UNIVERSITY OF LEEDS

This is a repository copy of *Infinite horizon stopping problems with (nearly) total reward criteria*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/81218/

Version: Accepted Version

Article:

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Infinite horizon stopping problems with (nearly) total reward criteria

Jan Palczewski[a], Łukasz Stettner[b]

[a]*School of Mathematics, University of Leeds, Leeds LS2 9JT, UK*
[b]*Institute of Mathematics, Polish Academy of Sciences, Sniadeckich 8, 00-956 Warszawa, Poland and Vistula University, Warszawa, Poland*

Abstract

We study an infinite horizon optimal stopping Markov problem which is either undiscounted (total reward) or with a general Markovian discount rate. Using ergodic properties of the underlying Markov process, we establish the feasibility of the stopping problem and prove the existence of optimal and ε-optimal stopping times. We show the continuity of the value function and its variational characterisation (in the viscosity sense) under different sets of assumptions satisfied by large classes of diffusion and jump-diffusion processes. In the case of a general discounted problem we relax a classical assumption that the discount rate is uniformly separated from zero.

Keywords: optimal stopping, infinite horizon, total reward, general Makovian discounting, non-uniformly ergodic Markov processes

1. Introduction

The theory of optimal stopping has recently seen its renaissance due to applications in finance and operations research (e.g., pricing of American options, optimal timing of a sale or valuation of natural resources, see [23] and references therein, or pricing of swing options with applications in energy trading [3]), and statistics (e.g., sequential hypothesis testing, see [22, 26] and references therein). In some applications, such as the valuation of American options, there is a natural bound for stopping times – the maturity date of the option. This results in finite horizon stopping problems. In others, like finding an optimal time to sell a stock/business or valuing natural resources, the horizon is infinite leading to an optimal stopping problem of the following form:

$$w(x) = \sup_{\tau} \limsup_{T \to \infty} \mathbb{E}^x \left\{ \int_0^{\tau \wedge T} e^{-\int_0^s r(X_u)du} f(X_s)ds + e^{-\int_0^{\tau \wedge T} r(X_u)du} g(X_{\tau \wedge T}) \right\}. \qquad (1)$$

Here, the stochastic process $X(t)$ models the underlying randomness in the world, such as the price of a stock, the price of a natural resource (gas, oil, etc.), or factors influencing the interest rate. The running profit/cost is represented by f while the proceeds from the final sale at time τ or the closure of the production facility are given by g. The function r corresponds to an instantaneous interest rate that is used for discounting of future cash-flows. For years, it had seemed sensible to assume that interest rates were uniformly separated from 0. However, recent developments in Japan and in Europe showed that (even nominal) interest rates for government bonds can get arbitrarily close to 0.

The uniform separation of the discount rate from 0 (i.e., $\inf_x r(x) > 0$) is in line with a large strand of literature in infinite-horizon optimal stopping [2, 22, 23, 24]; commonly, the discount rate is a positive constant. This ensures, under appropriate assumptions on the growth of f and g, that the value function is finite and allows to approximate the infinite horizon stopping problem with finite horizon ones. An abolition of the discounting or a relaxation of its separation from zero brings in a lot of difficulties. In particular, the integrability in the functional is at risk; in Section 2 we show that this is indeed the case. A solution to this problem is to impose restrictive assumptions on f and g. In martingale approaches it is common to assume that $\mathbb{E}^x\{\int_0^\infty e^{-\int_0^s r(X_u)du} f^-(X_s)ds\} < \infty$ for every x and that the family $\{e^{-\int_0^\tau r(X_u)du} g^-(X_\tau) : \ \tau\text{-a stopping time}\}$, is uniformly integrable, see [20, 22] and references therein. Alternatively, one can take f to be non-positive (then the integral term penalises for waiting only) [22, 26].

A general stopping problem without the restrictions on f was studied in [25, 27, 29] for uniformly geometrically ergodic Feller-Markov processes. Such processes converge exponentially fast to their invariant measure and the speed of this convergence is independent of the starting point. Examples are usually constrained to processes on compact state spaces. An attempt to generalise these results was made in [28], where the author assumed that trajectories of the process can be split into excursions with square integrable lengths between two compact sets. In this setting optimality was studied within a narrow class of stopping rules.

The aim of this paper is to analyse the optimal stopping problem (1) for *non-uniformly ergodic* Feller-Markov processes with minimal assumptions on f and g, and a general discount rate r which is only assumed to be non-negative. Specifically, assume that the state of the world is described by a right-continuous time homogeneous Markov process $\big(X(t)\big)$ defined on a locally compact separable space E endowed with a metric ρ with respect to which every closed ball is compact. The Borel σ-algebra on E is denoted by $\mathcal{E}$. Let P_t be the semigroup generated by the process (X_t), i.e., $P_t\phi(x) = \mathbb{E}^x\{\phi(X_t)\}$ for any bounded measurable function $\phi : E \to \mathbb{R}$. The transition measure is given by $P_t(x, A) = \mathbb{P}^x\{X_t \in A\}$ for $A \in \mathcal{E}$.

We make the following standing assumptions:

(A1) (Weak Feller property)
$$P_t\, \mathcal{C}_0 \subseteq \mathcal{C}_0,$$
where $\mathcal{C}_0$ is the space of continuous bounded functions $E \to \mathbb{R}$ vanishing in infinity.

(A2) (X_t) is ergodic, i.e., there is a unique probability measure μ on $\mathcal{E}$ such that for any $x \in E$
$$\lim_{t\to\infty} \|P_t(x, \cdot) - \mu(\cdot)\|_{TV} = 0,$$
where $\|\cdot\|_{TV}$ denotes the total variation norm.

(A3) Functions $f, g : E \to \mathbb{R}$ are continuous and bounded.

Assumption (A1) is commonly imposed and, in fact, necessary, to establish the existence of optimal stopping times for general weak Feller processes (for the necessity see the example at the end of Section 3.1 in [21]). The class of weakly Feller processes (A1) comprises Levy processes [1, Theorem 3.1.9], solutions to stochastic differential equations with continuous bounded coefficients driven by Levy processes (for details see [1, Theorem 6.7.2]) and many diffusions and jump-diffusions with unbounded coefficients encountered in modelling. Notice that a weak Feller process can be assumed to be a standard Markov process, see [8, Theorem 3.14].

The ergodicity assumption (A2) will be used to remove the requirement for the discount rate to be bounded from below by a strictly positive constant. It provides information about the behaviour of the integral part of the functional (1). Indeed, when the function f is not restricted to either positive or negative values, the integral of f with respect to the invariant measure μ determines the solution of the problem. For an undiscounted problem, i.e., $r \equiv 0$, it is optimal never to stop whenever $\mu(f) > 0$. On the other hand, the case of $\mu(f) < 0$ exhibits a much more interesting behaviour which is studied in detail in this paper.

The continuity in assumption (A3) can be relaxed when the process is strongly Feller, for example, for locally non-degenerate diffusions or jump-diffusions.

To formulate Theorem 1.1 below, we need to introduce two notions: of ε-optimal stopping times and strong Markov processes. The first notion is related to the approximation of the supremum in the optimal stopping problem (1): a stopping time τ is called ε-*optimal* if

$$\limsup_{T\to\infty} \mathbb{E}^x\left\{ \int_0^{\tau\wedge T} e^{-\int_0^s r(X_u)du}f(X_s)ds + e^{-\int_0^{\tau\wedge T} r(X_u)du}g(X_{\tau\wedge T})\right\} \geq w(x) - \varepsilon.$$

The semigroup of weakly Feller Markov processes, assumption (A1), may not sufficiently smooth out the payoff to guarantee continuity of the value function w under weaker assumptions. Instead, a stronger notion of strong Feller process can be used. A Markov process (X_t) satisfies the *strong Feller property* if its semigroup P_t, $t > 0$, maps measurable bounded functions into continuous functions. The family of strong Feller processes includes uniformly non-degenerate diffusions, and jump-diffusions with uniformly non-degenerate diffusion term as well as solutions to stochastic differential equations driven by Levy processes with a non-zero Brownian part or with an appropriately high intensity of small jumps (see also Section 6).

The most important of our contributions towards the solution of the stopping problem (1) are distilled in the following theorem:

THEOREM 1.1. *Assume that $\mu(f) < 0$, function r is non-negative, continuous and bounded, and either*

i) there is $K : E \to (0, \infty)$ bounded on compacts and $h : [0, \infty) \to \mathbb{R}_+$ such that

$$\|P_t(x, \cdot) - \mu(\cdot)\|_{TV} < K(x)h(t),$$

and $\int_0^\infty h(t)dt < \infty$,

ii) $\mathbb{E}^x\{K(X_T)\} < \infty$ for any $x \in E$ and $T \geq 0$,

iii) the process (X_t) is strongly Feller, or for any compact set $L \subset E$ there is $\alpha > 0$ such that $\sup_{x\in L,\, T\geq 0} \mathbb{E}^x\{K(X_T)^{1+\alpha}\} < \infty$,

or

i') for any $\delta, \varepsilon > 0$ and a compact set $K \subset E$ there is $N > 0$ and $p > 0$ such that for $n \geq N$ and $x \in K$ we have

$$\mathbb{P}^x\left\{\left|\frac{1}{\int_0^{n\delta} e^{-\int_0^s r(X_u)du}ds} \int_0^{n\delta} e^{-\int_0^s r(X_u)du}\big(f(X_s) - \mu(f)\big)ds\right| > \varepsilon\right\} \leq e^{-p(n\delta)}.$$

3

Then the stopping time $\tau_\varepsilon = \inf\{t \geq 0 : g(X_t) \geq w(X_t) - \varepsilon\}$ is ε-optimal for any $\varepsilon > 0$, the value function w is continuous and satisfies, in a viscosity sense, the variational inequality

$$\min\big(- \mathcal{A}w + rw - f, w - g\big) = 0,$$

where $\mathcal{A}$ is the infinitesimal generator of (X_t).

If, additionally, $R = \lim_{t\to\infty} \int_0^t r(X_s)ds < \infty$ $\mathbb{P}^x$-a.s. *(for example, $r \equiv 0$) then the stopping time $\tau^* = \inf\{t \geq 0 : g(X_t) \geq w(X_t)\}$ is optimal for w.*

REMARK 1.2. For a stopping problem with the total reward criterion

$$w(x) = \sup_\tau \limsup_{T\to\infty} \mathbb{E}^x\Big\{ \int_0^{\tau\wedge T} f(X_s)ds + g(X_{\tau\wedge T})\Big\}$$

Theorem 1.1 takes the following form: *Assume that $\mu(f) < 0$ and either (i)-(iii) hold, or*

i') *for any $\delta, \varepsilon > 0$ and a compact set $K \subset E$ there is $N > 0$ and $p > 0$ such that for $n \geq N$ and $x \in K$ we have*

$$\mathbb{P}^x\Big\{ \Big| \frac{1}{n\delta} \int_0^{n\delta} f(X_s)ds - \mu(f)\Big| > \varepsilon \Big\} \leq e^{-p(n\delta)}.$$

Then the value function w is continuous and satisfies, in a viscosity sense, the variational inequality

$$\min\big(- \mathcal{A}w - f, w - g\big) = 0.$$

Moreover, τ^* is optimal for w. ◁

Examples of processes satisfying (i)-(iii) are given in Section 6. They include a large class of (non-uniformly) exponentially or sub-exponentially ergodic diffusions (e.g., polynomially ergodic), including an Ornstein-Uhlenbeck process. Assumption (i') is mostly applicable in the undiscounted case ($r \equiv 0$) to processes on compact state spaces (for example, for reflected diffusions). It then follows from the weak Feller property (A1).

Theorem 1.1 combines selected results from Sections 2-4. For the undiscounted problem ($r \equiv 0$), conditions (i)-(ii) or (i') imply the main assumption (B1) at the beginning of Subsection 2.1 through results in Subsections 2.2 and 2.3. Theorem 2.11 shows that τ^* is an optimal stopping time. Continuity of w is proved in Subsection 2.4. The case with a general discount rate r is studied in Section 3. ε-optimality of τ_ε is demonstrated in Theorem 3.7. Theorem 3.11 shows that w is continuous. Optimality of τ^* follows from Theorem 3.8. Variational characterisation of the value function is established in Section 4.

The structure of the paper is as follows. Section 2 presents results for the undiscounted functional, $r \equiv 0$. These results are generalised to an arbitrary discount rate in Section 3. Variational characterisation of the value function is presented in Section 4. In Section 5 we show that when $\mu(r) > 0$ and a certain upper bound for the large deviations of the empirical process $r(X_t)$ holds, then the stopping problem (1) inherits many properties from the problem with a constant positive discount rate, in particular, the value function is finite for an arbitrary function f (not necessarily satisfying $\mu(f) < 0$).

4

2. Undiscounted functional

The study of the stopping problem (1) commences with a special case when the discount rate $r(x)$ is zero:

$$w(x) = \sup_{\tau} \limsup_{T\to\infty} \mathbb{E}^x \left\{ \int_0^{\tau \wedge T} f(X_s)ds + g(X_{\tau \wedge T}) \right\}. \tag{2}$$

This particular stopping problem will highlight main differences between a standard discounted functional and the functional with the discount rate that is not uniformly separated from zero. To appreciate the difficulty of the problem notice that by the law of large numbers for ergodic processess (see Lemma 2.1 below) we have

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t f^+(X_s)ds = \mu(f^+) \qquad \text{and} \qquad \lim_{t\to\infty} \frac{1}{t} \int_0^t f^-(X_s)ds = \mu(f^-)$$

$\mathbb{P}^x$-a.s. for all $x \in E$. When both the positive part f^+ of f and the negative part f^- of f are μ-a.s. non-zero then the integral $\int_0^\infty f(X_s)ds$ is undefined because it is a difference of two infinities. Ergodicity of the process (X_t) introduces a delicate trade-off between these integrals ensuring that one infinity "cancels" the other.

LEMMA 2.1. *Under assumption (A2), for any measurable function $h : E \to \mathbb{R}$ such that $\mu(|h|) < \infty$, we have for μ almost all $x \in E$*

$$\frac{1}{t} \int_0^t h(X_s)ds \to \mu(h) \qquad \mathbb{P}^x\text{-a.s. as } t \to \infty,$$

where $\mu(h)$ denotes $\int_E h(x)\mu(dx)$.

The proof of the above result follows from the law of large numbers for stationary processes [6, Chapter X, Theorem 2.1]; for details see [19, Theorem 17.1.2].

The solution of the stopping problem (2) is determined by the value $\mu(f)$. The case $\mu(f) > 0$ is trivial: it is optimal *not* to stop and the value function is infinite:

LEMMA 2.2. *Assume (A2). If $\mu(f) > 0$ then $w(x) = \infty$ and it is optimal to take $\tau = \infty$, $\mathbb{P}^x$-a.s. for all $x \in E$.*

Proof. The ergodicity of the process (X_t) implies that $\lim_{t\to\infty} P_t f(x) = \mu(f) > 0$, so for $\tau = \infty$ we have

$$\limsup_{T\to\infty} \mathbb{E}^x \left\{ \int_0^T f(X_s)ds + g(X_T) \right\} \geq \limsup_{T\to\infty} \int_0^T P_s f(x)ds - \|g\|_\infty = \infty.$$

$\square$

The case $\mu(f) < 0$ is the main topic of this paper and results in a much more interesting behaviour. In particular, the value function is often finite. The case of $\mu(f) = 0$ is a grey area – a transition point between interesting and trivial results. Our methods do not allow us to provide a thorough study of this case. However, a glimpse at possible results can be caught if one takes $f = 0$. Obviously, the value function is finite, but the existence of an optimal stopping time depends on the interplay between the process and the function g as the following example shows.

EXAMPLE. Let (X_t) be an Ornstein-Uhlenbeck process

$$dX_t = \theta(\mu - X_t)dt + \sigma dW_t.$$

Take $f \equiv 0$ and $g(x) = \arctan(x)$. Obviously, $w(x) \leq \max_x g(x) = \pi$. Take a stopping time $\tau_\varepsilon = \inf\{t \geq 0 : X_t \geq K(\varepsilon)\}$, where $K(\varepsilon) = \tan(\pi - \varepsilon)$. This stopping time is ε-optimal because $g(X_{\tau_\varepsilon}) \geq \pi - \varepsilon \geq w(x) - \varepsilon$. This also proves that $w(x) = \pi$. However, there is no optimal stopping time as the objective function $g(x)$ grows when x increases. ◁

2.1. Optimal stopping times

The developments of this subsection are pursued under the following additional assumption:

(B1) For any $x \in E$, there is $d(x) < 0$ such that

$$\gamma(x) = \sup_\tau \limsup_{T \to \infty} \mathbb{E}^x\left\{ \int_0^{\tau \wedge T} \big(f(X_s) - d(x)\big)ds \right\} < \infty.$$

Sufficient conditions will be provided in the following subsections.

Although the above assumption does not state explicitly that $\mu(f) < 0$, this is a necessary condition. If $\mu(f) \geq 0$ then $\mu(f) - d(x) > 0$ and by ergodicity

$$\lim_{t \to \infty} \mathbb{E}^x\{f(X_t) - d(x)\} = \mu(f) - d(x) > 0,$$

which implies $\gamma(x) = \infty$.

REMARK 2.3. It is vital for applications to notice that if assumption (B1) is satisfied for a function f, it is also satisfied for any function $f' \leq f$. This trivial observation greatly extends the applicability of sufficient conditions for (B1) discussed in Subsections 2.2 and 2.3. ◁

Assumption (B1) allows us to restrict the set of stopping times in (2) to those with finite expectation.

LEMMA 2.4. *For every integrable ε-optimal stopping time σ we have*

$$\mathbb{E}^x\{\sigma\} \leq \frac{\gamma(x) + 2\|g\| + \varepsilon}{-d(x)}.$$

Proof. Since σ is integrable, we have

$$\mathbb{E}^x\left\{ \int_0^\sigma f(X_s)ds + g(X_\tau) \right\} = \mathbb{E}^x\left\{ \int_0^\sigma \big(f(X_s) - d(x)\big)ds + d(x)\sigma + g(X_\sigma) \right\}$$
$$\leq \gamma(x) + \|g\| + d(x)\mathbb{E}^x\{\sigma\}.$$

On the other hand,

$$\mathbb{E}^x\left\{ \int_0^\sigma f(X_s)ds + g(X_\tau) \right\} \geq w(x) - \varepsilon \geq -\|g\| - \varepsilon.$$

Combining the above two inequalities,

$$-\|g\| - \varepsilon \leq \gamma(x) + \|g\| + d(x)\mathbb{E}^x\{\sigma\},$$

completes the proof. □

It follows from (2) that the value function can be defined as a supremum over bounded (hence, integrable) stopping times. The above lemma proves that the optimisation can be restricted to those stopping times whose expectation is bounded by

$$M(x) = \frac{\gamma(x) + 2\|g\| + 1}{-d(x)}. \tag{3}$$

Hence, the value function $w(x)$ is bounded from above by $\|f\|M(x) + \|g\|$.

REMARK 2.5. By the dominated convergence theorem and quasi left-continuity of (X_t) (see [8, Theorem 3.13]) whenever $\mathbb{E}^x\{\tau\} < \infty$ we have

$$\limsup_{T\to\infty} \mathbb{E}^x\Big\{ \int_0^{\tau\wedge T} f(X_s)ds + g(X_{\tau\wedge T}) \Big\} = \mathbb{E}^x\Big\{ \int_0^{\tau} f(X_s)ds + g(X_{\tau}) \Big\}.$$

◁

LEMMA 2.6. *The value function w is lower semi-continuous.*

Proof. Let

$$w_T(x) = \sup_{\tau \leq T} \mathbb{E}^x\Big\{ \int_0^{\tau} f(X_s)ds + g(X_{\tau}) \Big\}. \tag{4}$$

Due to the continuity and boundedness of functions f and g as well as the weak Feller property of the underlying process (X_t), the value function w_T is continuous, see [21, Corollary 3.6]. Functions w_T converge pointwise to w and they form an increasing sequence. Hence, w is lower semicontinuous. □

The construction of an optimal stopping time requires a few intermediate steps achieved in lemmas below.

LEMMA 2.7. *For every $x \in E$, there exists an non-decreasing sequence σ_m of bounded $\frac{1}{m}$-optimal stopping times for $w(x)$.*

Proof. The stopping problem $w(x)$ can be equivalently written as a supremum over bounded stopping times. Hence, $w_T(x)$ defined in (4) approximate $w(x)$ from below. By [21, Corollary 3.6], these stopping problems admit optimal solutions of the form

$$\tau_T = \inf\{t \in [0, T] : g(X_t) \geq w_{T-t}(X_t)\}.$$

As w_T are non-decreasing in T, it is easy to see that τ_T are non-decreasing in T as well. By setting $T(m) = \inf\{T \geq 0 : w(x) - w_T(x) \leq \frac{1}{m}\}$, we can set $\sigma_m = \tau_{T(m)}$. □

In what follows we shall denote by (σ_m) the sequence of stopping times from the above lemma.

LEMMA 2.8. *The process $Z_t := \int_0^t f(X_s)ds + w(X_t)$ is a right-continuous $\mathbb{P}^x$-supermartingale for any $x \in E$. Moreover, for a bounded stopping time σ and an arbitrary stopping time τ*

$$\mathbb{E}^x\Big\{ \int_0^{\sigma} f(X_s)ds + g(X_{\sigma}) \Big\} \leq \mathbb{E}^x\Big\{ \int_0^{\tau\wedge\sigma} f(X_s)ds + 1_{\{\sigma<\tau\}} g(X_{\sigma}) + 1_{\{\sigma\geq\tau\}} w(X_{\tau}) \Big\}. \tag{5}$$

Proof. Define $Z_t^T = \int_0^t f(X_s)ds + w_{T-t}(X_t)$, $t \in [0, T]$. This process is a Snell envelope of the process $t \mapsto \int_0^t f(X_s)ds + g(X_t)$, so, in particular, it is a right-continuous supermartingale [22, Theorem 2.4]. These supermartingales are increasing in T as so is w_{T-t}. Notice that $Z_t = \sup_{T \geq t} Z_t^T$ $\mathbb{P}^x$-a.s. Theorem T16 in [18] and the remark following it imply that (Z_t) is a right-continuous supermartingale.

Since σ is a bounded stopping time, the optional sampling theorem yields $\mathbb{E}^x\{Z_\sigma | \mathcal{F}_{\tau \wedge \sigma}\} \leq Z_{\tau \wedge \sigma}$. This reads

$$\int_0^{\tau \wedge \sigma} f(X_s)ds + w(X_{\tau \wedge \sigma}) \geq \mathbb{E}^x\left\{ \int_0^\sigma f(X_s)ds + w(X_\sigma) \Big| \mathcal{F}_{\tau \wedge \sigma} \right\}.$$

Hence,

$$\int_0^{\tau \wedge \sigma} f(X_s)ds + 1_{\{\sigma \geq \tau\}} w(X_{\tau \wedge \sigma}) \geq \mathbb{E}^x\left\{ \int_0^\sigma f(X_s)ds + 1_{\{\sigma \geq \tau\}} w(X_\sigma) \Big| \mathcal{F}_{\tau \wedge \sigma} \right\}$$

$$\geq \mathbb{E}^x\left\{ \int_0^\sigma f(X_s)ds + 1_{\{\sigma \geq \tau\}} g(X_\sigma) \Big| \mathcal{F}_{\tau \wedge \sigma} \right\}.$$

Adding $1_{\{\sigma < \tau\}} g(X_\sigma)$ to both sides completes the proof. $\qquad\square$

Due to the lower semi-continuity of w and the continuity of g, the random variable

$$\tau_\varepsilon = \inf\{t \geq 0 : g(X_t) \geq w(X_t) - \varepsilon\}$$

is a stopping time and $g(X_{\tau_\varepsilon}) \geq w(X_{\tau_\varepsilon}) - \varepsilon$ on $\{\tau_\varepsilon < \infty\}$.

LEMMA 2.9. *For any $x \in E$, we have $\mathbb{P}^x\{\sigma_m < \tau_\varepsilon\} \leq \dfrac{1}{m\varepsilon}$.*

Proof. Apply (5) to σ_m and τ_ε

$$w(x) - \frac{1}{m} \leq \mathbb{E}^x\left\{ \int_0^{\sigma_m} f(X_s)ds + g(X_\sigma) \right\}$$

$$\leq \mathbb{E}^x\left\{ \int_0^{\sigma_m \wedge \tau_\varepsilon} f(X_s)ds + 1_{\{\sigma_m < \tau_\varepsilon\}} g(X_{\sigma_m}) + 1_{\{\sigma_m \geq \tau_\varepsilon\}} w(X_{\tau_\varepsilon}) \right\}$$

$$\leq \mathbb{E}^x\left\{ \int_0^{\sigma_m \wedge \tau_\varepsilon} f(X_s)ds + 1_{\{\sigma_m < \tau_\varepsilon\}} \left(w(X_{\sigma_m}) - \varepsilon \right) + 1_{\{\sigma_m \geq \tau_\varepsilon\}} w(X_{\tau_\varepsilon}) \right\}$$

$$\leq \mathbb{E}^x\left\{ \int_0^{\sigma_m \wedge \tau_\varepsilon} f(X_s)ds + w(X_{\sigma_m \wedge \tau_\varepsilon}) \right\} - \varepsilon \mathbb{P}^x\{\sigma_m < \tau_\varepsilon\}$$

$$\leq w(x) - \varepsilon \mathbb{P}^x\{\sigma_m < \tau_\varepsilon\},$$

where the third inequality follows from the observation that $g(X_{\sigma_m}) < w(X_{\sigma_m}) - \varepsilon$ on $\{\sigma_m < \tau_\varepsilon\}$ and the last inequality is the result of (5) applied for $\tau \equiv 0$ and $\sigma = \sigma_m \wedge \tau_\varepsilon$. $\qquad\square$

LEMMA 2.10. *Let (X_n) be a sequence of positive random variables with $\mathbb{E} X_n \leq 1$. If Y is a positive random variable (with a possibly infinite expectation) and $\lim_{n \to \infty} \mathbb{P}\{Y > X_n\} = 0$ then $\mathbb{E} Y \leq 1$.*

Proof. Assume first that $Y \leq K$ a.s. for some constant $K > 0$. Then

$$\mathbb{E}\{Y\} = \mathbb{E}\{1_{\{Y \leq X_n\}} Y\} + \mathbb{E}\{1_{\{Y > X_n\}} Y\}$$

$$\leq \mathbb{E}\{1_{\{Y \leq X_n\}} X_n\} + \mathbb{E}\{1_{\{Y > X_n\}} Y\}$$

$$\leq \mathbb{E}\{X_n\} + K\mathbb{P}\{Y > X_n\} \leq 1 + K\mathbb{P}\{Y > X_n\}$$

and the last term tends to 0 as $n \to \infty$. Take now an arbitrary Y satisfying assumptions of the lemma. For each $K > 0$, we have $\lim_{n \to \infty} \mathbb{P}\{Y \wedge K > X_n\} = 0$. Previous considerations imply $\mathbb{E}\{Y \wedge K\} \leq 1$. By monotone convergence, $\mathbb{E}\{Y\} = \lim_{K \to \infty} \mathbb{E}\{Y \wedge K\} \leq 1$. $\qquad\square$

THEOREM 2.11. *Under assumptions (A1)-(A3) and (B1) an optimal stopping time is given by*

$$\tau^* = \inf\{t \geq 0 : g(x) \geq w(x)\}.$$

Moreover, $\mathbb{E}^x\{\tau^*\} \leq M(x)$, where $M(x)$ is defined in (3).

Proof. By Lemma 2.4, for every m the stopping time σ_m has a bounded expectation: $\mathbb{E}^x\{\sigma_m\} \leq M(x)$, where $M(x)$ is defined in (3). This bound does not depend on m. Combining this with conclusions of Lemma 2.9 and Lemma 2.10 (with $Y = \tau_\varepsilon/M(x)$ and $X_n = \sigma_n/M(x)$) yields that $\mathbb{E}^x\{\tau_\varepsilon\} \leq M(x)$.

Apply (5) to σ_m and τ_ε and notice that $w(X_{\tau_\varepsilon}) - \varepsilon \leq g(X_{\tau_\varepsilon})$ due to the continuity of g, lower semicontinuity of w and right-continuity of the process (X_t):

$$
\begin{aligned}
w(x) - \frac{1}{m} &\leq \mathbb{E}^x\Big\{ \int_0^{\sigma_m} f(X_s)ds + g(X_{\sigma_m}) \Big\} \\
&\leq \mathbb{E}^x\Big\{ \int_0^{\sigma_m \wedge \tau_\varepsilon} f(X_s)ds + 1_{\{\sigma_m < \tau_\varepsilon\}} g(X_{\sigma_m}) + 1_{\{\sigma_m \geq \tau_\varepsilon\}} w(X_{\tau_\varepsilon}) \Big\} \\
&\leq \mathbb{E}^x\Big\{ \int_0^{\sigma_m \wedge \tau_\varepsilon} f(X_s)ds + 1_{\{\sigma_m < \tau_\varepsilon\}} g(X_{\sigma_m}) + 1_{\{\sigma_m \geq \tau_\varepsilon\}} \big(g(X_{\tau_\varepsilon}) + \varepsilon\big) \Big\}.
\end{aligned}
$$

When $m \to \infty$, the dominated convergence theorem implies

$$w(x) \leq \mathbb{E}^x\Big\{ \int_0^{\tau_\varepsilon} f(X_s)ds + g(X_{\tau_\varepsilon}) + \varepsilon \Big\},$$

so τ_ε is ε-optimal. Put $\tau_0 = \lim_{\varepsilon \to 0} \tau_\varepsilon$. Since this is an increasing sequence of random variables with the expectation bounded by $M(x)$ then τ_0 is well defined and $\mathbb{E}^x\{\tau_0\} \leq M(x)$. For $\eta > 0$ and any $0 < \varepsilon \leq \eta$ we have $g(X_{\tau_\varepsilon}) \geq w(X_{\tau_\varepsilon}) - \eta$. By the quasi left-continuity of (X_t) and lower semicontinuity of w, taking the limit $\varepsilon \to 0$ and then $\eta \to 0$ yields

$$g(X_{\tau_0}) \geq w(X_{\tau_0}).$$

So $\tau_0 \geq \tau^*$. This means that τ^* is finite and $\mathbb{E}^x\{\tau^*\} \leq M(x)$. Its optimality follows in the same way as ε-optimality of τ_ε. $\qquad\square$

The stopping time τ^* can be characterised as a limit of optimal stopping times for problems with a finite stopping horizon when the horizon increases to infinity.

LEMMA 2.12. *We have $\tau^* = \lim_{T \to \infty} \tau_T$, where τ_T is an optimal stopping time for the problem* (4).

Proof. Recall that $\tau_T = \inf\{t \geq 0 : g(X_t) \geq w_{T-t}(X_t)\}$. The sequence $(\tau_T)_{T \geq 0}$ of random variables is increasing and each element has its expectation bounded by $M(x)$. Hence, $\tau_\infty = \lim_{T \to \infty} \tau_T$ exists. Since $\tau_T \leq \tau^*$, we conclude that $\tau_\infty \leq \tau^*$. We also know that $g(X_{\tau_T}) = w_{T-\tau_T}(X_{\tau_T})$, so $\lim_{T \to \infty} g(X_{\tau_T}) = \lim_{T \to \infty} w_{T-\tau_T}(X_{\tau_T})$. By the quasi left-continuity of (X_t) we have $\lim_{T \to \infty} g(X_{\tau_T}) = g(X_{\tau_\infty})$. For each $\omega \in \Omega$, any $T > \tau^*(\omega)$ and $0 < d < T - \tau_T(\omega)$ we have $w_{T-\tau_T}(X_{\tau_T}) \geq w_d(X_{\tau_T})$ (dependence on ω is omitted here for clarity of notation). Hence, for each positive d we have $\liminf_{T \to \infty} w_{T-\tau_T}(X_{\tau_T}) \geq w_d(X_{\tau_\infty})$ $\mathbb{P}^x$-a.s. and letting $d \to \infty$ we obtain $\liminf_{T \to \infty} w_{T-\tau_T}(X_{\tau_T}) \geq w(X_{\tau_\infty})$. Finally, $g(X_{\tau_\infty}) \geq w(X_{\tau_\infty})$, P^x - a.s., so $\tau_\infty \geq \tau^*$. $\qquad\square$

Summarising, under the assumptions of this section we showed that the value function $w(x)$ is lower semi-continuous and the stopping time τ^*, the first time the process enters the closed set $\{x \in E : g(x) \geq w(x)\}$, has a finite expectation and is optimal.

2.2. Zero-potential bounded from below

This and the following subsection provide sufficient conditions for (B1). Here, they are first expressed in terms of a zero-potential

$$q(x) = \limsup_{T \to \infty} \mathbb{E}^x \left\{ \int_0^T \big(f(X_s) - \mu(f) \big) ds \right\}.$$

They are later derived from certain conditions related to the speed of convergence of transition probabilities of the process (X_t) to the invariant measure μ.

(C1) $q(x)$ is continuous and bounded from below

(C2) for any bounded stopping time σ

$$q(x) = \mathbb{E}^x \left\{ \int_0^\sigma \big(f(X_s) - \mu(f) \big) ds + q(X_\sigma) \right\}.$$

(C3) $\mu(f) < 0$

REMARK 2.13. Assumption (C2) is equivalent to requesting that the process $Z_t = \int_0^t \big(f(X_s) - \mu(f) \big) ds + q(X_t)$ is a martingale. ◁

REMARK 2.14. Under assumption (C2), for any bounded stopping time τ

$$\mathbb{E}^x \left\{ \int_0^\tau f(X_s) ds + g(X_\tau) \right\} = q(x) + \mathbb{E}^x \{ \mu(f)\tau + (g - q)(X_\tau) \}.$$

The optimal stopping problem (2) can be equivalently written as

$$\hat{w}(x) := w(x) - q(x) = \sup_\tau \limsup_{T \to \infty} \mathbb{E}^x \{ \mu(f)(\tau \wedge T) + (g - q)(X_{\tau \wedge T}) \}. \tag{6}$$

This transforms the problem with a possibly unbounded integral term into a problem with an unbounded terminal reward and a strictly negative running cost. ◁

REMARK 2.15. The limit $\limsup_{T \to \infty}$ can be omitted in the definition of $\hat{w}$ and w and the set of stopping times can be restricted to those with a finite expectation. Indeed, if $\mathbb{E}^x \{\tau\} = \infty$, then

$$\limsup_{T \to \infty} \mathbb{E}^x \{ \mu(f)(\tau \wedge T) + (g - q)(X_{\tau \wedge T}) \} \leq \mu(f) \lim_{T \to \infty} \mathbb{E}^x \{\tau \wedge T\} + \Gamma = -\infty,$$

where $\Gamma = \sup_x (g - q)(x) < \infty$ by assumption (C1). ◁

The following lemma shows that the assumptions introduced at the beginning of this subsection are sufficient for (B1).

LEMMA 2.16. *Under (C1)-(C3), assumption (B1) is satisfied with any $d(x) \in (\mu(f), 0)$. Moreover, $\gamma(x) \leq q(x) - A$, where $A = \inf_y q(y)$.*

Proof. Take any stopping time τ and $T > 0$. Then

$$\mathbb{E}^x\Big\{\int_0^{\tau \wedge T} \big(f(X_s) - d(x)\big)ds\Big\}$$

$$= \mathbb{E}^x\Big\{\int_0^{\tau \wedge T} \big(f(X_s) - \mu(f)\big)ds + \big(\mu(f) - d(x)\big)(\tau \wedge T)\Big\}$$

$$\leq \mathbb{E}^x\Big\{\int_0^{\tau \wedge T} \big(f(X_s) - \mu(f)\big)ds\Big\}$$

$$= q(x) - \mathbb{E}^x\{q(X_{\tau \wedge T})\} \leq q(x) - A,$$

where A is a lower bound for q and in the first inequality we used $\mu(f) - d(x) < 0$. From the arbitrariness of τ and T we conclude that $\gamma(x) \leq q(x) - A < \infty$. $\qquad\square$

The existence and continuity of the zero-potential is ensured by (D1) below, while the integrability required in (C2) follows from (D2).

(D1) there is $K : E \to (0, \infty)$ bounded on compacts and $h : [0, \infty) \to \mathbb{R}_+$ such that

$$\|P_t(x, \cdot) - \mu(\cdot)\|_{TV} < K(x)h(t),$$

and $\int_0^\infty h(t)dt < \infty$,

(D2) (integrability) $\mathbb{E}^x\{K(X_T)\} < \infty$ for any $x \in E$ and $T \geq 0$.

REMARK 2.17. Ergodicity assumption (D1) is commonly studied in the theory of stochastic processes (see, e.g., references in Section 6):

1. (geometric ergodicity) there are functions $K, \lambda : E \to (0, \infty)$ bounded on compacts such that

$$\|P_t(x, \cdot) - \mu(\cdot)\|_{TV} < K(x)e^{-\lambda(x)t},$$

2. (polynomial ergodicity) there is a function $K : E \to (0, \infty)$ bounded on compacts and a constant $p > 1$ such that

$$\|P_t(x, \cdot) - \mu(\cdot)\|_{TV} < \frac{K(x)}{(1+t)^p}.$$

$\triangleleft$

LEMMA 2.18. *Assume (D1) and (D2). Then the zero-potential q is finite-valued and continuous, and assumption (C2) is satisfied.*

Proof. Define

$$\bar{q}(x) = \Big\{\int_0^\infty \big|P_s f(x) - \mu(f)\big|ds\Big\}.$$

This function is finite-valued by assumption (D1). This means that $q(x)$ is finite valued and continuous by the weak Feller property (and dominated convergence theorem). By (D2), for any $T > 0$ and $x \in E$, the random variable $\bar{q}(X_T)$ is $\mathbb{P}^x$-integrable.

Denote $\hat{f}(x) = f(x) - \mu(f)$. For $\alpha > 0$ and $T > 0$, due to the boundedness of $\hat{f}$ and Markov property of (X_t) we have

$$\mathbb{E}^x\Big\{\int_0^T e^{-\alpha s}\hat{f}(X_s)ds\Big\} = q_\alpha(x) - e^{-\alpha T}\mathbb{E}^x\{q_\alpha(X_T)\}, \qquad (7)$$

where

$$q_\alpha(x) = \mathbb{E}^x\Big\{ \int_0^\infty e^{-\alpha s}\hat{f}(X_s)ds\Big\}.$$

Clearly, $|q_\alpha(x)| \le \bar{q}(x)$, so $|q_\alpha(X_T)|$ is dominated by $\bar{q}(X_T)$. The integrability of the latter and the dominated convergence theorem yield that when $\alpha \to 0$ the equation (7) converges to

$$\mathbb{E}^x\Big\{ \int_0^T \hat{f}(X_s)ds\Big\} = q(x) - \mathbb{E}^x\{q(X_T)\}.$$

Take now a stopping time σ bounded by T. Then

$$|q_\alpha(X_\sigma)| = \Big|\mathbb{E}^x\Big\{ \int_\sigma^\infty e^{-\alpha(s-\sigma)}\hat{f}(X_s)ds\Big|\mathcal{F}_\sigma\Big\}\Big|$$

$$\le \Big|\mathbb{E}^x\Big\{ \int_\sigma^T e^{-\alpha(s-\sigma)}\hat{f}(X_s)ds\Big|\mathcal{F}_\sigma\Big\}\Big| + \Big|e^{-\alpha(T-\sigma)}\mathbb{E}^x\Big\{ \int_T^\infty e^{-\alpha(s-T)}\hat{f}(X_s)ds\Big|\mathcal{F}_\sigma\Big\}\Big|$$

$$\le T\|\hat{f}\| + \Big|\mathbb{E}^x\Big\{ e^{-\alpha(T-\sigma)}\mathbb{E}^x\Big\{ \int_T^\infty e^{-\alpha(s-T)}\hat{f}(X_s)ds\Big|\mathcal{F}_T\Big\}\Big|\mathcal{F}_\sigma\Big\}\Big|$$

$$= T\|\hat{f}\| + e^{-\alpha(T-\sigma)}\Big|\mathbb{E}^x\{q_\alpha(X_T)|\mathcal{F}_\sigma\}\Big| \le T\|\hat{f}\| + \Big|\mathbb{E}^x\{\bar{q}(X_T)|\mathcal{F}_\sigma\}\Big|,$$

which gives a bound for $|q_\alpha(X_\sigma)|$ independent of α. Strong Markov property gives

$$\mathbb{E}^x\Big\{ \int_0^\sigma e^{-\alpha s}\hat{f}(X_s)ds\Big\} = q_\alpha(x) - \mathbb{E}^x\{e^{-\alpha\sigma}q_\alpha(X_\sigma)\}. \tag{8}$$

Using the bound obtained above, we can take the limit $\alpha \to 0$ on both sides and conclude the proof. $\square$

To ensure that q is bounded from below, we have to assume something about the function f. Knowing that under ergodicity, the process spends most of the time in a ball with an appropriately big radius, we want to allow for flexibility of the function f there. However, to give us control of the integral the effect of f should be under control outside of the ball. The following lemma shows that it is sufficient to assume that $f > \mu(f)$ outside of a compact set.

LEMMA 2.19. *Assume (D1)-(D2). If $\mu(f) < 0$ and the set $L = \{x \in E : f(x) \le \mu(f)\}$ is compact, then q is bounded from below.*

Proof. Let B be a closed ball such that $\mu(B) > 0$. It is easy to see (for example, by the law of large numbers for ergodic processes) that $\tau_B < \infty$ $\mathbb{P}^x$-a.s. for any $x \in E$, where $\tau_B = \inf\{t \ge 0 : X_t \in B\}$. Let $\tilde{B}$ be the smallest closed ball containing B and L. Then $\tau_{\tilde{B}} < \infty$ $\mathbb{P}^x$-a.s. for $x \in E$. We have $f(x) - \mu(f) > 0$ for $x \in \tilde{B}^c$. For any $T > 0$, equation (8) reads

$$q_\alpha(x) = \mathbb{E}^x\Big\{ \int_0^{\tau_{\tilde{B}}\wedge T} e^{-\alpha s}\big(f(X_s) - \mu(f)\big)ds\Big\} + \mathbb{E}^x\{e^{-\alpha(\tau_{\tilde{B}}\wedge T)}q_\alpha(X_{\tau_{\tilde{B}}\wedge T})\}$$

$$\ge \mathbb{E}^x\{e^{-\alpha(\tau_{\tilde{B}}\wedge T)}q_\alpha(X_{\tau_{\tilde{B}}\wedge T})\}.$$

By the dominated convergence theorem and finiteness of $\tau_{\tilde{B}}$ we obtain

$$q_\alpha(x) \ge \mathbb{E}^x\{e^{-\alpha\tau_{\tilde{B}}}q_\alpha(X_{\tau_{\tilde{B}}})\}. \tag{9}$$

12

Considerations in the proof of Lemma 2.18 imply that $q_\alpha(x)$ converges to $q(x)$ when $\alpha \downarrow 0$, but they do not apply to the right-hand side of the above inequality since $\tau_{\tilde{B}}$ does not have to be bounded. However, $X_{\tau_{\tilde{B}}} \in \tilde{B}$ by the right-continuity of (X_t), so it is sufficient to prove uniform convergence of q_α to q on compact sets. We have the following estimate:

$$
\begin{aligned}
|q(x) - q_\alpha(x)| &\leq \int_0^\infty \left(1 - e^{-\alpha s}\right)|\mathbb{E}^x\{f(X_s) - \mu(f)\}ds \\
&\leq \int_0^\infty \left(1 - e^{-\alpha s}\right)\|f\|\|P_t(x, \cdot) - \mu(\cdot)\|_{TV}ds \\
&\leq \|f\|K(x)\int_0^\infty \left(1 - e^{-\alpha s}\right)h(s)ds.
\end{aligned}
$$

By the dominated convergence theorem, this proves uniform convergence of q_α to q as $\alpha \downarrow 0$ on compact sets since K is bounded on compacts. Again, by the dominated convergence theorem, (9) yields

$$
q(x) \geq \mathbb{E}^x\{q(X_{\tau_{\tilde{B}}})\}.
$$

The function q is continuous hence bounded on $\tilde{B}$, so $q(x) \geq \inf_{y \in \tilde{B}} q(y) > -\infty$. □

REMARK 2.20. It is easy to see that the set L from the above lemma is compact when $f \in \mathcal{C}_0$. This compactness property is also true for any continuous function f with $\mu(f) < 0$ and bounded from below by a $\mathcal{C}_0$ function ϕ. Indeed, $L \subset \{x \in E : \phi(x) \leq \mu(f)\} = L_\phi$ and the set L_ϕ is compact since $\mu(f) < 0$. Examples include functions f that are positive outside of a sufficiently large ball but with negative values in a set of a large μ-measure. ◁

It is not clear yet that q can be unbounded from above and bounded from below. The lemma below shows that this is natural for a large class of processes.

LEMMA 2.21. *Assume (D1)-(D2). If $\mu(f) < 0$ and there is $\delta > 0$ such that the set $L_\delta = \{x \in E : f(x) \leq \mu(f) + \delta\}$ is compact, then $\lim_{\|x\|\to\infty} q(x) = \infty$ if and only if $\lim_{\|x\|\to\infty} \mathbb{E}^x\{\tau_{B_n}\} = \infty$ for all n, where $B_n = \{x \in E : \|x\| \leq n\}$.*

Proof. Assume $\lim_{\|x\|\to\infty} \mathbb{E}^x\{\tau_{B_n}\} = \infty$ for all n. Let N be such that $L_\delta \subset B_N$. As in the proof of Lemma 2.19, for any $\alpha > 0$

$$
\begin{aligned}
q_\alpha(x) &= \mathbb{E}^x\left\{ \int_0^{\tau_{B_N}} e^{-\alpha s}\left(f(X_s) - \mu(f)\right)ds\right\} + \mathbb{E}^x\{e^{-\alpha\tau_{B_N}}q_\alpha(X_{\tau_{B_N}})\} \\
&> \delta\mathbb{E}^x\left\{\int_0^{\tau_{B_N}} e^{-\alpha s}ds\right\} + \mathbb{E}^x\{e^{-\alpha\tau_{B_N}}q_\alpha(X_{\tau_{B_N}})\}.
\end{aligned}
$$

By the monotone convergence theorem $\lim_{\alpha\downarrow 0} \mathbb{E}^x\{\int_0^{\tau_{B_N}} e^{-\alpha s}ds\} = \mathbb{E}^x\{\tau_{B_N}\}$. In the proof of Lemma 2.19 we have shown that q_α converges to q uniformly on compact sets. Hence,

$$
q(x) \geq \delta\mathbb{E}^x\{\tau_{B_N}\} + \inf_{y \in B_N} q(y)
$$

and the proof is complete.

Assume now that $\lim_{\|x\|\to\infty} q(x) = \infty$ and take any $N > 0$. Then for $x \notin B_N$

$$
q_\alpha(x) \leq 2\|f\|\mathbb{E}^x\left\{ \int_0^{\tau_{B_N}} e^{-\alpha s}ds\right\} + \mathbb{E}^x\left\{e^{-\alpha\tau_{B_N}}q_\alpha(X_{\tau_{B_N}})\right\},
$$

which, when $\alpha \downarrow 0$, yields

$$q(x) \leq 2\|f\|\mathbb{E}^x\{\tau_{B_N}\} + \sup_{y \in B_N} q(y).$$

Hence, $\lim_{\|x\| \to \infty} \mathbb{E}^x\{\tau_{B_N}\} = \infty$. $\qquad\square$

Notice that the compactness of L_δ is not required in the proof of the left implication.

The above lemma suggests another criterion for the boundedness of q from below. The proof is omitted.

LEMMA 2.22. *If q is continuous and $q(x) \to \infty$ when $\|x\| \to \infty$, then q is bounded from below.*

2.3. Large deviations type condition

The second criterion for (B1) is provided by an upper bound for large deviations of the empirical process:

(L) For any $\delta, \varepsilon > 0$ and a compact set $K \subset E$ there is $N > 0$ and $p > 0$ such that for $n \geq N$ and $x \in K$ we have

$$\mathbb{P}^x\left\{\left|\frac{1}{n\delta}\int_0^{n\delta} f(X_s)ds - \mu(f)\right| > \varepsilon\right\} \leq e^{-p(n\delta)}.$$

There are two important classes of Markov processes that satisfy assumption (L):

- weakly Feller processes over a compact state space with a unique invariant measure (see [7, Theorem 3] and also [4, 5]),

- weakly Feller processes on a locally compact separable matric space with a unique invariant measure for which there exists a Lyapunov function, i.e., a function $u \geq 1$ such that for each positive m the set $\{x \in E : \frac{u(x)}{P_\delta u(x)} \leq m\}$ is compact (for details see [7, Theorem 3] and Lemmas 4.1, 4.3 and Theorem 4.4 of [5] and also [4] and [15]).

Note that the first of the above condition means that if the state space E is compact then our standing assumptions (A1) and (A2) are sufficient for (L) to hold.

LEMMA 2.23. *Assume (L). For any $\varepsilon > 0$ and a compact set $K \subset E$ there is $S > 0$, $C > 0$ and $\rho > 0$ such that for $t \geq S$ and $x \in K$ we have*

$$\mathbb{P}^x\left\{\sup_{s \geq t}\left|\frac{1}{s}\int_0^s f(X_u)du - \mu(f)\right| > \varepsilon\right\} \leq Ce^{-\rho t}.$$

Proof. Fix a compact set $K \subset E$ and put $\delta = \varepsilon/(2\|f - \mu(f)\|)$. Take N and p from (assumption (L)) such that for $n \geq N$ and $x \in K$ we have

$$\mathbb{P}^x\left\{\left|\frac{1}{n\delta}\int_0^{n\delta} f(X_s)ds - \mu(f)\right| > \varepsilon/2\right\} \leq e^{-p(n\delta)}.$$

The boundedness of f and the choice of δ imply that for $n \geq N$

$$\left\{\sup_{s \geq n\delta}\left|\frac{1}{s}\int_0^s f(X_u)du - \mu(f)\right| > \varepsilon\right\} \subset \left\{\sup_{i \in \mathbb{N}}\left|\frac{1}{(n+i)\delta}\int_0^{(n+i)\delta} f(X_u)du - \mu(f)\right| > \varepsilon/2\right\},$$

14

where $\mathbb{N}$ denotes the set of non-negative integers. The probability of the event on the right-hand side is bounded from above by

$$\sum_{i=0}^{\infty} \mathbb{P}^x \left\{ \left| \frac{1}{(n+i)\delta} \int_0^{(n+i)\delta} f(X_s)ds - \mu(f) \right| > \varepsilon/2 \right\} \leq \sum_{i=0}^{\infty} e^{-p((n+i)\delta)} = e^{-p(n\delta)} \frac{1}{1 - e^{-p\delta}}.$$

This proves the statement of the lemma for $t = n\delta$, $n \geq N$, and $\rho = p$. Extending it to any $t \geq N\delta =: S$ is straightforward due to the boundedness of f. $\qquad\square$

Notice that ρ in the above lemma corresponds to p in assumption (L) for $\varepsilon/2$.

LEMMA 2.24. *Assumption (L) and $\mu(f) < 0$ imply (B1).*

Proof. Choose $0 < \varepsilon < -\mu(f)$ and fix $x \in E$. By Lemma 2.23 there is $S > 0$ such that for $t \geq S$ we have $\mathbb{P}^x\{A_t\} \leq Ce^{-\rho t}$, where

$$A_t = \left\{ \sup_{s \geq t} \left| \frac{1}{s} \int_0^s f(X_u)du - \mu(f) \right| > \varepsilon \right\}.$$

Then for any integrable stopping time τ we have

$$\mathbb{E}^x \left\{ \int_0^{\tau} (f(x_s) - \mu(f))ds \right\} = \sum_{i=0}^{\infty} \mathbb{E}^x \left\{ 1_{i \leq \tau < i+1} \int_0^{\tau} (f(x_s) - \mu(f))ds \right\}$$

$$\leq \varepsilon \mathbb{E}^x \tau + 2(\lfloor S \rfloor + 1)\|f\| + \sum_{i=\lfloor S \rfloor + 1}^{\infty} 2\|f\|(i+1)\mathbb{P}^x\{A_i\}$$

$$\leq \varepsilon \mathbb{E}^x \tau + 2(\lfloor S \rfloor + 1)\|f\| + \sum_{i=\lfloor S \rfloor + 1}^{\infty} 2\|f\|(i+1)Ce^{-i\rho}$$

$$= \varepsilon \mathbb{E}^x \{\tau\} + \bar{C},$$

This is equivalent to

$$\mathbb{E}^x \left\{ \int_0^{\tau} \left(f(x_s) - \mu(f) - \varepsilon \right)ds \right\} \leq \bar{C}.$$

Hence, (B1) holds with $d(x) = \mu(f) + \varepsilon$ which, by the choice of ε, is strictly negative. $\qquad\square$

REMARK 2.25. The above lemma implies that on any compact set one can keep $d(x)$ in assumption (B1) constant and the resulting function $\gamma(x)$ is also bounded on this compact set. Hence, $M(x)$, defined in (3), is bounded on compact sets. $\qquad\triangleleft$

The estimate in assumption (L) is uniform over compact subsets of E. This is an excessive requirement for deducting (B1), but it will be needed for proving the continuity of the value function in the following section. Moreover, to the best of our knowledge most conditions for a weakly Feller process to satisfy (L) pointwise automatically imply that the bound is uniform on compact sets.

2.4. Continuity of the value function

So far it has only been proved that the value function is lower semi-continuous, which was sufficient for contructing an optimal stopping time. However, it is often important to know if the value function is continuous. We will require it in Section 4 to characterise the value function as a viscosity solution to an appropriate variational inequality.

Assumption (L)

Assumption (L) ensures the continuity of w without any additional conditions. Denote by $\mathcal{A}(x)$ the set of stopping times with the expectation bounded by $M(x)$, defined in (3). Then

$$0 \leq w(x) - w_T(x) \leq \sup_{\tau \in \mathcal{A}(x)} \mathbb{E}^x \Big\{ \int_T^{\tau \vee T} f(X_s)ds + 1_{\{\tau > T\}} \big(g(X_\tau) - g(X_T) \big) \Big\}$$

$$\leq \sup_{\tau \in \mathcal{A}(x)} \mathbb{E}^x \Big\{ \int_T^{\tau \vee T} f(X_s)ds \Big\} + 2\|g\| \frac{M(x)}{T}, \tag{10}$$

where w_T is a value function for the stopping problem with horizon T. Fix a compact set $L \subset E$. The second term converges to 0 uniformly on L since $M(x)$ is bounded on L (see Remark 2.25). Take any $\varepsilon > 0$. By Lemma 2.23 there is $S > 0$ such that for $t \geq S$ we have $\mathbb{P}^x\{A_t\} \leq Ce^{-\rho t}$, where

$$A_t = \Big\{ \sup_{s \geq t} \Big| \frac{1}{s} \int_0^s f(X_u)du - \mu(f) \Big| > \varepsilon \Big\}.$$

Since $\mu(f) < 0$, for any stopping time $\tau \in \mathcal{A}(x)$ we have

$$\mathbb{E}^x \Big\{ \int_T^{\tau \vee T} f(X_s)ds \Big\} = \mathbb{E}^x \Big\{ 1_{\{\tau \geq T\}} \int_T^\tau (f(x_s) - \mu(f))ds + 1_{\{\tau \geq T\}} \mu(f)(\tau - T) \Big\}$$

$$\leq \mathbb{E}^x \Big\{ 1_{\{\tau \geq T\}} \int_T^\tau (f(x_s) - \mu(f))ds \Big\}.$$

Then, for $T \geq S$:

$$\mathbb{E}^x \Big\{ 1_{\{\tau \geq T\}} \int_T^\tau (f(x_s) - \mu(f))ds \Big\} \leq \sum_{i=0}^\infty \mathbb{E}^x \Big\{ 1_{T+i \leq \tau < T+i+1} \int_T^\tau (f(x_s) - \mu(f))ds \Big\}$$

$$\leq \sum_{i=0}^\infty \mathbb{E}^x \big\{ 1_{T+i \leq \tau < T+i+1} 1_{\{A_{T+i}\}} (i+1)2\|f\| \big\} + \sum_{i=0}^\infty \mathbb{E}^x \Big\{ 1_{T+i \leq \tau < T+i+1} 1_{\{A_{T+i}^c\}} \int_T^\tau (f(x_s) - \mu(f))ds \Big\}$$

$$\leq 2\|f\|C \sum_{i=0}^\infty (i+1)e^{-\rho(T+i)} + \sum_{i=0}^\infty \mathbb{E}^x \Big\{ 1_{T+i \leq \tau < T+i+1} 1_{\{A_{T+i}^c\}} \Big[\varepsilon\tau + \Big| \int_0^T (f(x_s) - \mu(f))ds \Big| \Big] \Big\}.$$

The first term is bounded by $2\|f\|Ce^{-\rho T}(1 - e^{-\rho})^{-2}$. For the second term notice that $A_T^c \subset A_{T+i}^c$. Hence, it is bounded from above by

$$\sum_{i=0}^\infty \mathbb{E}^x \big\{ 1_{T+i \leq \tau < T+i+1} 1_{\{A_{T+i}^c\}} [\varepsilon\tau + \varepsilon T] \big\} \leq 2\varepsilon \mathbb{E}^x \{\tau\} \leq 2\varepsilon M(x).$$

In total, we have demonstrated that

$$0 \leq w(x) - w_T(x) \leq 2\|f\|Ce^{-\rho T} \frac{1}{(1 - e^{-\rho})^2} + 2\varepsilon M(x) + 2\|g\| \frac{M(x)}{T}$$

for sufficiently large T. This proves that when $T \to \infty$ the difference $w - w_T$ converges, uniformly on L, to a positive function that is bounded by $2\varepsilon M(x)$. Due to the arbitrariness of $\varepsilon > 0$, this proves that w_T converges to w uniformly on L. Recalling that the value functions w_T are continuous yields the continuity of w.

Negative f

Assume that $f \leq 0$. Recall that

$$w(x) = \sup_{\tau \in \mathcal{A}(x)} \mathbb{E}^x \Big\{ \int_0^\tau f(X_s)ds + g(X_\tau) \Big\},$$

where $\mathcal{A}(x)$ is a set of stopping times τ such that $\mathbb{E}^x\{\tau\} \leq M(x)$. Let $w_T(x)$ be the value function of the stopping problem with stopping times bounded by T. Then

$$0 \leq w(x) - w_T(x) \leq \sup_{\tau \in \mathcal{A}(x)} \mathbb{E}^x \Big\{ \int_T^{T \vee \tau} f(X_s)ds + g(X_\tau) - g(X_{\tau \wedge T}) \Big\}$$

$$\leq \sup_{\tau \in \mathcal{A}(x)} \mathbb{E}^x \{ 1_{\{\tau > T\}} 2\|g\| \}$$

$$\leq 2\|g\| \frac{M(x)}{T},$$

where the penultimate inequality is due to $f \leq 0$ and the last inequality follows from Tchebyshev's inequality. The function $M(x)$ is bounded on compacts when (a) assumption (L) holds (see Remark 2.25), or when (b) assumptions (C1), (C2) are satisfied: by Lemma 2.16 $\gamma(x) \leq q(x) - A$, where A is the lower bound for q, and q is continuous. This means that functions w_T converge to w uniformly on compact sets. As w_T are continuous, so is w.

Strong Feller property

Recall that a Markov process (X_t) satisfies the *strong Feller property* if its semigroup P_t, $t > 0$, maps measurable bounded functions into continuous functions.

Under assumptions (C1)-(C2), the optimal stopping problem (2) is equivalent to the stopping problem (6) which we recall for convenience:

$$\hat{w}(x) := w(x) - q(x) = \sup_\tau \limsup_{T \to \infty} \mathbb{E}^x \{ \mu(f)(\tau \wedge T) + (g - q)(X_{\tau \wedge T}) \}.$$

The function $\hat{w}$ is bounded from above because $\mu(f) < 0$ and q is bounded from below. Let $\delta > 0$. By the strong Feller property the mapping $x \mapsto \mathbb{E}^x \{ \hat{w}(X_\delta) \vee (-M) \}$ is continuous for every $M > 0$. By monotone convergence theorem, $\lim_{M \to \infty} \mathbb{E}^x \{ \hat{w}(X_\delta) \vee (-M) \} = \mathbb{E}^x \{ \hat{w}(X_\delta) \}$, so the function $P_\delta \hat{w}(x)$ is upper semicontinuous. We shall show now that $P_\delta \hat{w}$ converges to $\hat{w}$ uniformly on compact sets as $\delta \to 0$, which implies that $\hat{w}$ and $w = \hat{w} + q$ are upper semicontinuous. Combined with the lower semicontinuity of w and the continuity of q this yields that w is continuous.

For the proof that $P_\delta \hat{w}$ converges to $\hat{w}$ uniformly on compact sets we need the following properties of weakly Feller processes.

LEMMA 2.26. *([21, Proposition 2.1]) For any compact set $K \subset E$, $T > 0$ and $\varepsilon > 0$ there is N such that for all $n \geq N$*

$$\sup_{x \in K} \mathbb{P}^x \big\{ \forall t \in [0, T] \ X_t \in B(0, n) \big\} \geq 1 - \varepsilon,$$

where $B(y, n)$ is a closed ball of radius n centered at y.

LEMMA 2.27. *([8, Theorem 3.7]) For any compact set $K \subset E$ and any $\varepsilon, \eta > 0$ there is $h_0 > 0$ such that*

$$\sup_{0 \leq h \leq h_0} \sup_{x \in K} \mathbb{P}^x \{ X_h \notin B(x, \eta) \} < \varepsilon.$$

Notice that $\hat{w}(x) - P_\delta \hat{w}(x) = w(x) - P_\delta w(x) - q(x) + P_\delta q(x)$. Assumption (C2) gives

$$P_\delta q(x) - q(x) = \mathbb{E}^x\Big\{\int_0^\delta \big(f(X_s) - \mu(f)\big)ds\Big\}.$$

Hence, $|P_\delta q(x) - q(x)| \le 2\|f\|\delta$. By Lemma 2.8 we have $\mathbb{E}^x\{w(X_\delta) + \int_0^\delta f(X_s)ds\} \le w(x)$ which implies

$$P_\delta w(x) - w(x) \le \delta\|f\|.$$

On the other hand, Lemma 2.8 applied for $\sigma = \tau^* \wedge T$ (with $T \ge \delta$) and $\tau = \delta$ yields

$$
\begin{aligned}
w(x) &= \limsup_{T\to\infty} \mathbb{E}^x\Big\{\int_0^{\tau^*\wedge T} f(X_s)ds + g(X_{\tau^*\wedge T})\Big\}\\
&\le \mathbb{E}^x\Big\{\int_0^{\tau^*\wedge\delta} f(X_s)ds + 1_{\{\tau^*<\delta\}}\,g(X_{\tau^*}) + 1_{\{\tau^*\ge\delta\}}\,w(X_\delta)\Big\}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
w(x) - P_\delta w(x) &\le \delta\|f\| + \mathbb{E}^x\{1_{\{\tau^*<\delta\}}\big[g(X_{\tau^*}) - w(X_\delta)\big]\}\\
&\le \delta\|f\| + \mathbb{E}^x\{1_{\{\tau^*<\delta\}}\big[g(X_{\tau^*}) - g(X_\delta)\big]\}\\
&\le \delta\|f\| + \sup_{\tau\le\delta}\mathbb{E}^x\{g(X_\tau) - g(X_\delta)\}.
\end{aligned}
$$

Fix a compact set $K \subset E$. By Lemma 2.26, for any $\varepsilon > 0$ there is a compact set L such that

$$\sup_{x\in K} \mathbb{P}^x\{X_t \in L \,\forall\, t \in [0,1]\} \ge 1 - \varepsilon.$$

By Lemma 2.27, for any $\eta > 0$ there is $h_0 > 0$ such that for all $\delta < h_0$ we have

$$\sup_{0\le h\le h_0}\sup_{x\in L} \mathbb{P}^x\{X_h \notin B(x,\eta)\} < \varepsilon.$$

Take $\delta < h_0 \wedge 1$. Then

$$
\begin{aligned}
&\sup_{\tau\le\delta}\mathbb{E}^x\{g(X_\tau) - g(X_\delta)\}\\
&\le \sup_{\tau\le\delta}\mathbb{E}^x\big\{1_{\{X_\tau\in L\}}\mathbb{E}^{X_\tau}\{g(X_0) - g(X_{\delta-\tau})\} + 1_{\{X_\tau\notin L\}}\big(g(X_\tau) - g(X_\delta)\big)\big\}\\
&\le \big(\omega(\eta) + \varepsilon 2\|g\|\big) + \varepsilon 2\|g\|,
\end{aligned}
$$

where $\omega_L(\cdot)$ is the modulus of continuity of g restricted to the 1-neighbourhood of L, i.e., the set $\{x \in E : \|x - y\| \le 1 \text{ for some } y \in L\}$. Summarising, for any $\eta \in (0,1)$ and a sufficiently small $\delta > 0$

$$-\delta\|f\| \le w(x) - P_\delta w(x) \le \delta\|f\| + \omega_L(\eta) + 4\varepsilon\|g\|.$$

Hence, $P_\delta(x)$ converges to $w(x)$ uniformly on K. This completes the proof of uniform on compact sets convergence of $P_\delta\hat{w}$ to $\hat{w}$ when $\delta \to 0$.

Uniform integrability of $K(X_T)$

Assume (D1) and

(D2') for every compact set $L \subset E$ there is $\alpha > 0$ such that

$$\sup_{x \in L,\, T \geq 0} \mathbb{E}^x\{K(X_T)^{1+\alpha}\} < \infty,$$

where K is defined in assumption (D1).

This, in particular, implies (D2). In the proof of Lemma 2.18 we have established that $|q(x)| \leq \bar{q}(x) \leq CK(x)$, where $C = 2\|f\| \int_0^\infty h(t)dt$. Hence,

$$\sup_{x \in L,\, T \geq 0} \mathbb{E}^x\{|q(X_T)|^{1+\alpha}\} =: \Gamma < \infty. \tag{11}$$

This estimate will play a crucial role in showing that w_T converges to w uniformly on L. Recall (10):

$$w(x) - w_T(x) \leq \sup_{\tau \in \mathcal{A}(x)} \mathbb{E}^x\left\{ \int_T^{\tau \vee T} f(X_s)ds \right\} + 2\|g\|\frac{M(x)}{T}.$$

Since $M(x)$ is bounded on L the second term converges uniformly to 0. Take a stopping time τ with $\mathbb{E}^x\{\tau\} \leq M(x)$:

$$\mathbb{E}^x\left\{ \int_T^{\tau \vee T} f(X_s)ds \right\} = \mathbb{E}^x\{\mu(f)(\tau \vee T - T) + q(X_T) - q(X_{\tau \vee T})\}$$

$$\leq \mathbb{E}^x\{1_{\{\tau > T\}}\, q(X_T)\} - A\,\mathbb{P}^x\{\tau > T\},$$

where A is a lower bound for q which can always be taken negative. By Hölder inequality, for any $x \in L$

$$\mathbb{E}^x\{1_{\{\tau > T\}}\, q(X_T)\} \leq \left(\mathbb{P}^x\{\tau > T\}\right)^{1/q}\left(\mathbb{E}^x\{|q(X_T)|^{1+\alpha}\}\right)^{1/(1+\alpha)} \leq \left(\frac{M(x)}{T}\right)^{1/q}\Gamma^{1/(1+\alpha)},$$

where $q = (1+\alpha)/\alpha$ is adjoint to $1+\alpha$. Summarising,

$$w(x) - w_T(x) \leq \left(\frac{M(x)}{T}\right)^{1/q}\Gamma^{1/(1+\alpha)} + (2\|g\| - A)\frac{M(x)}{T},$$

which proves uniform convergence to 0 on L when $T \to \infty$.

Relaxation of the lower bound on q

Assumption (C1) requires q to be bounded from below. This can be interpreted as a guarantee that long waiting is not penalised severely. Lemma 2.19 establishes the lower boundedness of q under the assumption that $f > \mu(f)$ outside of a compact set. But, intuitively, harsh penalisation of waiting (i.e., lower f) should make the problem easier as it would provide incentives to stop earlier and, therefore, prevent the explosion in the functional. This is indeed the case as we shall show in the theorem below.

THEOREM 2.28. *Assume (D1)-(D2) and one of the following conditions*

- *the process (X_t) is strongly Feller, or*

- *assumption (D2').*

Then the value function w is continuous.

Proof. Assume first that there is a continuous bounded function $\bar{f}$ such that $\bar{f} \geq f$, $\mu(\bar{f}) < 0$ and the set $\{x \in E : \bar{f}(x) \leq \mu(\bar{f})\}$ is compact. We will relax it later. Let $z_n(x) = 1 - \rho(x, B_n) \wedge 1$, where $\rho(x, B_n)$ is the distance of x from the ball of radius n. Put

$$f_n(x) = \bar{f}(x) - z_n(x)\big(\bar{f}(x) - f(x)\big).$$

Clearly, $f_n \equiv \bar{f}$ on B_{n+1}^c and $f_n \leq \bar{f}$. Hence, $\mu(f_n) \leq \mu(\bar{f})$,

$$\{x \in E : f_n(x) \leq \mu(f_n)\} \subset B_{n+1} \cup \{x \in E : \bar{f}(x) \leq \mu(\bar{f})\},$$

and assumption (B1) holds for f_n by Lemma 2.18 and 2.19. Denote by w_n the value function corresponding to f_n, i.e.,

$$w_n(x) = \sup_{\tau} \limsup_{T \to \infty} \mathbb{E}^x \Big\{ \int_0^{\tau \wedge T} f_n(X_s)ds + g(X_{\tau \wedge T}) \Big\}.$$

By Theorem 2.11 and discussion in the previous subsections w_n is continuous and there is an optimal stopping time $\tau_n^* = \inf\{t \geq 0 : g(X_t) \geq w_n(X_t)\}$. Since f_n are decreasing in n then w_n are decreasing in n and so do τ_n^*. Hence, for any $T > 0$

$$\mathbb{E}^x \Big\{ \int_0^{\tau_n^*} f_n(X_s)ds + g(X_{\tau_n^*}) \Big\} - \mathbb{E}^x \Big\{ \int_0^{\tau_n^* \wedge T} f_n(X_s)ds + g(X_{\tau_n^* \wedge T}) \Big\}$$

$$\leq \|f_n\| \mathbb{E}^x\{(\tau_n^* - T)^+\} + 2\|g\| \mathbb{P}^x\{\tau_n^* > T\}$$

$$\leq (\|f\| + \|\bar{f}\|)\mathbb{E}^x\{(\tau_1^* - T)^+\} + 2\|g\|\mathbb{P}^x\{\tau_1^* > T\} =: H(T).$$

This implies that $w_n(x) - w_{n,T}(x) \leq H(T)$, where $w_{n,T}$ stands for the value function when stopping times are bounded by T. Notice that due to the integrability of τ_1^*, we have $\lim_{T \to \infty} H(T) = 0$. Hence,

$$w_n(x) - w(x) = \big(w_n(x) - w_{n,T}(x)\big) + \big(w_{n,T}(x) - w_T(x)\big) + \big(w_T(x) - w(x)\big)$$
$$\leq H(T) + \big(w_{n,T}(x) - w_T(x)\big) + 0.$$

Pointwise convergence of w_n to w is established once we show that $w_{n,T}(x) - w_T(x)$ can be made arbitrarily small for every fixed T. By Lemma 2.26 for each $\varepsilon > 0$ and a compact set $L \subset E$ there is N such that $\mathbb{P}^x\{\forall t \in [0, T]\ X_t \in B_n\} \geq 1 - \varepsilon$ for $n \geq N$ and $x \in L$, i.e., the process stays in the ball B_n over time $[0, T]$ with the probability at least $1 - \varepsilon$. Using the fact that f_n coincides with f on B_n, we have for every $n > N$

$$0 \leq w_{n,T}(x) - w_T(x) \leq T\|f_n - f\|\, \mathbb{P}^x\{\exists t \in [0, T]\ X_t \notin B_n\} \leq T(\|f\| + \|\bar{f}\|)\varepsilon.$$

Summarising, for every $T > 0$, $\varepsilon > 0$ and a compact set L there is N such that for all $n \geq N$ and $x \in L$ we have $w_n(x) - w(x) \leq H(T) + T(\|f\| + \|\bar{f}\|)\varepsilon$. This proves convergence of w_n to w uniform in L. Hence, the continuity of w_n implies the continuity of w.

Take now an arbitrary continuous bounded function f with $\mu(f) < 0$. Let N be such that $\int_{B_N^c} |f(x)|\mu(dx) < -\mu(f)/4$. Define $\hat{f}(x) = z_N(x)f(x)$ and $\bar{f} = f \vee \hat{f}$. Then

$$\mu(\bar{f}) \leq \int_{B_N} f(x)\mu(dx) + \int_{B_N^c} |f(x)|\mu(dx) \leq \mu(f) - \mu(f)/4 - \mu(f)/4 = \mu(f)/2.$$

Hence, $\mu(\bar{f}) < 0$. Moreover, $\bar{f}(x) \geq 0$ for $x \in B_{N+1}^c$, so the set $\{x \in E : \bar{f}(x) \leq \mu(\bar{f})\}$ is contained in B_{N+1} and compact. The function $\bar{f}$ satisfies the conditions in the first part of the proof. $\quad\square$

REMARK 2.29. The function $\bar{f}$ in the above proof satisfies conditions of Lemma 2.19. Hence, assumption (B1) holds for $\bar{f}$. As it has been said earlier (see Remark 2.3), it also holds for any function dominated by $\bar{f}$, in particular, for f. Theorem 2.11 yields that $\tau^* = \inf\{t \geq 0 : g(X_t) \geq w(X_t)\}$ is optimal and has the expectation bounded by $M(x)$. Function $M(x)$ is bounded on compact sets because it is so for $\bar{f}$. ◁

3. Functional with a general discount rate

In the previous section we studied an optimal stopping problem without discounting. Here, we solve the problem with an arbitrary non-negative discount rate $r(x)$. The development will follow a similar line of thought as before but due to the presence of discounting many steps are more complicated.

For the sake of compactness of notation, we shall denote

$$\alpha_t = \int_0^t r(X_s)ds.$$

The optimal stopping problem (1) takes then the form

$$w(x) = \sup_\tau \limsup_{T\to\infty} \mathbb{E}^x\left\{\int_0^{\tau\wedge T} e^{-\alpha_s} f(X_s)ds + e^{-\alpha_{\tau\wedge T}} g(X_{\tau\wedge T})\right\}. \tag{12}$$

Standing assumptions for this section are (A1) (weak Feller property), (A3) (the continuity of f, g) and

(E1) For any $x \in E$, there is $d(x) < 0$ such that

$$\gamma_r(x) = \sup_\tau \limsup_{T\to\infty} \mathbb{E}^x\left\{\int_0^{\tau\wedge T} e^{-\alpha_s}\big(f(X_s) - d(x)\big)ds\right\} < \infty,$$

(E2) Function r is continuous bounded and $r(x) \geq 0$ for all $x \in E$.

Similarly as in the undiscounted case, when assumption (E1) holds for a function f it also holds for any function $f' \leq f$ with the same function d.

LEMMA 3.1. *For any $x \in E$ and any integrable ε-optimal stopping time σ we have*

$$\mathbb{E}^x\left\{\int_0^\sigma e^{-\alpha_s}ds\right\} \leq \frac{\gamma_r(x) + 2\|g\| + \varepsilon}{-d(x)}.$$

The optimisation in (12) can be constrained to stopping times σ with

$$\mathbb{E}^x\left\{\int_0^\sigma e^{-\alpha_s}ds\right\} \leq M(x) := \frac{\gamma_r(x) + 2\|g\| + 1}{-d(x)}$$

and $w(x) \leq \|f\|M(x) + \|g\| < \infty$.

The proof is similar to that of Lemma 2.4.

Define a value function for the stopping problem with a finite horizon:

$$w_T(x) = \sup_{\tau \leq T} \mathbb{E}^x\left\{\int_0^\tau e^{-\alpha_s} f(X_s)ds + e^{-\alpha_\tau} g(X_\tau)\right\}.$$

21

LEMMA 3.2. *Functions w_T are continuous and the value function w is lower semicontinuous.*

Proof. The discounted semigroup $P_t^r \phi(x) = \mathbb{E}^x\{e^{-\int_0^t r(X_u)du}\phi(X_t)\}$ maps continuous bounded functions into continuous bounded functions, [9, Chapter II, Section 5, Lemma 4]. This implies the continuity of w_T. As w can be approximated pointwise from below by continuous functions w_T it is itself lower semicontinuous. $\qquad\square$

Hence, the random variable (using the convention that $\inf\emptyset = \infty$)

$$\tau_\varepsilon = \inf\{t \geq 0 : g(X_t) \geq w(X_t) - \varepsilon\}$$

is a stopping time. On the set $\{\tau_\varepsilon < \infty\}$, due to the continuity of g and lower semicontinuity of w as well as the right-continuity of (X_t) we have

$$g(X_{\tau_\varepsilon}) \geq w(X_{\tau_\varepsilon}) - \varepsilon.$$

It is not possible to determine without further assumptions whether τ_ε admits only finite values. The following example explains why.

EXAMPLE. Consider an optimal stopping problem $w(x) = \sup_\tau \limsup_{T\to\infty} \mathbb{E}^x\{e^{-(\tau\wedge T)r}g(X_{\tau\wedge T})\}$ for $r > 0$ and $g \equiv -1$. It is easy to see that there is a unique optimal stopping time $\tau^* = \infty$ and the value function $w(x) = 0$. However, for any $\varepsilon \in (0, 1)$ the set $\{t : g(X_t) \geq w(X_t) - \varepsilon\}$ is empty, so $\tau_\varepsilon = \infty$. $\qquad\triangleleft$

This example suggests that although it cannot be guaranteed that τ_ε is finite-valued it can still be ε-optimal for w. This property is shared by the class of optimal stopping problems discussed in this section. The following technical lemmas provide us with tools required to demonstrate it. Proofs that are very similar to their counterparts in Section 2 are omitted.

LEMMA 3.3. *For every $x \in E$, there exists a non-decreasing sequence σ_m of bounded $\frac{1}{m}$-optimal stopping times for $w(x)$.*

In what follows we shall refer by (σ_m) to the sequence of stopping times from the above lemma.

LEMMA 3.4. *Assume only (E2). Then:*

- *The process $Z_t := \int_0^t e^{-\alpha s} f(X_s)ds + e^{-\alpha t}w(X_t)$ is a right-continuous $\mathbb{P}^x$-supermartingale for any $x \in E$.*

- *For a bounded stopping time σ and an arbitrary stopping time τ*

$$\mathbb{E}^x\left\{\int_0^\sigma e^{-\alpha s} f(X_s)ds + e^{-\alpha\sigma}g(X_\sigma)\right\}$$
$$\leq \mathbb{E}^x\left\{\int_0^{\sigma\wedge\tau} e^{-\alpha s}f(X_s)ds + 1_{\{\sigma<\tau\}}e^{-\alpha\sigma}g(X_\sigma) + 1_{\{\sigma\geq\tau\}}e^{-\alpha\tau}w(X_\tau)\right\}. \quad (13)$$

- *For a bounded stopping time σ*

$$\mathbb{E}^x\left\{\int_0^\sigma e^{-\alpha s}f(X_s)ds + e^{-\alpha\sigma}w(X_\sigma)\right\} \leq w(x). \quad (14)$$

Proof. First two statements are proved in a similar way as Lemma 2.8. Inequality (14) follows from (13) by taking $\tau \equiv 0$. $\qquad\square$

LEMMA 3.5.

1. $\mathbb{E}^x\{1_{\{\sigma_m < \tau_\varepsilon\}} e^{-\alpha_{\sigma_m}}\} \le \dfrac{1}{m\varepsilon}$.
2. $\lim\limits_{m \to \infty} \mathbb{E}^x\{1_{\{\tau_\varepsilon < \infty\}} 1_{\{\sigma_m \ge \tau_\varepsilon\}} e^{-\alpha_{\tau_\varepsilon}}\} = \mathbb{E}^x\{1_{\{\tau_\varepsilon < \infty\}} e^{-\alpha_{\tau_\varepsilon}}\}$.

Proof. We follow similar lines as in the proof of Lemma 2.9. Apply (13) to σ_m and τ_ε

$$
w(x) - \frac{1}{m} \le \mathbb{E}^x\left\{\int_0^{\sigma_m \wedge \tau_\varepsilon} e^{-\alpha_s} f(X_s)ds + 1_{\{\sigma_m < \tau_\varepsilon\}} e^{-\alpha_{\sigma_m}} g(X_{\sigma_m}) + 1_{\{\sigma_m \ge \tau_\varepsilon\}} e^{-\alpha_{\tau_\varepsilon}} w(X_{\tau_\varepsilon})\right\}
$$

$$
\le \mathbb{E}^x\left\{\int_0^{\sigma_m \wedge \tau_\varepsilon} e^{-\alpha_s} f(X_s)ds + 1_{\{\sigma_m < \tau_\varepsilon\}} e^{-\alpha_{\sigma_m}} \left(w(X_{\sigma_m}) - \varepsilon\right) + 1_{\{\sigma_m \ge \tau_\varepsilon\}} e^{-\alpha_{\tau_\varepsilon}} w(X_{\tau_\varepsilon})\right\}
$$

$$
\le \mathbb{E}^x\left\{\int_0^{\sigma_m \wedge \tau_\varepsilon} e^{-\alpha_s} f(X_s)ds + e^{-\alpha_{\sigma_m \wedge \tau_\varepsilon}} w(X_{\sigma_m \wedge \tau_\varepsilon})\right\} - \varepsilon\mathbb{E}^x\{1_{\{\sigma_m < \tau_\varepsilon\}} e^{-\alpha_{\sigma_m}}\}
$$

$$
\le w(x) - \varepsilon\mathbb{E}^x\{1_{\{\sigma_m < \tau_\varepsilon\}} e^{-\alpha_{\sigma_m}}\},
$$

where the second inequality follows from the observation that $g(X_{\sigma_m}) < w(X_{\sigma_m}) - \varepsilon$ on $\{\sigma_m < \tau_\varepsilon\}$ and the last inequality is the result of (14). This proves the first statement of the lemma.

The second statement is a consequence of the first one. Write

$$
\mathbb{E}^x\{1_{\{\tau_\varepsilon < \infty\}} e^{-\alpha_{\tau_\varepsilon}}\} = \mathbb{E}^x\{1_{\{\tau_\varepsilon < \infty\}} 1_{\{\sigma_m < \tau_\varepsilon\}} e^{-\alpha_{\tau_\varepsilon}}\} + \mathbb{E}^x\{1_{\{\tau_\varepsilon < \infty\}} 1_{\{\sigma_m \ge \tau_\varepsilon\}} e^{-\alpha_{\tau_\varepsilon}}\}.
$$

Recall that α_s is non-decreasing:

$$
\mathbb{E}^x\{1_{\{\tau_\varepsilon < \infty\}} 1_{\{\sigma_m < \tau_\varepsilon\}} e^{-\alpha_{\tau_\varepsilon}}\} \le \mathbb{E}^x\{1_{\{\tau_\varepsilon < \infty\}} 1_{\{\sigma_m < \tau_\varepsilon\}} e^{-\alpha_{\sigma_m}}\} \to 0
$$

as $m \to \infty$ by the first assertion of the lemma. $\square$

The following lemma unveils an important relation between τ_ε and σ_m.

LEMMA 3.6. *Under the standing assumptions of this section*

$$
\lim_{m \to \infty} (\sigma_m \wedge \tau_\varepsilon) = \tau_\varepsilon.
$$

Proof. Stopping times σ_m are non-decreasing, hence the limit $\sigma_\infty = \lim_{m \to \infty} \sigma_m$ exists (and is possibly infinite). From Lemma 3.5, by the dominated convergence theorem

$$
0 = \mathbb{E}^x\{\lim_{m \to \infty} 1_{\{\sigma_m < \tau_\varepsilon\}} e^{-\alpha_{\sigma_m}}\} \ge \mathbb{E}^x\{\lim_{m \to \infty} 1_{\{\sigma_m < \tau_\varepsilon\}} e^{-\|r\|\sigma_m}\}.
$$

Hence, on $\{\sigma_\infty < \infty\}$ we have $\lim_{m \to \infty} 1_{\{\sigma_m < \tau_\varepsilon\}} = 0$, which implies $\sigma_\infty \ge \tau_\varepsilon$. Obviously on $\{\sigma_\infty = \infty\}$ we have $\sigma_\infty \ge \tau_\varepsilon$. $\square$

We are now in a position to prove the main result of this section that τ_ε is an ε-optimal stopping time for w. This proof does not require τ_ε to be finite.

THEOREM 3.7. *Under the standing assumptions of this section the stopping time τ_ε is ε-optimal for* w.

Proof. Apply (13) to σ_m and τ_ε

$$w(x) - \frac{1}{m} \leq \mathbb{E}^x\left\{\int_0^{\sigma_m} e^{-\alpha_s} f(X_s)ds + e^{-\alpha_{\sigma_m}} g(X_{\sigma_m})\right\}$$

$$\leq \mathbb{E}^x\left\{\int_0^{\sigma_m \wedge \tau_\varepsilon} e^{-\alpha_s} f(X_s)ds + 1_{\{\sigma_m < \tau_\varepsilon\}} e^{-\alpha_{\sigma_m}} g(X_{\sigma_m}) + 1_{\{\sigma_m \geq \tau_\varepsilon\}} e^{-\alpha_{\tau_\varepsilon}} w(X_{\tau_\varepsilon})\right\}.$$

(15)

We now want to find out how the right-hand side of this inequality looks like when $m \to \infty$. By assertion 1 of Lemma 3.5,

$$\lim_{m \to \infty} \mathbb{E}^x\left\{1_{\{\sigma_m < \tau_\varepsilon\}} e^{-\alpha_{\sigma_m}} g(X_{\sigma_m})\right\} = 0.$$

On $\{\sigma_m \geq \tau_\varepsilon\}$ we have $\tau_\varepsilon < \infty$, so $w(X_{\tau_\varepsilon}) \leq g(X_{\tau_\varepsilon}) + \varepsilon$. Hence, for each m

$$\mathbb{E}^x\left\{1_{\{\sigma_m \geq \tau_\varepsilon\}} e^{-\alpha_{\tau_\varepsilon}} w(X_{\tau_\varepsilon})\right\} \leq \mathbb{E}^x\left\{1_{\{\sigma_m \geq \tau_\varepsilon\}} e^{-\alpha_{\tau_\varepsilon}} g(X_{\tau_\varepsilon})\right\} + \varepsilon$$

$$= \mathbb{E}^x\left\{1_{\{\tau_\varepsilon < \infty\}} 1_{\{\sigma_m \geq \tau_\varepsilon\}} e^{-\alpha_{\tau_\varepsilon}} g(X_{\tau_\varepsilon})\right\} + \mathbb{E}^x\left\{1_{\{\tau_\varepsilon = \infty\}} 1_{\{\sigma_m \geq \tau_\varepsilon\}} e^{-\alpha_{\tau_\varepsilon}} g(X_{\tau_\varepsilon})\right\} + \varepsilon.$$

Clearly, the second expectation is 0, while the first one converges to

$$\mathbb{E}^x\left\{1_{\{\tau_\varepsilon < \infty\}} e^{-\alpha_{\tau_\varepsilon}} g(X_{\tau_\varepsilon})\right\}$$

by the second assertion of Lemma 3.5. A proof of convergence of the first term on the right-hand side of (15) is slightly more involved. Notice that

$$\left|\int_0^{\sigma_m \wedge \tau_\varepsilon} e^{-\alpha_s} f(X_s)ds\right| \leq \|f\| \int_0^{\sigma_m \wedge \tau_\varepsilon} e^{-\alpha_s} ds.$$

The right-hand side is increasing in m. Hence, it is bounded by a random variable $\|f\| \int_0^{\sigma_\infty \wedge \tau_\varepsilon} e^{-\alpha_s} ds = \|f\| \int_0^{\tau_\varepsilon} e^{-\alpha_s} ds$ whose expectation does not exceed $\|f\| M(x)$. This allows us to use the dominated convergence theorem from which we obtain

$$\lim_{m \to \infty} \mathbb{E}^x\left\{\int_0^{\sigma_m \wedge \tau_\varepsilon} e^{-\alpha_s} f(X_s)ds\right\} = \mathbb{E}^x\left\{\int_0^{\tau_\varepsilon} e^{-\alpha_s} f(X_s)ds\right\}.$$

Since $\int_0^{\tau_\varepsilon} e^{-\alpha_s} ds \leq M(x)$ we also have

$$\mathbb{E}^x\left\{\int_0^{\tau_\varepsilon} e^{-\alpha_s} f(X_s)ds\right\} = \limsup_{T \to \infty} \mathbb{E}^x\left\{\int_0^{\tau_\varepsilon \wedge T} e^{-\alpha_s} f(X_s)ds\right\}.$$

Combining the above results gives

$$w(x) \leq \lim_{T \to \infty} \mathbb{E}^x\left\{\int_0^{\tau_\varepsilon \wedge T} e^{-\alpha_s} f(X_s)ds\right\} + \mathbb{E}^x\left\{1_{\{\tau_\varepsilon < \infty\}} e^{-\alpha_{\tau_\varepsilon}} g(X_{\tau_\varepsilon})\right\} + \varepsilon.$$

It remains to prove that

$$\limsup_{T \to \infty} \mathbb{E}^x\left\{1_{\{\tau_\varepsilon = \infty\}} e^{-\alpha_T} g(X_T)\right\} = 0.$$

(16)

By the first assertion of Lemma 3.5, we have

$$\mathbb{E}^x\left\{1_{\{\tau_\varepsilon = \infty\}} e^{-\sigma_m}\right\} \leq \frac{1}{m\varepsilon}.$$

24

By monotonicity of $t \mapsto \alpha_t$ and the dominated convergence theorem, this implies

$$\frac{1}{m\varepsilon} \geq \lim_{T\to\infty} \mathbb{E}^x \big\{ 1_{\{\tau_\varepsilon = \infty\}} \, e^{-\alpha_{\sigma_m \wedge T}} \big\} \geq \lim_{T\to\infty} \mathbb{E}^x \big\{ 1_{\{\tau_\varepsilon = \infty\}} \, e^{-\alpha_T} \big\}.$$

This is true for an arbitrary m, so the limit on the right-hand side is in fact 0. Boundedness of g then yields (16), which implies

$$w(x) \leq \limsup_{T\to\infty} \mathbb{E}^x \Big\{ \int_0^{\tau_\varepsilon \wedge T} e^{-\alpha_s} f(X_s) ds + e^{-\alpha_{\tau_\varepsilon \wedge T}} g(X_{\tau_\varepsilon \wedge T}) \Big\} + \varepsilon.$$

Hence τ_ε is ε-optimal. $\qquad\square$

We have shown that an ε-optimal stopping time exists in a standard form. Without further assumptions we are not able to construct an optimal stopping time. This should not be surprising as this is common for infinite horizon stopping problems even with a constant discount rate. Indeed, our main tool in the undiscounted case is the integral part of the functional which is interpreted as a penalty for a long wait before stopping – an infinite-valued stopping time drives the functional to $-\infty$. When discounting is applied, the integral part of the functional may remain bounded whatever the stopping time is applied; this is clearly the case when $\inf_x r(x) > 0$. However, when the discounting is limited, we are able to say more about an optimal stopping time and about properties of the value function.

(D3) The random variable $R := \lim_{t\to\infty} \alpha_t$ is $\mathbb{P}^x$-a.s. finite.

The random variable R is well-defined because $r(x) \geq 0$ so α_t is non-decreasing. It is satisfied if the discount rate $r(X_t)$ decreases quickly with time. In particular, a constant discount rate does not satisfy (D3).

(D3') There is $\delta > 0$ such that $\alpha_t \leq -\log(\delta)$, $\mathbb{P}^x$-a.s. for all $t \geq 0$.

The simplest example of a stopping problem satisfying (D3') is $r \equiv 0$, i.e., the problem studied in Section 2. Notice that assumption (D3') is a particular case of assumption (D3) with $R \leq -\log(\delta)$.

Let
$$\tau^* = \inf\{t \geq 0 : g(X_t) \geq w(X_t)\}.$$

On the set $\{\tau^* < \infty\}$, due to the continuity of g and lower semicontinuity of w as well as the right-continuity of (X_t) we have

$$g(X_{\tau^*}) \geq w(X_{\tau^*}).$$

We shall use this property in the proof of the following theorem.

THEOREM 3.8. *Assume (D3) in addition to the standing assumptions of this section. The stopping time τ^* is finite $\mathbb{P}^x$-a.s. and optimal for $w(x)$. Moreover, $\mathbb{E}^x \{\tau^* e^{-R}\} \leq M(x)$, where $M(x)$ is defined in Lemma 3.1. Under (D3') we additionally have $\mathbb{E}^x \{\tau_\varepsilon\} \leq M(x)/\delta$ and $\mathbb{E}^x \{\tau^*\} \leq M(x)/\delta$.*

Proof. We have

$$\mathbb{E}^x \{\sigma_m e^{-R}\} \leq \mathbb{E}^x \Big\{ \int_0^{\sigma_m} e^{-\alpha_s} ds \Big\} \leq M(x),$$

where the first inequality follows from the monotonicity of α_s and assumption (D3) while the second inequality is from Lemma 3.1. By the first assertion of Lemma 3.5 we obtain

$$\mathbb{E}^x \{1_{\{\sigma_m < \tau_\varepsilon\}} e^{-R}\} \leq \frac{1}{m\varepsilon}.$$

The factor e^{-R} can be interpreted as an unnormalised Radon-Nikodym density. Using this density to change the probability measure we end up in the setting of Lemma 2.10. Hence, $\mathbb{E}^x\{\tau_\varepsilon e^{-R}\} \le M(x)$. In Theorem 3.7 we showed that τ_ε is ε-optimal. Put $\tau_0 = \lim_{\varepsilon \to 0} \tau_\varepsilon$. Since this is an increasing sequence of random variables with the weighted expectation bounded by $M(x)$ then τ_0 is well defined and $\mathbb{E}^x\{\tau_0 e^{-R}\} \le M(x)$. Since R is finite-valued, this implies that $\tau_0 < \infty$ $\mathbb{P}^x$-a.s. We also have that $g(X_{\tau_\varepsilon}) \ge w(X_{\tau_\varepsilon}) - \varepsilon$. By the quasi left-continuity of (X_t), taking the limit $\varepsilon \to 0$ we obtain

$$g(X_{\tau_0}) \ge w(X_{\tau_0}).$$

So $\tau_0 \ge \tau^*$. This means that τ^* is finite-valued and $\mathbb{E}^x\{\tau^* e^{-R}\} \le M(x)$. Clearly, $\tau_\varepsilon \le \tau^*$, so $\tau_0 = \tau^*$. The optimality of τ^* follows immediately by letting $\varepsilon \to 0$ in

$$w(x) - \varepsilon \le \mathbb{E}^x\Big\{ \int_0^{\tau_\varepsilon} e^{-\alpha_s} f(X_s) ds + e^{-\alpha_{\tau_\varepsilon}} g(X_{\tau_\varepsilon}) \Big\}.$$

Under assumption (D3') we have $e^{-R} \ge \delta$. Then $\mathbb{E}^x\{\tau_\varepsilon \delta\} \le \mathbb{E}^x\{\tau_\varepsilon e^{-R}\} \le M(x)$ and, analogously, $\mathbb{E}^x\{\tau^* \delta\} \le M(x)$. Recalling that δ is a constant completes the proof. $\qquad \square$

3.1. Sufficient conditions for (E1)

We shall present two conditions that imply (E1). They generalise corresponding conditions from Section 2.

Define a potential

$$q_r(x) = \limsup_{T \to \infty} \mathbb{E}^x\Big\{ \int_0^T e^{-\alpha_s} \big(f(X_s) - \mu(f)\big) ds \Big\}.$$

and assume:

(C1') $q_r(x)$ is continuous and bounded from below,

(C2') for any bounded stopping time σ

$$q_r(x) = \mathbb{E}^x\Big\{ \int_0^\sigma e^{-\alpha_s} \big(f(X_s) - \mu(f)\big) ds + e^{-\alpha_\sigma} q_r(X_\sigma) \Big\}.$$

These assumptions collapse to their counterparts in Section 2 when $r \equiv 0$, i.e., $\alpha_t = 0$. Sufficient conditions for (C1') under uniform (geometric) ergodicity of (X_t) can be found in [29, Theorem 4] (see also [12]). An easy adaptation of Lemma 2.18 and 2.19 shows that conditions (C1')-(C2') are satisfied when $\mu(f) < 0$, the set $L = \{x \in E : f(x) \le \mu(f)\}$ is compact and assumptions (D1)-(D2) are fulfilled.

The statement and the proof of the following lemma resembles closely Lemma 2.16.

LEMMA 3.9. *Assume (C1')-(C2') and (C3). Assumption (E1) is satisfied with any $d(x) \in (\mu(f), 0)$. Moreover, $\gamma_r(x) \le q_r(x) - A$, where $A = \min\big(0, \inf_y q_r(y)\big)$.*

The second criterion for (E1) is provided by an upper bound for large deviations of the discounted empirical process (cf. assumption (L)).

LEMMA 3.10. *Assume*

(L$_r$) *For any $\delta, \varepsilon > 0$ and a compact set $K \subset E$ there is $N > 0$ and $p > 0$ such that for $n \geq N$ and $x \in K$ we have*

$$\mathbb{P}^x\left\{\left|\frac{1}{\int_0^{n\delta} e^{-\alpha_s} ds} \int_0^{n\delta} e^{-\alpha_s} f(X_s) ds - \mu(f)\right| > \varepsilon\right\} \leq e^{-p(n\delta)}.$$

Then (E1) holds and $M(x)$ defined in Lemma 3.1 is bounded on compact sets.

Proof. This proof is similar to the proof of Lemma 2.24. Choose $0 < \varepsilon < -\mu(f)$ and fix $x \in E$. Let $A_t = \left\{\sup_{u \geq t}\left|\left(\int_0^u e^{-\alpha_s} ds\right)^{-1} \int_0^u e^{-\alpha_s} f(X_s) ds - \mu(f)\right| > \varepsilon\right\}$. Similarly as in Lemma 2.23 there is $N > 0$ such that for $t \geq N$ we have $\mathbb{P}^x\{A_t\} \leq Ce^{-\rho t}$. Then for any integrable stopping time τ we have

$$\mathbb{E}^x\left\{\int_0^\tau e^{-\alpha_s}(f(x_s) - \mu(f)) ds\right\} = \sum_{i=0}^\infty \mathbb{E}^x\left\{1_{i \leq \tau < i+1} \int_0^\tau e^{-\alpha_s}(f(x_s) - \mu(f)) ds\right\}$$

$$\leq \varepsilon\mathbb{E}^x\left\{\int_0^\tau e^{-\alpha_s} ds\right\} + 2\|f\| + \sum_{i=1}^\infty 2\|f\|(i+1)\mathbb{P}^x\{A_i\}$$

$$\leq \varepsilon\mathbb{E}^x\left\{\int_0^\tau e^{-\alpha_s} ds\right\} + 2\|f\| + \sum_{i=1}^\infty 2\|f\|(i+1)Ce^{-i\rho}$$

$$= \varepsilon\mathbb{E}^x\left\{\int_0^\tau e^{-\alpha_s} ds\right\} + \bar{C}.$$

This is equivalent to

$$\mathbb{E}^x\left\{\int_0^\tau e^{-\alpha_s}\left(f(x_s) - \mu(f) - \varepsilon\right) ds\right\} \leq \bar{C}.$$

Hence, (E1) holds with $d(x) = \mu(f) + \varepsilon$ which, by the choice of ε, is strictly negative. Since $\bar{C}$ can be taken bounded on compact sets, we immediately obtain that d can be taken constant on compact sets and then $M(x)$ is bounded on compact sets. $\square$

3.2. Continuity of the value function

In three out of four cases studied in the theorem below, the continuity of w results from the convergence of w_T to w uniformly on compact sets and proofs from Subsection 2.4 apply. The proof when the process (X_t) is strongly Feller retains its validity due to the boundedness of the discount rate r (see assumption (E2)).

THEOREM 3.11. *Assume (E1)-(E2) and (A1). Either of the following conditions is sufficient for the continuity of the value function w:*

1. *Assumption (L$_r$) holds.*
2. *$f(x) \leq 0$ for $x \in E$.*
3. *The process (X_t) is strongly Feller and assumptions (C1')-(C2') hold.*
4. *Assumption (D1) and (D2') hold.*

Proof. The proof is an easy adaptation of the reasoning in Subsection 2.4. $\square$

4. Variational characterisation of the value function

Previous sections discuss properties of the value function and existence of optimal stopping times. This section provides a more explicit description of the value function as a solution to a variational inequality

$$\min\big(- \mathcal{A}w + rw - f, w - g\big) = 0, \tag{17}$$

where $\mathcal{A}$ is a generator for the process (X_t). In general, it is unlikely that the value function is in the domain $D_{\mathcal{A}}$ of the generator, so this variational formulation cannot be interpreted in a classical sense. Instead, we shall show that the value function is a viscosity solution of (17). We shall employ a more stringent definition than commonly used, i.e., we shall use a larger class of test functions.

DEFINITION 4.1. A continuous function u is a *viscosity subsolution* of (17) if for each $x \in E$ and $\psi \in D_{\mathcal{A}}$ such that $u(x) = \psi(x)$ and $\psi \geq u$ we have

$$\min\big(- \mathcal{A}\psi(x) + r(x)\psi(x) - f(x), u(x) - g(x)\big) \leq 0. \tag{18}$$

A continuous function u is a *viscosity supersolution* of (17) if for each $x \in E$ and $\phi \in D_{\mathcal{A}}$ such that $u(x) = \phi(x)$ and $\phi \leq u$ we have

$$\min\big(- \mathcal{A}\phi(x) + r(x)\phi(x) - f(x), u(x) - g(x)\big) \geq 0. \tag{19}$$

A continuous function u is a *viscosity solution* of (17) if it is both super- and subsolution.

In practice, the domain of a generator is rarely known. Instead, one considers a *core*, a linear subspace of $D_{\mathcal{A}}$ that defines the generator A on $D_{\mathcal{A}}$ uniquely via the closure of its graph. Test functions in the definition of the viscosity solutions are then restricted to belong to that core. For diffusions, it is common to consider C_0^2 functions or even C_0^∞ functions. Our choice to use all functions from the domain of the generator as test functions in the definition of viscosity solutions gives a flexibility for the selection of a core in the subsequent quest to solve (17) and to prove the uniqueness of solutions. This also comes at no additional cost.

Assume (A1)-(A3), (E1)-(E2) and (D3'). They ensure that there is an optimal stopping time with a finite expectation. We also require that the value function w is continuous. Theorem 3.11 provides sufficient conditions.

The domain of a generator of a weakly Feller process is a dense subset of the space of continuous functions vanishing in infinity [10, Theorem 17.4]. The range of the generator is within the same space – we shall use this continuity property in the proof of the following theorem.

THEOREM 4.2. *Under the standing assumptions of this section, the value function w is a viscosity solution of* (17).

Proof. Supersolution property: Take $x \in E$ and $\phi \in D_{\mathcal{A}}$ as in the definition of supersolution. Obviously, $\phi(x) - g(x) = w(x) - g(x) \geq 0$. By Lemma 3.4 for any $t > 0$ we have

$$w(x) \geq \mathbb{E}^x\Big\{ \int_0^t e^{-\alpha_s} f(X_s)ds + e^{-\alpha_t} w(X_t)\Big\}.$$

Thus

$$\mathbb{E}^x\big\{e^{-\alpha_t}\phi(X_t)\big\} - \phi(x) + \mathbb{E}^x\Big\{ \int_0^t e^{-\alpha_s} f(X_s)ds\Big\} \leq 0.$$

By Dynkin's formula

$$\mathbb{E}^x\Big\{\int_0^t e^{-\alpha_s}\Big(\mathcal{A}\phi(X_s) - r(X_s)\phi(X_s) + f(X_s)\Big)ds\Big\} \le 0.$$

Dividing by t both sides of the last inequality and letting $t \to 0$ we obtain $\mathcal{A}\phi(x) - r(x)\phi(x) + f(x) \le 0$.

Subsolution property: Take $x \in E$ and $\psi \in D_\mathcal{A}$ as in the definition of subsolution. If $w(x) = g(x)$ then (18) is trivially satisfied. Otherwise, $\tau^* > 0$ $\mathbb{P}^x$-a.s. By Lemma 3.4, the process $Z_t := \int_0^t e^{-\alpha_s} f(X_s)ds + e^{-\alpha_t}w(X_t)$ is a right-continuous $\mathbb{P}^x$-supermartingale. Since $\mathbb{E}^x\{Z_{\tau^*}\} = Z_0$ by the optimality of τ^*, the process $Z_{t\wedge\tau^*}$ is a martingale and therefore

$$\mathbb{E}^x\Big\{e^{-\alpha_{\tau^*\wedge t}}\psi(X_{\tau^*\wedge t}) - \psi(x) + \int_0^{\tau^*\wedge t} e^{-\alpha_s}f(X_s)ds\Big\} \ge 0.$$

Dynkin's formula yields for any $t > 0$

$$\mathbb{E}^x\Big\{\int_0^{\tau^*\wedge t} e^{-\alpha_s}\Big(\mathcal{A}\psi(X_s) - r(X_s)\psi(X_s) + f(X_s)\Big)ds\Big\} \ge 0.$$

Dividing both sides of the last inequality by t and letting $t \to 0$, taking into account that $\tau^* > 0$ $\mathbb{P}^x$-a.s., we obtain that $\mathcal{A}\psi(x) - r(x)\psi(x) + f(x) \ge 0$. $\qquad\square$

Demonstration of uniqueness of viscosity solutions to the variational inequality (17) is much more involved. We shall only conjecture here that when the process X_t is a weakly Feller jump-diffusion on $\mathbb{R}^d$ and viscosity solutions are defined by a core of C_0^2 test function then one can obtain uniqueness of solutions within a family of functions with a polynomial growth. The sufficiency of this notion depends very much on whether the value function has a polynomial growth. This growth is related to the expectation of the optimal stopping time τ^*. Under the setting of (D1)-(D2), this expectation is of the order $O(K(x))$, where $K(x)$ is from assumption (D1).

5. Dichotomy of discounting

In this section we drop the assumption $\mu(f) < 0$ and consider a general problem with a continuous bounded discount rate function r. We show that under certain assumptions (to be specified below) there are effectively two distinct regimes of the optimal stopping problem. When $\mu(r) > 0$, the stopping problem exhibits features of the classical optimal stopping problem with the discount rate separated from zero (cf. [24, 33]). In particular, the value function is continuous (hence, finite) for any continuous bounded function f and w_T approximates w uniformly on compact sets. On the other hand, when $\mu(r) = 0$ it behaves as if the discount rate was equal to 0. Indeed, the discount rate is non-negative so the assumption that $\mu(r) = 0$ implies that $r \equiv 0$ on the support of the invariant measure μ. It is common in many applications that the invariant measure is supported by the whole space E so $\int_0^t r(s)ds = 0$, $\mathbb{P}^x$-a.s. for all $t \ge 0$. In an unlikely case when the space contains a transient set (μ is not supported by the whole space) and the time until reaching the recurrent set is integrable, assumption (D3) is satisfied and we can employ Theorem 3.8.

Consider now $\mu(r) > 0$ and assume that the following upper bound for large deviations of the empirical process holds:

(L_d) For any $\delta, \varepsilon > 0$ and a compact set $K \subset E$ there is $N > 0$ and $p > 0$ such that for $n \geq N$ and $x \in K$ we have

$$\mathbb{P}^x \left\{ \left| \frac{1}{n\delta} \int_0^{n\delta} r(X_s)ds - \mu(r) \right| > \varepsilon \right\} \leq e^{-p(n\delta)}.$$

By an identical argument as in Lemma 2.23 there is a constant $C > 0$ and $\rho > 0$ such that for $t \geq S := N\delta$

$$\mathbb{P}^x \left\{ \sup_{s \geq t} \left| \frac{1}{s} \int_0^s r(X_u)du - \mu(r) \right| > \varepsilon \right\} \leq Ce^{-\rho t}, \qquad \forall\, x \in K.$$

Choose $0 < \varepsilon < \mu(r)$ and let

$$A_t = \left\{ \sup_{s \geq t} \left| \frac{1}{s} \int_0^s r(X_u)du - \mu(r) \right| > \varepsilon \right\}.$$

Consider a modified stopping problem

$$\tilde{w}(x) = \sup_\tau \mathbb{E}^x \left\{ \int_0^\tau e^{-(\alpha_s \vee \lambda s)} f(X_s)ds + e^{-(\alpha_\tau \vee \lambda \tau)} g(X_\tau) \right\},$$

where $\lambda = \mu(r) - \varepsilon$. One can solve this problem in a standard way (cf. [24]). Denoting by $\tilde{w}_T$ the value function with stopping times bounded by T we have

$$0 \leq \tilde{w}(x) - \tilde{w}_T(x) \leq \frac{1}{\lambda} e^{-\lambda T} \|f\| + 2e^{-\lambda T} \|g\|.$$

Functions $\tilde{w}_T$ are continuous [21, Corollary 3.6]. They approximate $\tilde{w}$ uniformly on compact sets, which implies that $\tilde{w}$ is continuous. ε-optimal stopping times take a standard form, while optimal stopping times might not exist as in the case with a constant discount rate.

It is interesting to note that under (L_d) the difference $w - \tilde{w}$ is bounded. Indeed, for $x \in K$

$$w(x) - \tilde{w}(x) \leq \sup_{\tau, \text{bounded}} \mathbb{E}^x \left\{ \int_0^\tau \left(e^{-\alpha_s} - e^{-(\alpha_s \vee \lambda s)} \right) f(X_s)ds + \left(e^{-\alpha_\tau} - e^{-(\alpha_\tau \vee \lambda \tau)} \right) g(X_\tau) \right\}.$$

Recall that $\mathbb{P}^x\{A_t\} \leq Ce^{-\rho t}$ for $t \geq S$. Let $n = \lfloor S \rfloor + 1$. Since on A_t^c there is $\alpha_s \geq \lambda s$ for $s \geq t$, we have

$$w(x) - \tilde{w}(x) \leq \sup_{\tau, \text{bounded}} \mathbb{E}^x \left\{ \sum_{i=n}^\infty \mathbb{1}_{\{i \leq \tau\}} \mathbb{1}_{\{A_i\}} \|f\| \right\} + n\|f\| + \|g\|$$

$$\leq \|f\| \sum_{i=n}^\infty \mathbb{P}^x\{A_i\} + n\|f\| + \|g\|$$

$$\leq \|f\| \left(n + \frac{C}{\rho} e^{-\rho n} \right) + \|g\|.$$

Similarly, one can estimate the difference $\tilde{w}(x) - w(x)$. This means that $w(x)$ is finite, so the integral term cannot explode. Moreover, this similarity between the two optimal stopping problem guides the following lemma.

LEMMA 5.1. *Assume (L_d) and $\mu(r) > 0$. The value function w is continuous and approximated uniformly on compact sets by w_T.*

Proof. Fix a compact set $K \subset E$ and $\varepsilon < \mu(r)$. Take $T \geq S$, where S is introduced below the large deviations assumption (L_d). For any $x \in E$

$$w(x) - w_T(x) \leq \sup_{\tau, \text{bounded}} \mathbb{E}^x \left\{ \int_{\tau \wedge T}^{\tau} e^{-\alpha s} f(X_s) ds + 1_{\{\tau > T\}} \left(e^{-\alpha_\tau} g(X_\tau) - e^{-\alpha_{\tau \wedge T}} g(X_{\tau \wedge T}) \right) \right\}$$

$$\leq \sup_{\tau, \text{bounded}} \mathbb{E}^x \left\{ \sum_{i=0}^{\infty} 1_{\{\tau \geq T+i\}} e^{-\alpha_{T+i}} \|f\| + 1_{\{\tau > T\}} e^{-\alpha_T} 2\|g\| \right\}$$

$$\leq \sup_{\tau, \text{bounded}} \mathbb{E}^x \left\{ \sum_{i=0}^{\infty} 1_{\{\tau \geq T+i\}} \left(1_{\{A_{T+i}\}} e^{-\alpha_{T+i}} + 1_{\{A_{T+i}^c\}} e^{-\lambda(T+i)} \right) \|f\| \right.$$

$$\left. + 1_{\{A_T\}} e^{-\alpha_T} 2\|g\| + 1_{\{A_T^c\}} e^{-\lambda T} 2\|g\| \right\}$$

$$\leq \|f\| \sum_{i=0}^{\infty} \mathbb{P}^x\{A_{T+i}\} + \|f\| e^{-\lambda T} \frac{1}{1 - e^{-\lambda}} + \mathbb{P}^x\{A_T\} 2\|g\| + e^{-\lambda T} 2\|g\|$$

$$\leq \|f\| \left(\frac{C}{p} e^{-\rho T} + e^{-\lambda T} \frac{1}{1 - e^{-\lambda}} \right) + 2\|g\| \left(C e^{-\rho T} + e^{-\lambda T} \right).$$

Hence, value functions w_T converge to w uniformly (and exponentially fast) on compact sets. Since w_T are continuous, so is w. $\qquad\square$

It is now standard to show that $\tau_\varepsilon = \inf\{t \geq 0 : g(X_t) + \varepsilon \geq w(X_t)\}$ is an ε-optimal stopping time for any $\varepsilon > 0$.

6. Examples

Example 1. Consider a diffusion on $\mathbb{R}^d$

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \tag{20}$$

where (W_t) is a standard d_1-dimensional Brownian motion with $d_1 \geq d$, $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^{d \times d_1}$ are bounded Borel measurable functions. The boundedness of b outside of a large enough ball is actually implied by the conditions spelled out below, so it is enough to assume local boundedness at this stage. Let $a(x) = \sigma(x)\sigma^T(x)$ and assume *uniform non-degeneracy*, i.e., $\inf_{\|\xi\|=1} \inf_x \xi^T a(x)\xi > 0$. The speed of convergence of X_t to its invariant measure will depend on the quantity

$$\eta(x) = \frac{x^T}{\|x\|} b(x).$$

The following list summarises three cases we will describe in detail. Each set of conditions implies assumptions (D1), (D2) and (D2'). It is easy to verify that the Ornstein-Uhlenbeck process belongs to the first class.

- If $\eta(x) \leq -r$ for sufficiently large x and some $r > 0$, then the process converges at an exponential speed.

- If $\eta(x) \leq -r/\|x\|^p$ for sufficiently large x and some $r > 0$, $0 < p < 1$, then the process converges at a subexponential speed.

- If $\eta(x) \leq -r/\|x\|$ for sufficiently large x and some $r > d/2 + 1$, then the process converges at a polynomial speed. In particular, it suffices to show that $\limsup_{\|x\| \to \infty} x^T b(x) = -\infty$.

Details are collected in two lemmas below (Lemma 6.1 and Lemma 6.2). We shall omit references for the first class as it is contained in the second one, which provides sufficient conditions for this paper. Although the same could be said about the relationship between the second and third class, we present details for both cases as they are less well known.

LEMMA 6.1 ([11]). *Assume that there is $M > 0$, $r > 0$ and $0 < p < 1$ such that $\eta(x) \leq -r/\|x\|^p$ for $\|x\| > M$. Then there are a unique invariant measure μ and constants C_0, C_1, A_0, A_1 such that*

$$\sup_{t \geq 0} \mathbb{E}^x \{e^{2A_0\|X_t\|^\alpha}\} \leq C_0 e^{2A_0\|x\|^\alpha}$$

and

$$\|P_t(x, \cdot) - \mu(\cdot)\|_{TV} \leq C_1 e^{A_0\|x\|^\alpha - A_1 t^\delta},$$

where $\alpha = 1 - p$ and $\delta = (1 - p)/(1 + p)$.

LEMMA 6.2 ([30]). *Assume that there is $M > 0$ and $r > d/2 + 1$ such that $\eta(x) \leq -r/\|x\|$ for $\|x\| > M$. Then there are a unique invariant measure μ and constants $m, k, C > 0$, such that*

$$\|P_t(x, \cdot) - \mu(\cdot)\|_{TV} \leq C(1 + \|x\|^m)(1 + t)^{-(1+k)}$$

and

$$\sup_{t \geq 0} \mathbb{E}^x \{\|X_t\|^{m+\delta}\} \leq C_0(1 + \|x\|^{m+\delta})$$

for sufficiently small $\delta > 0$.

The rate of convergence of transition probabilities to an invariant measure is closely related to so called mixing coefficients. In [31] it is shown that the bounds from Lemmas 6.1 and 6.2 are almost optimal.

Example 2. The reader is referred to [17] for details. Consider a jump-diffusion

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t + \int_{|z| \leq 1} \zeta(X_{t-}, z)\tilde{\mu}(dt, dz) + \int_{|z| > 1} \zeta(X_{t-}, z)\mu(dt, dz),$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathbb{R}^{d \times d_1}$ and $\zeta : \mathbb{R}^d \times \mathbb{R}^{d_2} \to \mathbb{R}^d$, (W_t) is a d_1-dimensional standard Brownian motion and μ is a Poisson measure on $[0, \infty) \times (\mathbb{R}^{d_2} \setminus \{0\})$ with the intensity $\nu(dz)dt$. The Levy measure ν satisfies $\int (|z| \wedge 1)\nu(dz) < \infty$. $\tilde{\mu}$ is the compensated measure μ, i.e., $\tilde{\mu}(dt, dz) = \mu(dt, dz) - \nu(dz)dt$. Let Assumptions 1, 2 and 3* from [17] be satisfied. The first one enforces linear growth for b, σ, ζ, the second one is concerned with absolute continuity of the the transition probability with respect to the Lebesgue measure. The last assumption is a Foster-Lyapunov criterion. Theorem 2.2 in the aforementioned paper shows the existence of a unique invariant measure. Coupled with Proposition 3.8 therein it proves our assumptions (D1)-(D2). As a special case consider an Ornstein-Uhlenbeck process

$$dX_t = -QX_t dt + dZ_t,$$

where (Z_t) is a (multi-dimensional) Levy process with the generating triplet (b, A, ν) and all eigenvalues of Q have positive real parts. If $\int_{|z|>1} |z|^q \nu(dz) < \infty$ for some $q > 0$ then the process (X_t) satisfies (D1)-(D2) [17, Theorem 2.6]; in fact, the convergence is exponentially fast. It can also be shown that it is strongly Feller if either (a) A is of full rank or (b) $\nu(\mathbb{R}^d) = \infty$ and ν is absolutely continuous with respect to the Lebesgue measure [16, Theorem 3.1]. Other conditions for the strong Feller property of jump diffusions can be found in [14, 32]. If the diffusion term vanishes, $\sigma \equiv 0$, i.e., when (X_t) is driven only by a jump noise, a sufficient condition for (D1) (with an exponential speed) and (D2) can be found in [13].

References

[1] D. Applebaum, *Lévy processes and stochastic calculus*, Cambridge studies in advanced mathematics, Cambridge University Press, 2004.

[2] A. Bensoussan and J.L. Lions, *Applications des inéquations variationnelles en contrôle stochastique*, Dunod, 1978.

[3] R. Carmona and N. Touzi, *Optimal multiple stopping and valuation of swing options*, Mathematical Finance **18** (2008), 239–268.

[4] M.D. Donsker and S.R.S. Varadhan, *Asymptotic evaluation of certain Markov processes for large time i*, Communications on Pure and Applied Mathematics **28** (1975), 1–47.

[5] ———, *Asymptotic evaluation of certain Markov processes for large time iii*, Communications on Pure and Applied Mathematics **29** (1976), 389–461.

[6] J.L. Doob, *Stochastic processes*, John Wiley & Sons, 1953.

[7] T. Duncan, B. Pasik-Duncan, and Ł. Stettner, *Adaptive control of discrete time Markov processes by large deviations method*, Applicationes Mathematicae **27** (2000), no. 3, 265–285.

[8] E.B. Dynkin, *Markov processes*, vol. 1, Springer-Verlag, 1965 (English).

[9] I.I. Gikhman and A.V. Skorokhod, *The theory of stochastic processes*, vol. 2, Springer-Verlag, Berlin, New York, 1975.

[10] O. Kallenberg, *Foundations of modern probability*, Springer, 1997.

[11] S.A. Klokov and A.Yu. Veretennikov, *On subexponential mixing rate for Markov processes*, Theory of Probability & Its Applications **49** (2005), no. 1, 110–122.

[12] I. Kontoyiannis and S.P. Meyn, *Spectral theory and limit theorems for geometrically ergodic Markov processes*, The Annals of Applied Probability **13** (2003), 304–362.

[13] A. M. Kulik, *Exponential ergodicity of the solutions to sdes with a jump noise*, Stochastic Processes and their Applications **119** (2009), 602–632.

[14] Y. Kwon and C. Lee, *Strong feller property and irreducibility of diffusions with jumps*, Stochastics and Stochochastic Reports **67** (1999), 147–157.

[15] R. Liptser, *Large deviations for occupation measures of Markov processes: Discrete time, noncompact case*, Theory of Probability & Its Applications **41** (1997), no. 1, 35–54.

[16] H. Masuda, *On multidimensional Ornstein-Uhlenbeck processes driven by a general Levy process*, Bernoulli **10** (2004), no. 1, 97–120.

[17] ———, *Ergodicity and exponential β-mixing bounds for multidimensional diffusions with jumps*, Stochastic Processes and their Applications **117** (2007), 35–56.

[18] P.A. Meyer, *Probability and potentials*, Blaisdell Publishing Company, 1966.

[19] S. Meyn and R.L. Tweedie, *Markov chains and stochastic stability*, 2nd ed., Cambridge University Press, New York, NY, USA, 2009.

[20] B. Øksendal and A. Sulem, *Applied stochastic control of jump diffusions*, Springer, 2005.

[21] J. Palczewski and Ł. Stettner, *Finite horizon optimal stopping of discontinuous functionals with applications to impulse control with delay*, SIAM Journal on Control and Optimization **48(8)** (2010), 4874–4909.

[22] G. Peskir and A.N. Shiryaev, *Optimal stopping and free-boundary problems*, Birkhäuser Verlag, 2006.

[23] H. Pham, *Continuous-time stochastic control and optimization with financial applications*, Springer, 2009.

[24] M. Robin, *Controle impulsionnel des processus de Markov*, Ph.D. thesis, University of Paris IX, 1978.

[25] _____, *On some impulse control problems with long run average cost*, SIAM Journal of Control and Optimization **19** (1981), 333–358.

[26] A.N. Shiryaev, *Optimal stopping rules*, Springer, 1978.

[27] Ł. Stettner, *On the Poisson equation and optimal stopping of ergodic Markov processes*, Stochastics **18** (1986), 25–48.

[28] _____, *On ergodic stopping and impulsive control problems for nonuniformly ergodic Markov processes*, Applied Mathematics and Optimization **19** (1989), 75–95.

[29] _____, *On some stopping and impulsive control problems with a general discount rate criteria*, Probability and Mathematical Statistics **10** (1989), no. 2, 223–245.

[30] A.Yu. Veretennikov, *On polynomial mixing bounds for stochastic differential equations*, Stochastic Processes and their Applications **70** (1997), 115–127.

[31] _____, *On lower bounds for mixing coefficients of markov diffusions*, From Stochastic Calculus to Mathematical Finance, Springer Berlin Heidelberg, 2006, pp. 623–633 (English).

[32] J. Wang, *Regularity of semigroups generated by levy type operators via coupling*, Stochastic Processes and their Applications **120** (2010), 1680–1700.

[33] J. Zabczyk, *Stopping problems in stochastic control*, Proceedings of the International Congress of Mathematicians, Warsaw, Poland, 1984, pp. 1425–1437.