10-4-2016

On Some Optimal Stopping Problems with Constraint

J. L. Menaldi
Wayne State University, menaldi@wayne.edu

M. Robin
Digiteo

SIAM J. CONTROL OPTIM.
Vol. 54, No. 5, pp. 2650–2671

ON SOME OPTIMAL STOPPING PROBLEMS WITH CONSTRAINT*

J. L. MENALDI† AND M. ROBIN‡

Abstract. We consider the optimal stopping problem of a Markov process $\{x_t : t \geq 0\}$ when the controller is allowed to stop only at the arrival times of a signal, that is, at a sequence of instants $\{\tau_n : n \geq 1\}$ independent of $\{x_t : t \geq 0\}$. We solve in detail this problem for general Markov–Feller processes with compact state space when the interarrival times of the signal are independent identically distributed random variables. In addition, we discuss several extensions to other signals and to other cases of state spaces. These results generalize the works of several authors where $\{x_t : t \geq 0\}$ was a diffusion process and where the signal arrives at the jump times of a Poisson process.

Key words. Markov–Feller processes, information constraints

AMS subject classifications. Primary, 49J40; Secondary, 60J60, 60J75

DOI. 10.1137/15M1040001

1. Introduction.

The usual (or standard) optimal stopping time problem refers to a simple stochastic control problem where the controller may, at any time, choose to stop the system with a terminal reward or to continue (and stop later).

This type of control problem has been studied extensively both with probabilistic methods and analytical methods using the link between optimal stopping problems and variational inequalities; see Bensoussan and Lions [2] and the bibliography therein.

In contrast, the constrained optimal stopping time problem imposes conditions on the decision of stopping or not, the state of the system.

In this paper, we consider the following constraint: the controller is allowed to stop the system only when a signal is received. For instance, the system evolves according to a Wiener process w_t (with drift b and diffusion σ) and the signal is the jumps of a Poisson process N_t, independent of the Wiener process. This particular model was studied by Depuis and Wang [6] for a geometric Brownian process. The dynamic programming equation (or HJB equation) takes the form

$$Au - \alpha u + \lambda[\psi - u]^+ = 0,$$

(where A is the infinitesimal generator of a Wiener process, λ is the intensity of an independent Poisson process, and ψ is the reward) and the verification theorem is based on an explicit solution of the HJB equation and on the solution of a discrete time problem.

Note that, for the standard stopping problem, the dynamic programming method leads to a variational inequality (see Bensoussan and Lions [2]).

The aim of the present paper is to generalize the above problem in two directions:

- to replace the one-dimensional Wiener process by a general Markov–Feller process in finite or infinite dimension,
- to replace the Poisson process by a more general counting process.

We treat in detail the case where the intervals between the arrival times of the signal are independent identically distributed random variables (IID case), indepen-

*Received by the editors September 15, 2015; accepted for publication (in revised form) August 17, 2016; published electronically October 4, 2016.
http://www.siam.org/journals/sicon/54-5/M104000.html
†Wayne State University, Department of Mathematics, Detroit, MI 48202 (menaldi@wayne.edu).
‡Digiteo, Fondation Campus Paris-Saclay, 91190 Saint-Aubin, France (maurice.robin@digiteo.fr).

dent from the process to be stopped, and we will discuss extensions to various other cases.

We obtain existence and uniqueness results for the HJB equation and the existence of an optimal stopping time. Obviously, in the case of "general" Markov processes, one cannot rely on an explicit solution of the HJB equation; the solution is obtained via a discrete time stopping problem associated to the continuous time one.

Let us mention that a variant of the problem with Poisson constraints was studied by Lempa [11] and that Liang [12] studied the same kind of problems for multidimensional diffusions and a Poisson signal, as well as a switching problem in Liang and Wei [13].

Let us mention that other classes of stopping problems with constraint have been studied. An example is given in Bensoussan and Lions [2], where it is allowed to stop only when the state x of the process lies in some given domain in the state space. Egloff and Leippold [7] studied an application of this kind of model to American options.

The paper is organized as follows: section 2 is devoted to the statement of the problem for the IID case and to the definition of the associated discrete time problem. In section 3, the discrete time problem is completely treated. In section 4, the solution of the continuous time problem is obtained. In section 5, the extension of the results of sections 3 and 4 to unbounded rewards and various types of signal processes is discussed.

2. Statement of the problem. In an abstract probability space $(\Omega, \mathcal{F}, P)$, let $\{x_t : t \geq 0\}$ be a continuous-time homogeneous Markov process with semigroup $\Phi(t)$, infinitesimal generator A, and transition probability function $p(x, t, B)$, i.e.,

$$P\{x_t \in B \,|\, x_s = x\} = p(x, t-s, B) \quad \forall x \in E,\ t > s \geq 0,\ B \in \mathcal{B}(E),$$

where E is a Polish space (complete, separable, and metrizable space) with its Borel σ-algebra $\mathcal{B}(E)$. If this process x_t represents the state of the system at time t, and then the typical optimal stopping time problem is to decide when to stop the system to maximize the reward functional

$$J_x(\theta) = \mathbb{E}_x\big\{\mathrm{e}^{-\alpha\theta}\psi(x_\theta)\big\} \quad \forall\theta \quad \text{stopping time}, \tag{2.1}$$

where ψ is a nonnegative Borel measurable function and $\mathbb{E}_x$ is the conditional expectation given $x_0 = x$. This amounts to describing the optimal reward function

$$\sup\big\{J_x(\theta) : \theta \text{ stopping time}\big\}$$

and to obtaining an optimal stopping time $\hat{\theta} = \hat{\theta}(x)$ for each initial state x. In this model, the constant $\alpha > 0$ is the discount factor and the terminal reward function ψ satisfies suitable conditions.

2.1. The optimal stopping with constraint. The usual optimal stopping problem as presented above is well known, but our interest here is to restrict the stopping action (of the controller) to certain instants when a signal arrives. A signal is given by a sequence $\{0 \leq \tau_1 < \tau_2 < \cdots\}$ of nonnegative random variables independent of $\{x_t : t \geq 0\}$.

The simplest model is when $\{\tau_n : n \geq 1\}$ are the jumps of a Poisson process, i.e., the time between two consecutive jumps $\{T_1 = \tau_1,\, T_2 = \tau_2 - \tau_1,\, T_3 = \tau_3 - \tau_2, \ldots\}$ is necessarily an IID sequence, exponentially distributed. In this paper, we focus on the situation where $\{T_n : n \geq 1\}$ is a sequence of IID random variables with common law π_0 (not exponential in general), and some possible extensions are discussed later.

Since the distribution is not longer exponential (i.e., not memoryless), it is useful to introduce an homogeneous Markov process $\{y_t : t \geq 0\}$ with states in $[0, \infty[$, independent of $\{x_t : t \geq 0\}$, which represents *the time elapsed since the last signal* (i.e., it is as if time is reset to zero when a signal is received and that time is measured), so that almost surely, the cad-lag paths $t \mapsto y_t$ are piecewise differentiable with derivative equals to one, with jumps only back to zero, and the form of infinitesimal generator is expected to be

$$A_1\varphi(y) = \partial_y\varphi(t) + \lambda(y)[\varphi(0) - \varphi(y)] \quad \forall y \geq 0, \tag{2.2}$$

where the intensity $\lambda(y) \geq 0$ is a Borel measurable function. Thus, the signals are defined as functionals on $\{y_t : t \geq 0\}$, namely,

$$\tau_n = \inf\big\{t > \tau_{n-1} : y_t = 0\big\}, \quad n \geq 1. \tag{2.3}$$

The couple $\{(x_t, y_t) : t \geq 0\}$ yields an homogeneous Markov process in continuous time, and a signal arrives at a random instant τ if and only if $\tau > 0$ and $y_\tau = 0$.

To define the family of *admissible* stopping times, first suppose that a cad-lag realization of the Markov process $\{(x_t, y_t) : t \geq 0\}$ is given, with its filtration $\mathbb{F} = \{\mathcal{F}_t : t \geq 0\}$ satisfying the usual conditions (of right-continuity and completeness). This construction also yields a realization of the Markov chain $\{(x_{\tau_n}, y_{\tau_n}, \tau_n) : n \geq 0\}$ with a filtration $\mathbb{G} = \{\mathcal{G}_n : n \geq 0\}$. With this setting, each signal τ_n, $n \geq 1$, is an $\mathbb{F}$-stopping time. Note that if η is a $\mathbb{G}$-stopping time with values in $\mathbb{N} = \{1, 2, \dots\}$, then the composition $\theta(\omega) = \tau_{\eta(\omega)}(\omega)$ is an $\mathbb{F}$-stopping time. All this leads to the following.

Definition 2.1. *An $\mathbb{F}$-stopping time θ is called "admissible" if for every almost surely ω there exists $n = \eta(\omega) \geq 1$ such that $\theta(\omega) = \tau_{\eta(\omega)}(\omega)$, and equivalently, if θ almost surely satisfies $\theta > 0$ and $y_\theta = 0$.* ◻

Now, without any loss of generality, we may assume that the final cost is nonnegative, $\psi \geq 0$, to describe our optimal control problem (or sequential decision problem) as

$$V(x, y) = \sup\big\{J_x(\theta) : \theta \text{ is an admissible stopping time}\big\} \tag{2.4}$$

with the sequence of signals $\{\tau_n = \tau_n^y : n \geq 1\}$ and Definition 2.1.

Moreover, if $\mathbb{1}_{\{y(\theta)=0\}} = 0$ when $y(\theta) \neq 0$ and $\mathbb{1}_{\{y(\theta)=0\}} = 1$ only when $y(\theta) = 0$, then the reward function

$$J_{x,y}(\theta) = \mathbb{E}_{x,y}\big\{\mathrm{e}^{-\alpha\theta}\psi(x_\theta)\mathbb{1}_{\{y(\theta)=0\}}\big\} \quad \forall\theta \quad \mathbb{F}\text{-stopping time} \tag{2.5}$$

can be used to give an equivalent formulation of our optimal control problem as

$$V(x, y) = \sup\big\{J_{x,y}(\theta) : \theta \text{ is a stopping time and } \theta > 0\big\} \tag{2.6}$$

for any initial value (x, y) in $E \times [0, \infty[$.

An auxiliary Markovian model very similar to (2.4) is given by the value function either

$$V_0(x, y) = \sup\big\{J_x(\theta) : \theta \text{ is a zero admissible stopping time}\big\},$$

where "zero admissible" stopping time means that $n = \eta(\omega) \geq 0$ instead of $\eta(\omega) \geq 1$ in Definition 2.1 with $\tau_0 = 0$, or equivalently,

$$V_0(x, y) = \sup\big\{J_{x,y}(\theta) : \theta \text{ is an } \mathbb{F}\text{-stopping time}\big\}, \tag{2.7}$$

where the constraint $\theta > 0$ has been removed. Again, the final reward ψ is assumed nonnegative. This is to say that $V_0(x, y)$ is the optimal reward of an "usual" optimal stopping time problem; the only different point is that its final reward function $(x, y) \mapsto \psi(x)\mathbb{1}_{\{y=0\}}$ has a discontinuity at $y = 0$.

2.2. Discrete time model. Based on the admissible stopping times as in Definition 2.1, it is natural to associate several discrete-time problems (DSTP1 and DSTP2) to the original problem. In a discrete-time model, a discrete stopping time η (having values $0, 1, \dots, \infty$) is the control, instead of an admissible stopping time θ, which has values in $[0, \infty]$ as in a continuous-time model. Note that capital letters are used for the various optimal rewards in continuous time V, V_0, while lower case letters like v_0, v are used in discrete time.

DSTP1 is defined as follows: The signals are given through an IID sequence $\{T_i : i \geq 0\}$ of nonnegative random variables, i.e., $\tau_1 = T_1, \dots, \tau_n = T_1 + T_2 + \cdots + T_n$. Since the signals are independent of $\{x_t : t \geq 0\}$, the expressions $\tau_0 = 0$ and $\{z_n = (x_{\tau_n}, \tau_n) : n \geq 0\}$ define an homogeneous Markov chain.

As mentioned early, a realization of this Markov chain yields a filtration $\mathbb{G} = \{\mathcal{G}_n : n \geq 0\}$ as generated by the random variables $z_0, z_1, \dots, z_n$. Therefore, for every $\mathbb{G}$-stopping time η, the reward functional is given by

$$K(z, \eta) = \mathbb{E}_z\big\{e^{-\alpha\tau_\eta}\psi(x_{\tau_\eta})\big\} \quad \forall z = (x, 0),\ x \in E, \tag{2.8}$$

with an optimal reward

$$v_0(x, 0) = \sup\big\{K(x, 0, \eta) : \eta \geq 0,\ \mathbb{G}\text{-stopping time}\big\} \quad \forall x. \tag{2.9}$$

Since

$$\mathbb{E}_0\big\{e^{-\alpha\tau_\eta}\big\} = \mathbb{E}_0\big\{\big(e^{-\alpha T_1}\big)^\eta\big\},$$

this description corresponds to a standard (or usual) discrete-time optimal stopping time problem. These type of Markovian models are well known, e.g., see the book by Shiryaev [19].

DSTP2 is defined as follows: It is a discrete-time optimal stopping time problem *with constraint*, namely, the controller is not allowed to stop the dynamic of the system at the initial time and must wait until the first signal arrives.

It uses the same homogeneous Markov chain $\{z_n : n \geq 0\}$ as above, but where the signals given through the homogeneous Markov process $\{y_t : t \geq 0\}$ represent the time elapsed since the last signal, i.e., τ_n is defined by (2.3) with initial value $y_0 = y \geq 0$. The reward is now

$$K(x, y, \eta) = \mathbb{E}_{x,y}\big\{e^{-\alpha\tau_\eta}\psi(x_{\tau_\eta})\big\} \quad \forall (x, y) \in E \times [0, +\infty[\tag{2.10}$$

and the optimal reward

$$v(x, y) = \sup\big\{K(x, y, \eta) : \eta \geq 1,\ \mathbb{G}\text{-stopping time}\big\} \quad \forall x, y. \tag{2.11}$$

Since no stopping can be applied before τ_1, it is intuitive that both problems (DSTP1 and DSTP2) are related by the equation

$$v(x, y) = \mathbb{E}_{x,y}\big\{e^{-\alpha\tau_1}v_0(x_{\tau_1}, 0)\big\}, \quad (x, y) \in E \times [0, +\infty[\tag{2.12}$$

with $\tau_1 = \tau_1^y$. If $y = 0$, then $\tau_1 = T_1$. This will be proved later.

2.3. Preliminary properties. There are several properties that are easily deduced from the formulation of the various problems.

From Definition 2.1 of admissible stopping times, it follows that an $\mathbb{F}$-stopping time θ is admissible if and only if there exists a discrete $\mathbb{G}$-stopping time $\eta \geq 1$ such that $\theta = \tau_\eta$. This simple assertion implies that taking the supremum of $J_{xy}(\theta)$ on θ admissible or taking the supremum of $K(x, y, \eta)$ on $\eta \geq 1$ yields the same values, i.e.,

$$V(x, y) = v(x, y) \quad \text{for any initial values} \quad (x, y) \in E \times [0, +\infty[. \tag{2.13}$$

Similarly,

$$V_0(x, 0) = v_0(x, 0) \quad \text{for any initial values} \quad x \in E. \tag{2.14}$$

Moreover, if $\hat{\eta}_0$ and $\hat{\eta}$ are optimal discrete stopping times for (DSTP1) and (DSTP2), respectively, then the continuous stopping times $\tau_{\hat{\eta}_0}$ and $\tau_{\hat{\eta}}$ are optimal admissible stopping times for the continuous time problems with optimal rewards $V_0(x, 0)$ and $V(x, y)$, given by (2.7) and (2.4). Conversely, if $\hat{\theta}_0$ and $\hat{\theta}$ are optimal admissible stopping times for the continuous time problems, then

$$\hat{\eta}_0 = \inf\left\{n \geq 0 : \tau_n = \hat{\theta}_0\right\} \quad \text{and} \quad \hat{\eta} = \inf\left\{n \geq 1 : \tau_n = \hat{\theta}\right\}$$

are optimal admissible stopping times for the discrete time problems with optimal rewards $v_0(x, 0)$ and $v(x, y)$, given by (2.9) and (2.11). Therefore, to solve our constrained continuous time problem with reward $V(x, y)$ given by (2.4), we can solve the discrete problem with reward $v(x, y)$ given by (2.11). This property is used later.

It is also clear from the definition of the auxiliary Markovian model with reward $V_0(x, y)$ given by (2.7) that

$$\text{if } y > 0, \text{ then} \quad V_0(x, y) = V(x, y) \quad \forall x \in E, \tag{2.15}$$

while

$$V_0(x, 0) \geq V(x, 0) \quad \forall x \in E. \tag{2.16}$$

Note that we have defined the discrete time optimal rewards $v_0(x, 0)$ and $v(x, y)$, but there is no definition for an optimal reward like $v_0(x, y)$ with $y > 0$ (which is not used), as in the case of continuous time like $V_0(x, y)$.

As discussed later, to actually prove the key relation (2.12), between the two discrete stopping time problems (DSTP1) and (DSTP2), the discrete dynamic programming will be involved.

2.4. Remarks on the model.

(1) The homogeneous Markov process $\{y_t : t \geq 0\}$ representing the *time elapsed since the last signal* can be easily constructed from a given sequence $\{\tau_n : n \geq 1\}$. Indeed, for initial conditions $y_0 = 0$ define $\tau_0 = 0$ and then by induction

$$y_t = t - \tau_{n-1} \quad \text{if} \quad \tau_{n-1} \leq t < \tau_n \quad \text{and} \quad y_{\tau_n} = 0, \quad n \geq 1. \tag{2.17}$$

However, if $y_0 = y > 0$, then conditional probability must be used to define y_t as beginning at time "$-y$" conditional to "having the first jump at sometime $t \geq 0$." This means that if the initial IID sequence $\{T_1, T_2, T_3, \dots\}$ of *waiting time between two consecutive signals* has its common law π_0 supported on $[0, \infty[$, and $y > 0$ is the initial condition at time $t = 0$, then first consider a nonnegative random variable τ^y

independent of $\{T_1, T_2, \ldots\}$ and of the Markov process $\{x_t : t \geq 0\}$ with distribution

$$P\{\tau^y \in]a, b]\} = P\{\tau_1 \in]a+y, b+y] \mid \tau_1 \geq y\} = \frac{\pi_0(]a+y, b+y])}{\pi_0(]y, +\infty[)} \tag{2.18}$$

for any $b > a \geq 0$. Now define the sequence of signals

$$\tau_0^y = 0, \quad \tau_1^y = \tau^y \quad \text{and} \quad \tau_{n+1}^y = \tau_n^y + T_n \quad \forall n \geq 1, \tag{2.19}$$

and the process $\{y_t : t \geq 0\}$ with $y_0 = y$ by the expressions

$$y_t = y_{\tau_{n-1}^y} + t - \tau_{n-1}^y \quad \text{if} \quad \tau_{n-1}^y \leq t < \tau_n^y \quad \text{and} \quad y_{\tau_n^y} = 0 \quad \forall n \geq 1, \tag{2.20}$$

$y_{\tau_n^y} = 0$, which agree with (2.17) when $y = 0$, while the process $\xi_t = \sum_{i=1}^{\infty} \mathbb{1}_{\tau_i \leq t}$ counts the jumps. Note that if the law π_0 is an exponential distribution, then τ^y has also the same exponential distribution π_0 (i.e., the jumps of y_t do not depend on the initial value y_0, in other words, τ^y can be regarded as one of T_i), and therefore, there is no need to introduce the Markov process $\{y_t : t \geq 0\}$ in the model. Moreover, if the law π_0 satisfies $\pi_0(]y_{\max}, \infty[) = 0$ (with $0 < y_{\max} < \infty$), then the initial value y should be taken either $0 < y < y_{\max}$ (if $\pi_0(\{y_{\max}\}) = 0$) or $0 < y \leq y_{\max}$ (if $\pi_0(\{y_{\max}\}) > 0$). ◻

(2) As mentioned early, we focus our attention on the case when $\{y_t : t \geq 0\}$ is an homogeneous Markov process, independent of $\{x_t : t \geq 0\}$, with values in $[0, \infty)$, and such that almost surely its cad-lag paths $t \mapsto y_t$ are piecewise linear with slope equal to one, except when it jumps back to zero. Thus the recurrence formula $\tau_0^y = 0$, and

$$\tau_n^y = \inf\big\{t > \tau_{n-1}^y : y_t = 0,\ y_0 = y\big\}, \quad n \geq 1, \tag{2.21}$$

defines the sequence $\{\tau_n^y : n \geq 1\}$ of signals, with a general initial condition $y \geq 0$, even if only $y = 0$ could be of main interest. Note that the strong Markov property implies that the law of $\{y_{t+\tau_n} : t \geq 0\}$ conditioned to $\{y_s : s \leq \tau_n\}$ depends only on $y_{\tau_n} = 0$, which proves that $T_n = \tau_{n+1} - \tau_n$ for $n \geq 1$ is an IID sequence. ◻

(3) First, recall three facts:

(a) if $\{\tau_1, \tau_2 - \tau_1, \ldots\}$ is a nonnegative IID sequence independent of a (homogeneous) Markov process $\{z_t : t \geq 0\}$, then the sequence $\{z_{\tau_n} : n \geq 1\}$ is a (homogeneous) Markov chain;

(b) if $\{z_t : t \geq 0\}$ is a (homogeneous) Markov process and Γ is a Borel subset, then the following sequence given by induction $\tau_0 = 0$ and $\tau_n = \inf\{t > \tau_{n-1} : z_t \in \Gamma\}$ can be used to define the (homogeneous) Markov chain $\{z_{\tau_n} : n \geq 0\}$, e.g., $z = (x, y)$ and $\Gamma = E \times \{0\}$;

(c) if $\{T_1, T_2, \ldots\}$ is a nonnegative IID sequence and g, h are Borel functions (on suitable Borel spaces), then for a given initial condition (ξ_0, τ_0), the expressions $\xi(\tau_0) = \xi_0$,

$$\begin{aligned} \tau_n &= \tau_{n-1} + T_n, \quad n \geq 1, \\ \xi(t) &= \xi_{n-1} + g(\xi(\tau_{n-1}))(t - \tau_{n-1}), \quad \tau_{n-1} \leq t < \tau_n, \quad n \geq 1, \\ \xi_n &= \xi(\tau_n) = h(\xi(\tau_n -)) = h\big(\xi_{n-1} + g(\xi(\tau_{n-1}))T_n\big) \end{aligned}$$

define an homogeneous Markov chain $\{\xi_n : n \geq 0\}$. However, the continuous time process $\{\xi(t) : t \geq 0\}$ is an homogeneous Markov process only if the distribution of the IID sequence $\{T_n\}$ is exponential. In any case, if the process $\{y(t) = y_t : t \geq 0\}$ as above (2.20) is added, then the couple $\{(\xi(t), y(t)) : t \geq 0\}$ is an homogeneous Markov process.

Second, note that in particular for (c), if $g = 1$ and $h(\xi) = \xi$, then $\xi_n = \xi_{n-1} + T_n$ for every $n \geq 1$, and therefore $\xi_n - \tau_n = \xi_0 - \tau_0$ for every $n \geq 1$, i.e., the sequences $\{\tau_n\}$ and $\{\xi_n\}$ are the same if the initial $x_0 = \tau_0$. Hence, if $\{\tau_n : n \geq 1\}$ represents the sequence of signals, then either $\{\xi_n : n \geq 0\}$ or $\{\tau_n : n \geq 0\}$ is an homogeneous Markov chain, but it is clear that the difference $t - \tau_{n-1}$ for $\tau_{n-1} \leq t < \tau_n$ represents the "time elapsed since the last signal" only for $n \geq 2$. On the other hand, if $\{T_1, T_2, \ldots\}$ is a sequence of independent nonnegative random variables, then (1) the construction in (c) also yields a Markov chain $\{\xi_n : n \geq 0\}$ and a Markov process $\{\xi(t) : t \geq 0\}$ (if the common distribution is exponential) and (2) the sequence in (a) defined by $\tau_0 = 0$ and $\tau_n = \tau_{n-1} + T_n$, $n \geq 1$, yields also a Markov chain (nonhomogeneous in general) $\{z_{\tau_n} : n \geq 0\}$, provided $\{T_1, T_2, \ldots\}$ is independent of the Markov process $\{z_t : t \geq 0\}$. In particular, $\{\tau_n : n \geq 0\}$ is a Markov chain (which is homogeneous if the $T_1, T_2, \ldots$ are identically distributed, i.e., if $\{T_n\}$ is an IID sequence), but note that the sequence $\{\tau_n : n \geq 0\}$ describes only the n-transition from 0 to τ_n of the Markov chain, and the n-transition from any value $\tau \geq 0$ is given by the conditional probability

$$P\{\tau_n \in]a + \tau, b + \tau] \,|\, \tau_n \geq \tau\} = \frac{\pi_n(]a + \tau, b + \tau])}{\pi_n(]\tau, +\infty[)} \quad \forall b > a \geq 0,$$

where π_n is the distribution of T_n, as in the construction of the Markov process $\{y_t : t \geq 0\}$. □

(4) In continuous time, the model with optimal reward $V_0(x, y)$ given by (2.7) is not "completely" Markovian even if the couple (x, y) defines an homogeneous Markov process $\{(x_t, y_t) : t \geq 0\}$. Indeed, "to wait for a signal" or "to wait for a jumps" cannot be restated in terms of the "state" (x, y), i.e., a constraint as "stopping is allowed only when a signal is received" cannot be translated into a condition either like "on any state of the form $(x, 0)$ the controller can stop the system if desired" (which implies that at the beginning stopping is allowed) or like "on any state of the form $(x, 0)$ the controller cannot stop the system" (which implies that the system can never be stopped). Essentially, the constraint "wait until ..." at the moment the dynamic starts is non-Markovian time-homogeneous (it uses the memory to know whether a signal has arrived early) for (x, y) as the state of the system. Another way of describing the situation is to say that the control problem is not time-homogeneous, i.e., to fully describe the system we need the term (x, y, t), so that stopping is allowed only when $y = 0$ and $t > 0$. □

(5) In contrast to Remark 2.4, the discrete time problem (DSTP1) is an usual discrete optimal stopping time problem. The underlying Markov chain is $\{z_n = (x_{\tau_n}, \tau_n) : n \geq 0\}$. In terms of the Markov chain $\{X_n = x_{\tau_n} : n \geq 0\}$ and the (variable) discount factor $\beta_n = \mathrm{e}^{-\alpha T_n}$, the reward functional can be written as

$$K(z, \eta) = \mathbb{E}_z \left\{ \psi(x_{\tau_\eta}) \prod_{n=0}^{\eta} \beta_n \right\} \quad \forall z = (x, 0),\ x \in E. \tag{2.22}$$

Alternatively, if the state $\{\tilde{X}_n : n \geq 0\}$ is the Markov chain with transition kernel

$$Pf(x) = \frac{1}{\beta} \int_0^\infty \mathrm{e}^{-\alpha t} \Phi(t) f(x) \pi_0(\mathrm{d}t), \qquad \beta = \int_0^\infty \mathrm{e}^{-\alpha t} \pi_0(\mathrm{d}t),$$

where π_0 is the common distribution of the IID sequence $\{T_n : n \geq 0\}$, then the same reward can be written as

$$K(z,\eta) = \mathbb{E}_z\big\{\beta^\eta \psi(\tilde{X}_\eta)\big\}, \quad \forall z = (x,0),\ x \in E,$$

which is a standard discounted discrete time optimal stopping problem. ◻

(6) If the Markov process $\{x_t\}$ is nonhomogeneous, then it suffices to add a new variable to have an homogeneous process, i.e., to consider the couple (x,t) as the initial state. In this case, the "time elapsed since the last signal" for a given initial condition $y(t_0) = y_0$ can also be defined as in Remark 2.4 and in terms of the sequence $\{T_n : n \geq 1\}$ of independent random variables (with distributions depending on n) representing the "waiting time between two consecutive signals," namely, $\tau_0 = t_0$, and

$$\begin{aligned}
&\tau_n = \tau_{n-1} + T_n \quad \text{and} \quad y_{\tau_n} = 0, \quad n \geq 1,\\
&y_t = y + t \quad \text{if} \quad \tau_0 \leq t < \tau_1,\\
&y_t = t - \tau_{n-1} \quad \text{if} \quad \tau_n \leq t < \tau_{n+1}, \quad n \geq 1,
\end{aligned}$$

and in this case, $\{(x_t,t) : t \geq 0\}$, $\{(y_t,t) : t \geq 0\}$ and $\{(x_t,y_t,t) : t \geq 0\}$ are all homogeneous Markov processes. Moreover, if the signals are nonhomogeneous (i.e., $\{T_n\}$ is a sequence of independent random variables, nonnecessarily with the same distribution) then additional variable is necessary as well. ◻

(7) Actually, if the sequence $\{T_1, T_2, \ldots\}$ is only conditional independent with respect to $\{x_t : t \geq 0\}$, then the (conditional) intensity may also be depending on the variable x, i.e., $\lambda(x,y)$,

$$\pi_0(x,]a,b]) = \int_a^b \exp\left(-\int_0^y \lambda(x,t)\mathrm{d}t\right)\lambda(x,y)\mathrm{d}y, \quad b > a \geq 0,$$

and the above construction (1) still works fine. In this case, the couple $\{(x_t,y_t) : t \geq 0\}$ or $\{x_t : t \geq 0\}$ is a homogeneous Markov process, but not necessary $\{y_t : t \geq 0\}$ alone is a Markov process. ◻

(8) As mentioned earlier, the signals can be given (a) either via the homogeneous Markov process $\{y_t : t \geq 0\}$ representing *the time elapsed since the last signal* (b) or via an IID sequence $\{T_n : n \geq 1\}$ of random variables representing the *waiting time between two consecutive signals*. The simplest model is the signals as the jumps of a Poisson process, where the common law π_0 of the IID sequence $\{T_n : n \geq 1\}$ is the exponential distribution and the infinitesimal generator of the homogeneous Markov process $\{y_t : t \geq 0\}$ is

$$A_1\varphi(y) = \partial_y\varphi(y) + \lambda[\varphi(0) - \varphi(y)] \quad \forall y \geq 0,$$

where ∂_y is the derivative and $\lambda > 0$ is the intensity. However, the whole Markov process $\{y_t : t \geq 0\}$ is of no use in this case. In general, the distribution π_0 must have support in either $[0,\infty[$ or $[0, y_{\max}]$ for some $y_{\max} < \infty$, and the form of infinitesimal generator is expected to be

$$A_1\varphi(y) = \partial_y\varphi(y) + \lambda(y)[\varphi(0) - \varphi(y)] \quad \forall y \geq 0, \tag{2.23}$$

where the intensity $\lambda(y) \geq 0$ is a Borel measurable function. If the common distribution π_0 has a density $\dot{\pi}_0$, i.e.,

$$\pi_0([0,y]) = \int_0^y \dot{\pi}_0(s)\mathrm{d}s \quad \forall y \geq 0,$$

then

$$\lambda(y) = \frac{\dot{\pi}_0(y)}{1 - \pi_0([0, y])} = \lim_{h \to 0} \frac{P\{y \le T_1 \le y + h \mid T_1 \ge y\}}{h} \quad \forall y \ge 0,$$

or equivalently

$$\lambda(y) = \left\{ -\ln \left[1 - \int_0^t \dot{\pi}_0(s)\mathrm{d}s \right] \right\}' \quad \forall y \ge 0,$$

which yields the conditional distribution

$$\pi(y,]a, b]) = \frac{\pi_0(]a, b] \cap]y, \infty[)}{\pi_0(]y, \infty[)} = \int_a^b \exp\left\{ -\int_0^t \lambda(y + s)\mathrm{d}s \right\} \lambda(t + y)\mathrm{d}t$$

for any real numbers $b > a \ge y \ge 0$. However, for a common distribution π_0 without a density (with respect to the Lebesgue measure) the form of the infinitesimal generator A_1 may not be known. In any case, it should be clear that almost surely the cad-lag paths $t \mapsto y_t$ are piecewise linear with slop equals to one, and with jumps only back to zero. ∎

3. Discrete time HJB equations. For the sake of simplicity and to fully understand the difficulties of this problem, it is convenient to impose throughout this section a restrictive assumption on the model. In section 4.2, some generalizations in various directions are discussed.

Assume that the homogeneous Markov process $\{x_t : t \ge 0\}$, with semigroup $\{\Phi(t) : t \ge 0\}$ and infinitesimal generator A, is a Feller process and E is a compact metric space, i.e.,

$$\Phi(t) \text{ is a continuous semigroup on } E, \text{ compact metric space.} \tag{3.1}$$

The discount factor α is positive and the reward function ψ is continuous, i.e.,

$$\alpha > 0 \quad \text{and} \quad \psi \in C(E). \tag{3.2}$$

The signals $\{\tau_n : n \ge 1\}$ besides being independent of $\{x_t : t \ge 0\}$ are given either though an IID sequence $\{T_n : n \ge 1\}$ of waiting time between two consecutive signals with a common distribution π_0 satisfying

$$\mathbb{E}\{\mathrm{e}^{-\alpha T_1}\} = \int_{[0,\infty[} \mathrm{e}^{-\alpha t} \pi_0(\mathrm{d}t) = k_0 < 1, \tag{3.3}$$

or equivalently, the homogeneous Markov process $\{y_t : t \ge 0\}$ of the time elapsed since the last signal satisfies

$$\mathbb{E}\{\mathrm{e}^{-\alpha\tau}\} = k_0 < 1 \quad \text{with} \quad \tau = \inf\{t > 0 : y_t = 0,\, y_0 = 0\}. \tag{3.4}$$

Sometimes, we will assume that the common distribution π_0 has a density.

These assumptions will be modified later to include a local compact metric space E or a Banach space and a reward with polynomial growth.

3.1. Discrete time standard optimal stopping. The dynamic programming principle in discrete time for (DSTP1) with an initial state $(x,0)$ is expressed as follows: the controller can decide either (a) to stop with a reward $\psi(x)$ or (b) to wait (i.e., to postpone the decision) with a discounted reward $\mathbb{E}_{x,0}\mathrm{e}^{-\alpha\tau_1}u_0(x_{\tau_1},0)$. This yields the following HJB equation:

$$u_0(x,0) = \max\big\{\psi(x), \mathbb{E}_{x,0}\mathrm{e}^{-\alpha T_1}u_0(x_{T_1},0)\big\} \quad \forall x \in E. \tag{3.5}$$

THEOREM 3.1. *Under the assumptions* (3.1), (3.2), *and* (3.4), *there exits one and only one solution u_0 in $C(E)$ of the HJB equation* (3.5).

Proof. In view of the assumptions, the operator

$$T_0 w(x) = \max\big\{\psi(x), \mathbb{E}_x\{\mathrm{e}^{-\alpha\tau_1}w(x_{\tau_1})\}\big\}$$

satisfies

$$|T_0w_1(x) - T_0w_2(x)| \le \mathbb{E}_x\{\mathrm{e}^{-\alpha\tau_1}\}\|w_1 - w_2\| \le k_0\|w_1 - w_2\| \quad \forall x \in E,$$

where $\|\cdot\|$ is the sup-norm in $C(E)$. Thus, the operator T_0 is contraction mapping on $C(E)$, which proves that the HJB equation (3.5) has a unique solution $u_0(x,0)$ in $C(E)$. ∎

THEOREM 3.2. *Under the assumptions* (3.1), (3.2), *and* (3.4), *the unique solution u_0 in $C(E)$ of the HJB equation* (3.5) *is indeed the optimal reward* (2.9), *i.e., $u_0(x,0) = v_0(x,0)$. Moreover, the first exit time from the continuation region is optimal, i.e., the discrete stopping time*

$$\hat{\eta}_0 = \inf\{n \ge 0 : u_0(x_{\tau_n},0) = \psi(x_{\tau_n})\} \tag{3.6}$$

satisfies $u_0(x,0) = K(x,0,\hat{\eta}_0)$ as given by (2.8).

Proof. The HJB equation (3.5) yields

$$\begin{aligned} u_0(x,0) &= \mathbb{E}_{x,0}\big\{\mathrm{e}^{-\alpha\tau_1}u_0(x_{\tau_1},0)\big\} \\ &\quad + \mathbb{E}_{x,0}\big\{\mathrm{e}^{-\alpha\tau_1}[\psi(x_{\tau_1}) - u_0(x_{\tau_1},0)]^+\big\} \quad \forall x \in E. \end{aligned}$$

or equivalently,

$$u_0(x,0) = \big[Q\big(u_0(\cdot,0) + [\psi - u_0(\cdot,0)]^+\big)\big](x),$$

using the sub-Makovian kernel $\varphi \mapsto Q\varphi$, with

$$\big(Q\varphi\big)(x) = \mathbb{E}_{x,0}\big\{\mathrm{e}^{-\alpha\tau_1}\varphi(x_{\tau_1})\big\}.$$

Thus, substituting x for $x_{\tau_{n-1}}$, this can be written as

$$u_0(x_{\tau_{n-1}},0) = \big[Q\big(u_0(\cdot,0) + [\psi - u_0(\cdot,0)]^+\big)\big](x_{\tau_{n-1}}),$$

or equivalently,

$$\begin{aligned} u_0(x_{\tau_{n-1}},0) &= \mathbb{E}_{x_{\tau_{n-1}},\tau_{n-1}}\big\{\mathrm{e}^{-\alpha T_n}u_0(x_{\tau_n},0)\big\} \\ &\quad + \mathbb{E}_{x_{\tau_{n-1}},\tau_{n-1}}\big\{\mathrm{e}^{-\alpha T_n}[\psi(x_{\tau_n}) - u_0(x_{\tau_n},0)]^+\big\} \quad \forall n \ge 1, \end{aligned}$$

where $\mathbb{E}_{x,\tau}$ means the conditional expectation with respect to the initial condition $x_\tau = x$. Hence, the strong Markov property yields

$$\begin{aligned} u_0(x_{\tau_{n-1}},0) &= \mathbb{E}\big\{\mathrm{e}^{-\alpha T_n}u_0(x_{\tau_n},0)\,|\,\mathcal{G}_{n-1}\big\} \\ &\quad + \mathbb{E}\big\{\mathrm{e}^{-\alpha T_n}[\psi(x_{\tau_n}) - u_0(x_{\tau_n},0)]^+\,|\,\mathcal{G}_{n-1}\big\} \quad \forall n \ge 1, \end{aligned}$$

and multiplying by $\mathrm{e}^{-\alpha\tau_{n-1}}$, this becomes

$$\mathrm{e}^{-\alpha\tau_{n-1}}u_0(x_{\tau_{n-1}},0) = \mathbb{E}\big\{\mathrm{e}^{-\alpha\tau_n}u_0(x_{\tau_n},0)\,|\,\mathcal{G}_{n-1}\big\} + \mathbb{E}\big\{\mathrm{e}^{-\alpha\tau_n}[\psi(x_{\tau_n}) - u_0(x_{\tau_n},0)]^+\,|\,\mathcal{G}_{n-1}\big\} \quad \forall n \geq 1.$$

Actually, in this calculation, the integer value n could be replaced by any $\mathbb{G}$-stopping time $\eta \leq n$.

Now, iterate this argument to deduce

$$u_0(x,0) = \mathbb{E}_{x,0}\big\{\mathrm{e}^{-\alpha\tau_n}u_0(x_{\tau_n},0)\big\} + \sum_{k=1}^{n}\mathbb{E}_{x,0}\big\{\mathrm{e}^{-\alpha\tau_k}[\psi(x_{\tau_k}) - u_0(x_{\tau_k},0)]^+\big\}. \tag{3.7}$$

Moreover, since all expectation are uniformly integrable (due to the fact that u_0 is bounded), the integer value n in the equality (3.7) could be replaced again by any $\mathbb{G}$-stopping time (nonnecessarily bounded, this time).

Hence, the equality (3.7) proves that

$$u_0(x,0) \geq \mathbb{E}_x\big\{\mathrm{e}^{-\alpha\tau_\eta}u_0(x_{\tau_\eta},0)\big\} \geq \mathbb{E}_x\big\{\mathrm{e}^{-\alpha\tau_\eta}\psi(x_{\tau_\eta})\big\} \quad \forall x \in E,$$

and for every $\mathbb{G}$-stopping time η, i.e., $u_0(x,0) \geq v_0(x,0)$, for every x in E. On the other hand, the definition of first exit time from the continuation region (3.6) and the equality (3.7) show the optimality of $\hat{\eta}_0$, i.e.,

$$u_0(x,0) = \mathbb{E}_x\big\{\mathrm{e}^{-\alpha\hat{\tau}_0}\psi(x_{\hat{\tau}_0})\big\}, \quad \hat{\tau}_0 = \tau_{\hat{\eta}_0},$$

and therefore $u_0(x,0) = v_0(x,0)$ for every x in E. ∎

Remark 3.1. The martingale theory can be used to prove Theorem 3.2. Indeed, define the processes

$$a_n = \mathrm{e}^{-\alpha\tau_n}u_0(x_{\tau_n},0), \quad n \geq 0,$$
$$b_n = a_{\min\{n,\hat{\eta}_0\}} = \mathrm{e}^{-\alpha\theta_n}u_0(x_{\theta_n},0), \quad \theta_n = \tau_{\min\{n,\hat{\eta}_0\}}, \quad n \geq 0,$$

to check that the HJB equation (3.5) and the Markov property imply that $\{a_n : n \geq 0\}$ is a bounded supermartingale and $\{b_n : n \geq 0\}$ is a bounded martingale. Hence, the sampling theorem and the equality $u_0 \geq \psi$ yield

$$K(x,0,\eta) = \mathbb{E}_x\big\{\mathrm{e}^{-\alpha\tau_\eta}\psi(x_{\tau_\eta})\big\} \leq \mathbb{E}_x\{a_\eta\} \leq \mathbb{E}_x\{a_0\} = u_0(x,0)$$

for every $\mathbb{G}$-stopping time η, i.e., $u_0 \geq v_0$. Also, with $\hat{\tau}_0 = \tau_{\hat{\eta}_0}$, the martingale part implies

$$u_0(x,0) = \lim_n \mathbb{E}_x\{b_n\} = \mathbb{E}_x\big\{\mathrm{e}^{-\alpha\hat{\tau}_0}u_0(x_{\hat{\tau}_0},0)\big\} = \mathbb{E}_x\big\{\mathrm{e}^{-\alpha\hat{\tau}_0}\psi(x_{\hat{\tau}_0})\big\},$$

i.e., $\hat{\eta}_0$ is an optimal $\mathbb{G}$-stopping time and $u_0 = v_0$. ∎

3.2. Discrete time constrained optimal stopping. The dynamic programming principle in discrete time for (DSTP2) with an initial state (x,y) at the initial time is expressed as follows: the controller (cannot decide to stop immediately) have to wait (i.e., to postpone the decision) and either to stop at time $t = \tau_1$ with a reward $\psi(x_{\tau_1})$ or to continue with a reward $u(x_{\tau_1}, 0)$. These two possibilities should be discounted, i.e., the following HJB equation

$$u(x, y) = \mathbb{E}_{x,y}\big\{\mathrm{e}^{-\alpha\tau_1} \max\{\psi(x_{\tau_1}), u(x_{\tau_1}, 0)\}\big\} \quad \forall (x, y) \in E \times [0, \infty[. \tag{3.8}$$

Now, the model involves the Markov process $\{y_t : t \geq 0\}$ representing the time elapsed since the last signal. Depending on the common distribution π_0 of IID sequence $\{T_n : n \geq 1\}$, the process $\{y_t : t \geq 0\}$ is a Feller process. Actually only the fact that

$$(x, y) \mapsto \mathbb{E}_{x,y}\big\{\mathrm{e}^{-\alpha\tau_1}\varphi(x_{\tau_1})\big\} \quad \text{is continuous} \tag{3.9}$$

for every continuous and bounded function φ is needed.

THEOREM 3.3. *Under the assumptions* (3.1)*,* (3.2)*, and* (3.4)*, there exits one and only one solution $x \mapsto u(x, 0)$ in $C(E)$ of the HJB equation* (3.8) *with $y = 0$. Moreover, the same HJB equation* (3.8) *determines the values of $u(x, y)$ when $u(x, 0)$ is known, and $y \mapsto u(x, y)$ continuous if a nonnegative Borel measurable and bounded intensity $y \mapsto \lambda(y)$ exists.*

Proof. This is analogous to Theorem 3.1. Consider the HJB equation (3.8) with $y = 0$ and the operator

$$Tw(x) = \mathbb{E}_{x,0}\big\{\mathrm{e}^{-\alpha\tau_1} \max\{\psi(x_{\tau_1}), w(x_{\tau_1})\}\big\}, \quad \text{note that} \ \ \tau_1 = T_1,$$

which is also contraction mapping on $C(E)$. This proves that the HJB equation (3.8) has a unique solution $u(x, 0)$ in $C(E)$. Next, the values of $u(x, y)$ for $y > 0$ are obtained from $u(x, 0)$ by (3.8).

Now, if an intensity $\lambda(y)$ exists and φ is a nonnegative Borel measurable and bounded function, then

$$\mathbb{E}_{x,y}\{\mathrm{e}^{-\alpha\tau_1}\varphi(x_{\tau_1})\} = \int_y^\infty \exp\left\{-\alpha t - \int_y^t \lambda(\xi)\mathrm{d}\xi\right\} \lambda(t)[\Phi(t - y)f(x)]\mathrm{d}t, \tag{3.10}$$

and the last integral is equal to

$$\int_0^\infty \exp\left\{-\alpha t - \int_0^t \lambda(y + \xi)\mathrm{d}\xi\right\} \lambda(t + y)[\Phi(t)f(x)]\mathrm{d}t,$$

where $\{\Phi(t) : t \geq 0\}$ is the semigroup of the initial Markov process. Hence, the expression (3.10) proves that the function (3.9) depends continously on y, even if the intensity $\lambda(y)$ is only a nonnegative Borel measurable bounded function. Therefore, the last assertion holds. ∎

REMARK 3.2. If an intensity λ (density $\dot{\pi}$) exists for the common law π of the sequence $\{T_n : n \geq 1\}$ of the waiting time between to consecutive signals (or equivalently, the infinitesimal generator A_1 of the Markov process $\{y_t : t \geq 0\}$ representing the time elapsed since the last signal has the form (2.2), and λ satisfies $\lambda(t) \leq \lambda_0 < \infty$, a.e. $t \geq 0$), then the operator T in the proof of Theorem 3.1 can be rewritten as

$$T(w) = \int_0^\infty \exp\left\{-\alpha t - \int_0^t \lambda(r)\mathrm{d}r\right\} \lambda(t)\big[\Phi(t)\big(\psi \vee w\big)\big]\mathrm{d}t.$$

Hence, the inequality

$$\int_0^\infty \exp\left\{-\alpha t - \int_0^t \lambda(r)\mathrm{d}r\right\} \lambda(t)\mathrm{d}t$$
$$= - \int_0^\infty \mathrm{e}^{-\alpha t}\mathrm{d}\left[\exp\left\{-\int_0^t \lambda(r)\mathrm{d}r\right\}\right] = 1$$
$$- \alpha \int_0^\infty \exp\left\{-\alpha t - \int_0^t \lambda(r)\mathrm{d}r\right\} \mathrm{d}t \le 1 - \frac{\alpha}{\alpha+\lambda_0} = \frac{\lambda_0}{\alpha+\lambda_0}$$

shows also that T is a contraction, i.e.,

$$\|T(w_1) - T(w_2)\| \le \frac{\lambda_0}{\alpha+\lambda_0}\|w_1 - w_2\| \quad \forall w_1, w_2 \in C(E),$$

where $\|\cdot\|$ is the sup-norm on $C(E)$. A similar argument can be made for the operator T_0. ∎

THEOREM 3.4. *If the assumptions* (3.1), (3.2), *and* (3.4), *then HJB equation* (3.8) *becomes*

$$u(x,y) = \mathbb{E}_{x,y}\big\{\mathrm{e}^{-\alpha\tau_1} u_0(x_{\tau_1}, 0)\big\} \quad \forall (x,y) \in E \times [0,\infty[. \tag{3.11}$$

Moreover, the solution of HJB equations (3.5) *and* (3.8) *satisfy the relation* $u_0(x,0) = \max\{\psi(x), u(x,0)\}$ *for every* x *in* E.

Proof. Define $w(x) = \max\{\psi(x), u(x,0)\}$ and use the HJB equation (3.8) with $y = 0$ to substitute $u(x,0)$ and deduce that

$$\begin{aligned} w(x) &= \max\{\psi(x), u(x,0)\} \\ &= \max\Big\{\psi(x), \mathbb{E}_{x,y}\big\{\mathrm{e}^{-\alpha\tau_1}\max\{\psi(x), u(x_{\tau_1},0)\}\big\}\Big\} \\ &= \max\big\{\psi(x), \mathbb{E}_{x,y}\{\mathrm{e}^{-\alpha T_1} w(x_{T_1})\}\big\}. \end{aligned}$$

This means that w is a solution of the HJB equation (3.5), and by uniqueness we get $w = u_0(x,0)$. ∎

THEOREM 3.5. *Under the assumptions* (3.1), (3.2), *and* (3.4), *the unique solution* u *in* $C(E)$ *of the HJB equation* (3.8) *is indeed the optimal reward* (2.11), *i.e.,* $u(x,y) = v(x,y)$. *Moreover, the exit time from the continuation region is optimal, i.e., the discrete stopping time*

$$\hat{\eta} = \inf\{n \ge 1 : u_0(x_{\tau_n}, 0) = \psi(x_{\tau_n})\} \tag{3.12}$$

satisfies $u(x,y) = K(x,y,\hat{\eta})$ *as given by* (2.10).

Proof. All that is needed is to use the strong Markov property and the HJB equation (3.8), in relation to Theorem 3.2. Indeed, once an optimal stopping time has been found for the discrete stopping time problem (DSTP1) with optimal reward $v_0(x,0) = u_0(x,0)$, the HJB equation (3.11) yields an optimal stopping time $\hat{\eta}$ as the τ_1-translation of $\hat{\eta}_0$, i.e., for an initial value (x,y) no stopping is allowed, we are forced to wait until the next possible stopping time given by τ_1, and at this new initial value $x_1 = x_{\tau_1}$ with $y = 0$, the optimal stopping time $\hat{\eta}_0$ is used. All this can be shown by following the argument in Theorem 3.2 and recalling that

$$u(x,y) = \mathbb{E}_{x,y}\big\{\mathrm{e}^{-\alpha\tau_1} u_0(x_{\tau_1}, 0)\big\} \quad \text{and}$$
$$u_0(x_{\tau_1}, 0) = \mathbb{E}_{x_{\tau_1},0}\big\{\mathrm{e}^{-\alpha\tau_{\hat{\eta}_0}}\psi(x_{\tau_{\hat{\eta}_0}})\big\},$$

i.e., the HJB equation (3.11) and the optimal character of $\hat{\eta}_0$. This proves that $u(x,y) = v(x,y)$, the optimal reward (2.11), and that there exits an optimal stopping time, i.e., for (DSTP2) as defined in section 2.2 . This optimal stopping time can be defined as the exit time of the continuation region, after waiting for the first signal, i.e.,

$$\hat{\eta} = \inf\{n \geq 1 : u_0(x_{\tau_n}, 0) = \psi(x_{\tau_n})\}, \quad \hat{\tau} = \tau_{\hat{\eta}};$$

note the condition $n \geq 1$ instead of $n \geq 0$ as in (DSTP1).

Now, from the HJB equation (3.11) and the equalities $u(x,y) = v(x,y)$ and $u_0(x,0) = v_0(x,0)$ we deduce the assertion (2.12). Note that we could replace $u_0(x,0)$ in (3.12) with $\max\{\psi(x), u(x,0)\}$, i.e.,

$$\hat{\eta} = \inf\{n \geq 1 : u(x_{\tau_n}, 0) \leq \psi(x_{\tau_n})\}$$

as expected. ∎

4. Continuous time HJB equation.

4.1. Formal analysis. Let us look at the dynamic programming principle in continuous time for our initial problem of optimal stopping time with constraint, where the optimal reward is given by (2.6). This argument is better understood under the assumption that the common distribution π_0 has a density so that

$$\lim_{h \to 0} \frac{P\big\{T_1 \in [y, y+h[\,|\, T_1 \geq h > 0\big\}}{h} = \lambda(y), \tag{4.1}$$

where $\lambda(y)$ is a nonnegative Borel measurable function.

For the initial value (x,y) in $E \times [0,\infty[$ at the initial time $t = 0$, the optimal reward $V(x,y)$ can be expressed on a small time interval $[0,h[$ as either (a) the future reward if the controller decides to do nothing on the $[0,h[$ or (b) the future reward if the controller decides to stop at τ_1 if it appears on the time interval $[0,h[$ and to continue if it does not appear, i.e.,

$$\begin{aligned} &V(x,y) = \max\{(a),(b)\} \quad \text{with} \\ &\quad (a) = \mathbb{E}_{x,y}\{\mathrm{e}^{-\alpha h} V(x_h, y_h)\}, \\ &\quad (b) = \lambda(y) h \mathbb{E}_{x,y}\{\mathrm{e}^{-\alpha h}\psi(x_h)\} + \big(1 - \lambda(y)h\big)\mathbb{E}_{x,y}\{\mathrm{e}^{-\alpha h} V(x_h, y+h)\}, \end{aligned}$$

since $y_h = y + h$ if there is not jump on $[0,h[$. By means of infinitesimal generator $A_{x,y}$ of the homogeneous Markov process $\{(x_t, y_t) : t \geq 0\}$ this can be written as

$$\begin{aligned} &(\mathrm{a}) = V(x,y) + \int_0^h \mathrm{e}^{-\alpha t}\mathbb{E}_{x,y}\big\{A_{x,y}V(x_t,y_t) - \alpha V(x_t,y_t)\big\}\mathrm{d}t, \\ &(\mathrm{b}) = \lambda(y)h\psi(x) \\ &\qquad + (1 - \lambda(y)h)\bigg[V(x,y) + \int_0^h \mathrm{e}^{-\alpha t}\mathbb{E}_{x,y}\bigg\{[A_x V(x_t, y+t)] \\ &\qquad\qquad + \frac{\partial V(x_t, y+t)}{\partial y} - \alpha V(x_t, y+t)\bigg\}\mathrm{d}t\bigg], \end{aligned}$$

where

$$A_{x,y}\varphi(x,y) = A_x\varphi(x,y) + \frac{\partial \varphi(x,y)}{\partial y} + \lambda(y)[\varphi(x,0) - \varphi(x,y)]$$

with A_x the infinitesimal generator of $\{x_t : t \geq 0\}$. Hence, as $h \to 0$, (a) yields

$$A_{x,y} V(x,y) - \alpha V(x,y)$$

and (b) yields

$$A_x V(x,y) + \frac{\partial V(x,y)}{\partial y} + \lambda(y)[V(x,0) - V(x,y)] - \alpha V(x,y) \\ + \lambda(y)[\psi(x) - V(x,0)].$$

Both together yield

$$A_{x,y} V(x,y) - \alpha V(x,y) + \lambda(y)\big[\psi(x) - V(x,0)\big]^+ = 0, \tag{4.2}$$

which can also be written as

$$A_x V(x,y) + \frac{\partial V(x,y)}{\partial y} - [\alpha + \lambda(y)] V(x,y) + \lambda(y) \max\big\{\psi(x), V(x,0)\big\}. \tag{4.3}$$

Now, let us go back to the auxiliary Markovian model given by the optimal reward $V_0(x,y)$ as in (2.7), where the constraint $\theta > 0$ have been removed. In this case, the controller is allowed to stop the evolution of the system only when the state (x,y) belongs to the region $\{(x,y) \in E \times [0,\infty[: y = 0\}$, in short, when $y = 0$.

Therefore, the dynamic programming principle applied to the auxiliary Markovian model (which is also an optimal stopping time with constraint) would be solved by the equations derived from the following: (a) in the region where stopping is not allowed, only run the dynamics until exits the region, and (b) in the region where stopping is allowed, either stop or continue, whatever is better. Thus, use the generators $A_{x,y} = A_x + A_y = A + A_1$ of a the continuous time Markov process $\{(x_t, y_t) : t \geq 0\}$ to obtain the HJB equation

$$(A + A_1 - \alpha) V_0(x,y) = 0 \quad \text{if} \quad y > 0 \tag{4.4}$$

with the boundary condition

$$\max\big\{(A + A_1 - \alpha) V_0(x,y), \psi(x) - V_0(x,y)\big\} = 0 \quad \text{if} \quad y = 0, \tag{4.5}$$

and if an intensity of jumps exists, then $A_1 = \partial_y$ on the boundary $y = 0$. Moreover, since $V_0 \geq \psi$ and ψ could be taken strictly positive (without any loss of generality), we can combine these conditions as

$$\max\big\{(A + A_1 - \alpha) V_0(x,y), \psi(x) \mathbb{1}_{\{y=0\}} - V_0(x,y)\big\} = 0 \tag{4.6}$$

for every x in E and $y \geq 0$.

Actually, since we do not expect V_0 to belong to the domain of the infinitesimal generator, an equivalent expression in terms of the semigroup of the couple (x,y) is desired, e.g., V_0 should satisfy the inequalities

$$U_0(x,y) \geq \mathbb{E}_{x,y}\big\{\mathrm{e}^{-\alpha t} U_0(x_t, y_t)\big\} \quad \forall t \geq 0, \\ U_0(x,y) \geq \psi(x) \mathbb{1}_{\{y=0\}} \quad \forall x, y,$$

with $U_0 = V_0$; actually, V_0 should be the minimum solution of these inequalities, i.e., $V_0 \leq U_0$ for any other U_0 as above. By means of the discrete time problem, a better alternative to this equation is

$$V_0(x,y) = \max\big\{\psi(x), \mathbb{E}_{x,y}\{\mathrm{e}^{-\alpha \tau_y} V_0(x_{\tau_y}, 0)\}\big\} \quad \forall x \in E,\ y \geq 0, \tag{4.7}$$

where $\tau_y = \inf\{t > 0 : y_t = 0,\ y_0 = y\}$ for every x in E and $y \geq 0$. Moreover, the expression

$$\begin{aligned} V_0(x,y) &= \mathbb{E}_{x,y}\{\mathrm{e}^{-\alpha\tau_y} V_0(x_{\tau_y}, 0)\} \\ &= \mathbb{E}_x \left\{ \int_0^\infty \mathrm{e}^{-\alpha t} \exp\left(- \int_0^t \lambda(y+s)\mathrm{d}s \right) \lambda(y+t) V_0(x_t, 0)\mathrm{d}t \right\} \end{aligned}$$

yields the value $V_0(x,y)$ after $V_0(x,0)$ has been found.

Comparing with the optimal reward $V(x,y)$ as given by (2.4), the same HJB equation (4.4) and (4.5) should be satisfied by V, except that at the initial time (say, $t = 0$), (4.4) should hold true, even if $y = 0$. This means that when the evolution begins, the controller should wait for the signal to arrive, even if the "state" is $(x.y)$ with $y = 0$. In this model, the full state is (x, y, t), i.e., it is not time-homogeneous.

4.2. Solving HJB equation. Based on the previous arguments on the discrete time problems, we have the following.

THEOREM 4.1. *Under the assumptions* (3.1), (3.2), *and* (3.4), *the HJB equation* (4.3) *has one and only one weak solution (i.e., in term of the resolvent)* V *in* $C_b(E\times]0,\infty[)$, *which agrees with the optimal reward* (2.4). *Moreover, the exit time from the continuation region is optimal, i.e., the continuous stopping time* $\hat{\theta} = \tau_{\hat{\eta}}$ *with the discrete stopping time*

$$\hat{\eta} = \inf\{n \geq 1 : V(x_{\tau_n}, 0) \leq \psi(x_{\tau_n})\} \tag{4.8}$$

satisfies $V(x,y) = J_x(\hat{\theta})$ *as given by* (2.1). *Furthermore, if the intensity of jumps* $y \mapsto \lambda(y)$ *is a continuous function, then the unique solution* V *the HJB equation* (4.2) *belongs to the domain* $\mathcal{D}_{x,y} \subset C_b(E\times]0,\infty[)$ *of the infinitesimal generator* $A_{x,y}$.

Proof. All that is necessary to point out is that the HJB equation (4.3) can be written as

$$V(x,y) = \int_0^\infty \mathrm{e}^{-\alpha t} \mathrm{e}^{-\int_0^t \lambda(y+r)\mathrm{d}r} \lambda(y+t)\Phi(t)\big(\max\{\psi, V(\cdot,0)\}\big)(x)\mathrm{d}t,$$

which is also

$$V(x,y) = \mathbb{E}_{x,y}\{\mathrm{e}^{-\alpha\tau_1} \max\{\psi(x_{\tau_1}), V(x_{\tau_1}, 0))\}\}.$$

This is the HJB equation (3.8) corresponding to the discrete time constrained optimal stopping (DTSP2) with optimal reward $v(x,y)$ given by (2.11).

At this point, we could repeat the discussion in Theorems 3.3–3.5 to conclude that there is a unique solution $V(x,y)$ and the fact that the exit time from the continuation region is optimal.

Furthermore, the HJB equation (4.2) and the assertion that λ is continuous imply that $A_{x,y}V$ is also continuous, i.e., which proves that V belongs to the domain $\mathcal{D}_{x,y}$ of the infinitesimal generator $A_{x,y}$. ∎

REMARK 4.1. The HJB equation (4.2) is similar to the penalized equation of the unconstrained problem, e.g., see Bensoussan and Lions [2]. Similarly, using the same method as in the penalized problem, if λ goes to zero (uniformly), then the solution V_λ converges to the solution (which is a function of x only) of the classical variational inequality of the unconstrained problem. ∎

5. Extensions. Some possible extensions are discussed below, without full details and only with precise indications. A full analysis could take much more space and is not suitable for a short publication.

5.1. Unbounded data. As presented in section 2, the state of the system to be controlled is represented by continuous-time homogeneous Markov process $\{x_t : t \geq 0\}$ with semigroup $\Phi(t)$, infinitesimal generator A, and transition probabilities function $p(x,t,B)$, x in E, $t \geq 0$, B in $\mathcal{B}(E)$, where E is a Polish space (complete, separable and metrizable space) with its Borel σ-algebra $\mathcal{B}(E)$. This general structure includes many situations, but, in proving the results in previous sections 3 and 4, the conditions (3.1) and (3.2), i.e., E compact, $\Phi(t)$ Feller, discount $\alpha > 0$, and reward ψ bounded continuous, and also the key assumption (3.4).

5.1.1. Locally compact or finite dimension. To include more practical models, we need to allow a reward ψ unbounded and a value space E like $\mathbb{R}^d$, i.e., a Polish space locally compact (not necessarily compact). For that purpose more conditions should be imposed on the semigroup $\{\Phi(t) : t \geq 0\}$ of the continuous-time homogeneous Markov process $\{x_t : t \geq 0\}$.

To simplify a little, assume that $E = \mathbb{R}^d$ and for convenience, instead of using $\{x_t : t \geq 0\}$ as the notation of the Markov process, now we switch to $\{X(t,x) : t \geq 0\}$ (in short $X(t,x)$, where x refers to the initial condition at the initial time $t = 0$) as a homogeneous Markov–Feller process on $\mathbb{R}^d$ such that the following hold:

(1) $x \mapsto X(t,x)$ is locally uniformly continuous (in x), locally uniformly continuous for any t in $[0,\infty)$, i.e., for any $\varepsilon > 0$ there is a $\delta > 0$ such that for any x, $\bar{x}$ in $\mathbb{R}^d$ satisfying $|x - \bar{x}| < \delta$, $|x| \leq 1/\varepsilon$ and $|\bar{x}| \leq 1/\varepsilon$ we have

$$P\left\{\sup_{0\leq t\leq 1/\varepsilon} |X(t,x) - X(t,\bar{x})| \geq \varepsilon\right\} < \varepsilon. \tag{5.1}$$

(2) $t \mapsto X(t,x)$ is locally uniformly continuous (in t) for any x in $\mathbb{R}^d$, i.e., for any x in $\mathbb{R}^d$ and for any $\varepsilon > 0$ there is a $\delta > 0$ such that

$$P\left\{\sup_{0\leq t\leq \delta}\ \sup_{0\leq s\leq 1/\varepsilon} |X(t+s,x) - X(s,x)| \geq \varepsilon\right\} < \varepsilon. \tag{5.2}$$

(3) For any $p > 0$ there are positive constants α_0 and μ sufficiently large such that the estimate

$$\mathbb{E}\left\{\sup_{t\geq 0} e^{-\alpha_0 t}(\mu + |X(t,x)|^2)^{p/2}\right\} \leq K_p\,(\mu + |x|^2)^{p/2} \quad \forall t \geq 0,\ x \in \mathbb{R}^d, \tag{5.3}$$

holds, with some $K_p \geq 1$ and $K_p = 1$ if the sup is removed in the left-hand side. To this three conditions, add the following:

(4) $t \mapsto X(t,x)$ is continuous at $t = 0$, locally uniformly in x, i.e., for any $\varepsilon > 0$ there is a $\delta > 0$ such that for any x in $\mathbb{R}^d$ satisfying $|x| \leq 1/\varepsilon$ we have

$$P\left\{\sup_{0\leq t\leq \delta} |X(t,x) - x| \geq \varepsilon\right\} < \varepsilon. \tag{5.4}$$

Note that these properties are not easily expressed in terms of the semigroup $\{\Phi(t) : t \geq 0\}$, and therefore it is necessary to change of notation from x_t, P_x, and $\mathbb{E}_x$ to $X(x,t)$, P, and $\mathbb{E}$. Here, the expressions $P\{\cdot\}$ and $\mathbb{E}\{\cdot\}$ refer to the probability and the mathematical expectation relative to the canonical probability space (satisfying the usual conditions), where the Markov–Feller process is defined.

If $\Phi_\alpha(t)$ with a properly selected $\alpha \geq 0$ denotes the semigroup associated with the Markov–Feller process, i.e.,

$$\Phi_\alpha(t)h(x) = \mathbb{E}\{e^{-\alpha t}h(y(t,x))\}, \tag{5.5}$$

then $\Phi_\alpha(t)$ is a *strongly continuous* semigroup on $C_0(\mathbb{R}^d) = C_b(\mathbb{R}^d)$, the space of real (uniformly) continuous and bounded functions in $\mathbb{R}^d$. However, to include a reward function ψ with a polynomial growth, another space is used.

Let $C_p(\mathbb{R}^d)$ be the space of real uniformly continuous functions on any ball and with a growth bounded by the norm to the $p \geq 0$ power, in other words, the space of real functions h on $\mathbb{R}^d$ such that $x \mapsto h(x)(1+|x|^2)^{-p/2}$ is bounded and locally uniformly continuous, with the weighted sup-norm

$$\|h\| = \|h\|_{C_p} := \sup_{x\in\mathbb{R}^d}\{|h(x)|(\mu+|x|^2)^{-p/2}\}, \tag{5.6}$$

where μ is a positive constant sufficiently large to so that estimate (5.3) holds. It is clear that $C_b(\mathbb{R}^d) \subset C_q(\mathbb{R}^d) \subset C_p(\mathbb{R}^d)$ for any $0 \leq q < p$. Then the (linear) semigroup $\{\Phi_\alpha(t),\, t \geq 0\}$ with an α–exponential factor is a weakly continuous Markov–Feller semigroup in the space $C_p(\mathbb{R}^d)$, i.e.,

$$\begin{cases} \Phi_\alpha(t+s) = \Phi_\alpha(t)\Phi_\alpha(s) \quad \forall\, s,t \geq 0, \\ \|\Phi_\alpha(t)h\| \leq \|h\| \quad \forall\, h \in C_p(\mathbb{R}^d), \\ \Phi_\alpha(t)h(x) \to h(x) \;\text{ as }\; t \to 0 \quad \forall\, h \in C_p(\mathbb{R}^d), \\ \Phi_\alpha(t)h(x) \geq 0 \quad \forall\, h \geq 0, \quad h \in C_p(\mathbb{R}^d). \end{cases} \tag{5.7}$$

This follows immediately from the conditions (5.1), (5.2), and (5.3) imposed on the Markov–Feller process $y(t,x)$, provided $\alpha \geq \alpha_0$ with α_0 aad μ as in (5.3). Conditions (5.3) and (5.4) imply

$$\|\Phi_\alpha(t)h \to h\| \to 0 \;\text{ as }\; t \to 0 \quad \forall\, h \in C_p(\mathbb{R}^d), \tag{5.8}$$

which make $\{\Phi_\alpha(t),\, t \geq 0\}$ a strongly continuous semigroup in the Banach space $C_p(\mathbb{R}^d)$.

At this point, the arguments discussed in sections 3 and 4 can be repeated with $C_p(\mathbb{R}^d)$ instead of $C(E)$, E compact, and with a reward function ψ in $C_p(\mathcal{O})$. Certainly, depending on the growth (given by $p > 0$) assumption (5.3) on $\alpha_0 = \alpha_0(p,\lambda)$ shows its strength.

In the present case, u_0 or u will belong to $C_q(\mathbb{R}^d)$ for some q and, for the proof of the properties of Theorems 3.2 and 4.1, the uniform integrability of $u_0(x_{\tau_n})\exp(-\alpha\tau_n)$ will be obtained from (5.3).

REMARK 5.1. A good prototype where the above properties are satisfied is the diffusion with jumps as in Menaldi [14]. Moreover, the case where E is a convenient (unbounded) subset of $\mathbb{R}^d$ can also be treated with suitable modifications. □

5.1.2. Nonlocally compact or infinite dimension. If the continuous-time homogeneous Markov process $\{x_t : t \geq 0\}$ takes values in a Polish space E which is not locally compact, then another problem appears. For instance, if E is a Banach of Hilbert space of infinite dimension, then a prototype could be stochastic partial differential equations (stochastic PDE), where the semigroup $\{\Phi(t) : t \geq 0\}$ is not strongly continuous.

For instance, in trying to use the general semigroup formulation (on a separable Banach space E other than the Euclidean space) one finds the typical situation when the semigroup is not continuous, i.e., the semigroup corresponds to a strong Markov process, but not exactly a Feller process, e.g., the semigroup is defined on $E = L^2$, but it maps continuous functions on $E = H^1$ into continuous functions on $E = L^2$. For instance, this is the case of a stochastic PDE, of which an extreme situation is the stochastic Navier–Stokes equation, i.e., Menaldi and Sritharan [17]. In this case, there is a reasonable set of assumptions to be used, namely, (5.1), (5.2), (5.3), but certainly not (5.4). In this case we only has a weakly continuous Markov–Feller semigroup in the space $C_p(\mathbb{R}^d)$ satisfying (5.7), but not (5.8).

Since the semigroup is not strongly continuous, we cannot consider the *strong* infinitesimal generator as acting on a dense domain in $C_p(\mathbb{R}^d)$. However, this Markov–Feller semigroup $(\Phi_\alpha(t),\, t \geq 0)$ may be considered as acting on real Borel functions with p-polynomial growth, which is a Banach space with the norm (5.6) and is denoted by $B_p(\mathbb{R}^d)$. It is convenient to define the family of seminorms on $B_p(\mathbb{R}^d)$

$$p_0(h, x) := E\left\{\sup_{s \geq 0} |h(X(s, x))|\, e^{-\alpha_0 s}\right\} \quad \forall x \in \mathbb{R}^d, \tag{5.9}$$

where $2\alpha_0$ satisfies the estimate (5.3) with $2p$ and μ, and when $p = 0$ we may take $\alpha_0 = 0$. If a sequence $\{h_n\}$ of equi-bounded functions in $B_p(\mathbb{R}^d)$ satisfies $p_0(h_n - h, x) \to 0$ for any x in $\mathbb{R}^d$, we say that $h_n \to h$ boundedly pointwise relative to the above family of seminorms. In view of (5.2), it is clear that $p_0(\Phi_\alpha(t)h - h, x) \to 0$ as $t \to 0$, for every function h in $C_p(\mathbb{R}^d)$ and every x in $\mathbb{R}^d$.

DEFINITION 5.1. *Let $\bar{C}_p(\mathbb{R}^d)$ be the subspace of functions $\bar{h}$ in $B_p(\mathbb{R}^d)$ such that the mapping $t \mapsto \bar{h}[X(t, x)]$ is almost surely continuous on $[0, \infty)$ for any x in $\mathbb{R}^d$ and satisfies*

$$\lim_{t \to 0} p_0(\Phi_\alpha(t)\bar{h} - \bar{h}, x) = 0 \quad \forall\, x \in \mathbb{R}^d, \tag{5.10}$$

where $p_0(\cdot, \cdot)$ is the seminorm given by (5.9)*, and if $\alpha_0(p)$ denotes the constant appearing in* (5.3)*, then $2\alpha > \alpha_0(2p)$ with $p \geq 0$ and $\alpha_0(0) = 0$.* □

This is the space of functions (uniformly) continuous over the random field $X(\cdot, x)$, relative to the family of seminorms (5.9), and it is independent of α, as long as (5.3) holds. Hence, we may consider the semigroup $\{\Phi_\alpha(t) : t \geq 0\}$. Then, the tools exist to treat the optimal stopping problem for a class of infinite dimensional processes. Details can be found in Menaldi [16].

5.2. Other types of signals. There are several variants on this model; only a couple of them are mentioned below.

5.2.1. Pure jump Markov processes. Suppose we are given a pure jump Markov process $\{\xi_t : t \geq 0\}$ with infinitesimal generator

$$A_1\varphi(\xi) = \lambda(\xi)\big[Q\varphi(\xi) - \varphi(\xi)\big], \qquad Q\varphi(\xi) = \int \varphi(\zeta)q(\xi, \mathrm{d}\zeta)$$

with a Markovian kernel $q(\xi, \mathrm{d}\zeta)$ on a space E_1 and a bounded continuous intensity $\lambda(\xi)$. Then, define the signals $\{\tau_n : n \geq 1\}$ as the sequence of jumps times of the process $\{\xi_t : t \geq 0\}$. All this leads to the problem

$$V(x,\xi) = sup\big\{\mathbb{E}_{x,\xi}\{\mathrm{e}^{-\alpha\theta}\psi(x_\theta)\} : \theta \text{ admissible}\big\}$$

with the HJB equation either

$$A_x V + A_1 V - \alpha V + \lambda(\xi)[\psi - QV]^+ = 0$$

or

$$A_x V(x,y) - \big(\alpha + \lambda(\xi)\big)V(x,y) + \lambda(\xi)\max\big\{\psi(x), QV(x,y)\big\}.$$

The same kind of analysis as before allows one to obtain the existence and uniqueness of the solution of the HJB equation and the verification theorem.

Note that this case is not included in the IID case treated early (when π_0 is not exponential).

5.2.2. Semi-Markov processes. The case where the process giving the signals is a semi-Markov process allows one to cover both the IID case and the pure jump Markov processes. Assume that $\{y_t^1 : t \geq 0\}$ is a semi-Markov process with values in a space E_1 (with the discrete topology and the Borel σ-algebra) and $\{y_t = (y_t^1, y_t^2) : t \geq 0\}$ is the appropriated Markov process where $\{y_t^2 : t \geq 0\}$ is the elapsed time since the last jump of $\{y_t^1 : t \geq 0\}$, e.g., see Jacod [10], Gikhman and Skorokhod [8], and Robin [18], among others.

If we are given $\lambda : E_1 \times [0,\infty[\longrightarrow [0,\infty[$ satisfying

$$0 \leq \lambda(y^1, y^2) \leq M, \qquad y_2 \mapsto \lambda(y^1, y^2) \quad \text{continuous}$$

and a transition probability $q(y^1, y^2, \Gamma)$ with

$$y^2 \mapsto \int_{E_1} q(y^1, y^2, \mathrm{d}z)\varphi(z) \quad \text{continuous},$$

for every φ bounded measurable on E_1, one can show that $\{y_t : t \geq 0\}$ can be constructed as a Markov process with infinitesimal generator

$$A_1 f = \frac{\partial f}{\partial y^2} + \lambda(y^1, y^2)\left[\int_{E_1} q(y^1, y^2, \mathrm{d}z) f(z, 0) - f(y^1, y^2)\right].$$

Then, one can state the optimal stopping problem with constraint in the same way as before for $V(x,y) = V(x, y^1, y^2)$ with $y = (y^1, y^2)$ and obtain the HJB equation

$$A_x V + A_1 V - \alpha V + \lambda(y)[\psi - QV]^+ = 0,$$

where

$$QV(y) = QV(y^1, y^2) = \int_{E_1} q(y^1, y^2, \mathrm{d}z) V(x, z, y^2).$$

It is not difficult to prove similar results to section 3 and 4 for this class of processes and the above assumptions.

5.2.3. Piecewise-deterministic Markov processes. Refer to Davis [4] (or the book by Davis [5]) for the definition of piecewise-deterministic Markov processes (PDMPs). This class of Markov process contains both pure jump Markov processes and the IID case in the sense that the process $\{y_t :\geq 0\}$ of the IID case is a PDMP.

Actually, in the IID case, if $N_t = \sum_n \mathbb{1}_{\tau_n < t}$, then $\{N_t : t \geq 0\}$ is a semi-Markov process and the couple (N_t, y_t) is a PDMP.

Let us assume that $\{y_t :\geq 0\}$ is a PDMP with values in a space E_1 and satisfies the conditions of Davis [4, p. 77] in order to be a Feller process, namely, λ is bounded and continuous on E_1 and $y \mapsto Qf(y)$ is continuous for any f bounded continuous, where λ is the intensity of the jumps and Q is the Markovian kernel of the jumps. This means that the infinitesimal generator of $\{y_t :\geq 0\}$ is

$$A_1 f(y) = Df(y) + \lambda(y) \left[\int_{E_1} q(y, \mathrm{d}z) f(z) - f(z) \right] \tag{5.11}$$

with Df being the infinitesimal generator of a deterministic movement.

In that framework, we have again a similar HJB equation for $V(x, y)$ (with y being multidimensional here), and one can see that the same method as before can be applied to solve the constrained optimal stopping problem.

5.2.4. Diffusion processes with jumps. It is clear that diffusion processes with jumps could be used as a model for the signal, if the jumps do not accumulate, i.e., the source of the jumps is a composed Poisson process, instead of any general Poisson martingale measure (e.g., see Applebaum [1], Ikeda and Watanabe [9], Menaldi [14], among many others books). In this case, if the first order partial differential operator Df in the expression (5.11) of the infinitesimal generator A_1 in the previous section on PDMP is replaced by a second order partial differential operator Af corresponding to a diffusion process, then the whole argument in section 5.2.3 can be repeated to include the class of diffusion with jumps, where the jumps are generated by a composed Poisson process.

5.3. Impulse control problems. A typical impulse control problem is given by a controlled process such that

$$\mathrm{d}x_t = b\mathrm{d}t + \sigma \mathrm{d}t + \sum_{i \geq 1} \xi_i \delta(t - \theta_i), \qquad x_0 = x,$$

(with δ being the delta distribution) where the control is the sequence $\nu = \{\xi_i, \theta_i : i \geq 1\}$ of stopping times θ_i, $\theta_{i+1} \geq \theta_i$, at which ξ_i is added to the state $x_{\theta_i -}$ and can be chosen in order to minimize the cost

$$J(\nu) = \mathbb{E} \left\{ \int_0^\infty \mathrm{e}^{-\alpha t} f(x_t) \mathrm{d}t + \sum_{i \geq 1} \mathrm{e}^{-\alpha \theta_i} k(\xi_i) \right\},$$

where $k(\xi)$ is the cost of impulses. It is know (see Bensoussan and Lions [3], and also, e.g., Menaldi [15], Robin [18], among many others) that under suitable assumptions the optimal cost function $u(x)$ is the solution of a quasi-variational inequality

$$\begin{cases} -Au + \alpha u \leq f, \\ u \leq Mu := \inf_{\xi} \big\{ k(\xi) + u(x + \xi) \big\}, \\ (-Au + \alpha u - f)(u - Mu) = 0. \end{cases}$$

It is also known that u is the limit of a sequence of optimal stopping problems u_n. One can consider the above problem with a constraint such that θ_i has to be chosen in a sequence of signals $\{\tau_n : n \geq 1\}$ independent of x_t. The treatment of this kind of impulse control with constraint is the subject of on going work.

REFERENCES

[1] D. APPLEBAUM, *Lévy Processes and Stochastic Calculus*, Cambridge University Press, Cambridge, 2004.

[2] A. BENSOUSSAN AND J.-L. LIONS, *Applications des Inéquations Variationnelles en Contrôle Stochastique*, Dunod, Paris, 1978.

[3] A. BENSOUSSAN AND J.-L. LIONS, *Contrôle Impulsionnel et Inéquations Quasi-Variationnelles*, Gauthier-Villars, Paris, 1982.

[4] M. H. A. DAVIS, *Piecewise-deterministic Markov processes: A general class of nondiffusion stochastic models*, J. Roy. Statist. Soc. Ser. B, 46 (1984), pp. 353–388.

[5] M. H. A. DAVIS, *Markov Models and Optimization*, Chapman & Hall, London, 1993.

[6] P. DUPUIS AND H. WANG, *Optimal stopping with random intervention times*, Adv. Appl. Probab., 34 (2002), pp. 141–157.

[7] D. EGLOFF AND M. LEIPPOLD, *The valuation of American options with stochastic stopping time constraints*, Appl. Math. Finance, 16 (2009), pp. 287–305.

[8] I.I. GIKHMAN AND A.V. SKOROKHOD, *The Theory of Stochastic Processes.* II, Springer-Verlag, Berlin, 2004.

[9] N. IKEDA AND S. WATANABE, *Stochastic Differential Equations and Diffusion Processes*, 2nd ed., North-Holland, Amsterdam, 1989.

[10] J. JACOD, *Semi-groupes et mesures invariantes pour les processus semi-Markoviens à espace d'état quelconque*, Ann. Inst. H. Poincaré Sect. B, 9 (1973), pp. 77–112.

[11] J. LEMPA, *Optimal stopping with information constraint*, Appl. Math. Optim., 66 (2012), pp. 147–173.

[12] G. LIANG, *Stochastic control representations for penalized backward stochastic differential equations*, SIAM J. Control Optim., 53 (2015), pp. 1440–1463.

[13] G. LIANG AND W. WEI, *Optimal switching at poisson random intervention times*, Discrete Contin. Dyn. Syst. Ser. B, to appear.

[14] J.L. MENALDI, *Stochastic Differential Equations with Jumps*, http://math.wayne.edu/~menaldi/papers/m-book-06-2013.pdf (September 30, 2012).

[15] J.L. MENALDI, *Optimal impulse control problems for degenerate diffusions with jumps*, Acta Appl. Math., 8 (1987), pp. 165–198.

[16] J.L. MENALDI, *Stochastic hybrid optimal control models*, Aportaciones Mat. Investig., Soc. Mat. Mexicana 16, 2001, pp. 205–250.

[17] J.L. MENALDI AND S.S. SRITHARAN, *Impulse control of stochastic Navier-Stokes equations*, Nonlinear Anal., 52 (2003), pp. 357–381.

[18] M. ROBIN, *Contrôle impulsionnel des processus de Makov*, Thèse d'état, 1978; also available online from https://hal.archives-ouvertes.fr/tel-00735779/document.

[19] A. N. SHIRYAEV, *Optimal Stopping Rules*, Springer-Verlag, Berlin, 1978.