{E}[A_{\tau_n} - A_t],$$

since M^* is uniformly integrable. Hence, since A is non-decreasing,

$$0 \le \lim_{n\to\infty} \mathbb{E}[S_{\tau_n} - G_{\tau_n}] = -\lim_{n\to\infty} \mathbb{E}[A_{\tau_n} - A_t] \le 0,$$

and thus we have equalities throughout. By passing to a sub-sequence we can assume that

$$\lim_{n\to\infty} (S_{\tau_n} - G_{\tau_n}) = 0 = \lim_{n\to\infty} (A_{\tau_n} - A_t) \quad \text{a.s.} \tag{A.2}$$

The first equality in (A.2) implies that $K^\epsilon_t \le \tau_{n_0}$ a.s., for some large enough $n_0 \in \mathbb{N}$, and thus $A_{K^\epsilon_t} \le A_{\tau_n}$, for all $n_0 \le n$. Since A is non-decreasing, we also have that $0 \le A_{K^\epsilon_t} - A_t \le A_{\tau_n} - A_t$ a.s., $n_0 \le n$, and from the second equality in (A.2) we conclude that $A_{K^\epsilon_t} = A_t$ a.s. The indistinguishability follows from the right-continuity of G and S. □

A.1. Proofs of results in section 2.

Proof of Lemma 2.12. The completed filtration generated by a Feller process satisfies the *usual assumptions*, in particular, it is both right-continuous and quasi-left-continuous. The latter means that for any predictable stopping time σ, $\mathcal{F}_{\sigma-} = \mathcal{F}_\sigma$. Moreover, every càdlàg Feller process is left-continuous over stopping times and satisfies the strong Markov property. On the other hand, every Feller process admits a càdlàg modification (these are standard results and can be found, for example, in Revuz and Yor [42] or Rogers and Williams [44]). All that remains is to show that the addition of the functional F leaves (X, F) strong Markov. This is elementary from (2.7). □

A.2. Proofs of results in section 3.

Proof of Lemma 3.8. Let $(\tau_n)_{n\in\mathbb{N}}$ be a nondecreasing sequence of stopping times with $\lim_{n\to\infty} \tau_n = \tau$, for some fixed $\tau \in \mathcal{T}_{0,T}$. Since S is a supermartingale, $\mathbb{E}[S_{\tau_n}] \ge \mathbb{E}[S_\tau]$, for every $n \in \mathbb{N}$. For a fixed $\epsilon > 0$, $K^\epsilon_{\tau_n}$ (defined by (A.1)) is a stopping time, and by Lemma A.2, $A_{K^\epsilon_{\tau_n}} = A_{\tau_n}$ a.s. Therefore, since M^* is uniformly integrable,

$$\mathbb{E}[S_{K^\epsilon_{\tau_n}}] = \mathbb{E}[M^*_{K^\epsilon_{\tau_n}} - A_{K^\epsilon_{\tau_n}}] = \mathbb{E}[M^*_{\tau_n} - A_{\tau_n}] = \mathbb{E}[S_{\tau_n}].$$

Thus, by the definition of $K^\epsilon_{\tau_n}$,

$$\mathbb{E}[G_{K^\epsilon_{\tau_n}}] \ge \mathbb{E}[S_{K^\epsilon_{\tau_n}}] - \epsilon = \mathbb{E}[S_{\tau_n}] - \epsilon.$$

Let $\hat{\tau} := \lim_{n\to\infty} K^\epsilon_{\tau_n}$. Note that the sequence $(K^\epsilon_{\tau_n})_{n\in\mathbb{N}}$ is non-decreasing and dominated by K^ϵ_τ. Hence $\tau \le \hat{\tau} \le K^\epsilon_\tau$. Finally, using the regularity of G we obtain

$$\mathbb{E}[S_\tau] \ge \mathbb{E}[S_{\hat{\tau}}] \ge \mathbb{E}[G_{\hat{\tau}}] = \lim_{n\to\infty} \mathbb{E}[G_{K^\epsilon_{\tau_n}}] \ge \lim_{n\to\infty} \mathbb{E}[S_{\tau_n}] - \epsilon.$$

Since ϵ is arbitrary, the result follows. □

Proof of Lemma 3.11. For $n \ge 1$, define

$$\tau_n := \inf\{t \ge 0 : \int_0^t |dK_s| \ge n\}.$$

Clearly $\tau_n \uparrow \infty$ as $n \to \infty$. Then for each $n \geq 1$

$$\begin{aligned}\mathbb{E}[\int_0^{t\wedge\tau_n} |dK_s|] &\leq \mathbb{E}[\int_0^{\tau_n} |dK_s|] \\ &= \mathbb{E}[\int_0^{\tau_n -} |dK_s|] + |\Delta K_{\tau_n}|] \\ &\leq n + c.\end{aligned}$$

Therefore, since $X \in \mathbb{G}$,

$$||L^{\tau_n}||_{\mathcal{S}^1} \leq ||X^{\tau_n}||_{\mathcal{S}^1} + \mathbb{E}[\int_0^{\tau_n} |dK_s|] < \infty,$$

and thus, $||X^{\tau_n}||_{\mathcal{H}^1} < \infty$, for all $n \geq 1$. □