Communications in Applied Mathematics and Computational Science

NUMERICAL METHOD FOR EXPECTATIONS OF PIECEWISE DETERMINISTIC MARKOV PROCESSES

ADRIEN BRANDEJSKY, BENOÎTE DE SAPORTA AND FRANÇOIS DUFOUR

vol. 7 no. 1 2012

mathematical sciences publishers

COMM. APP. MATH. AND COMP. SCI.
Vol. 7, No. 1, 2012

msp

NUMERICAL METHOD FOR EXPECTATIONS OF PIECEWISE DETERMINISTIC MARKOV PROCESSES

ADRIEN BRANDEJSKY, BENOÎTE DE SAPORTA AND FRANÇOIS DUFOUR

We present a numerical method to compute expectations of functionals of a piecewise deterministic Markov process. We discuss time dependent functionals as well as deterministic time horizon problems. Our approach is based on the quantization of an underlying discrete-time Markov chain. We obtain bounds for the rate of convergence of the algorithm. The approximation we propose is easily computable and is flexible with respect to some of the parameters defining the problem. An example illustrates the paper.

1. Introduction

The aim of this paper is to propose a practical numerical method to approximate some expectations related to a piecewise deterministic Markov process thanks to the quantization of a discrete-time Markov chain naturally embedded within the continuous-time process.

Piecewise deterministic Markov processes (PDMP's) have been introduced by M. H. A. Davis in [5] as a general class of stochastic models. PDMP's are a

This work was supported by ARPEGE program of the French National Agency of Research (ANR), project "FAUTOCOES", number ANR-09-SEGI-004.
MSC2010: primary 60J25, 65C20; secondary 60K10.
Keywords: expectation, piecewise deterministic Markov processes, quantization, numerical method.

family of Markov processes involving deterministic motion punctuated by random jumps. The motion depends on three local characteristics namely the flow Φ, the jump rate λ and the transition measure Q, which specifies the postjump location. Starting from the point x, the motion of the process follows the flow $\Phi(x,t)$ until the first jump time T_1, which occurs either spontaneously in a Poisson-like fashion with rate $\lambda(\Phi(x,t))$ or when the flow $\Phi(x,t)$ hits the boundary of the state space. In either case, the location of the process at the jump time T_1 is selected by the transition measure $Q(\Phi(x,T_1),\cdot)$ and the motion restarts from this new point X_{T_1} denoted by Z_1. We define similarly the time S_2 until the next jump, $T_2 = T_1 + S_2$ with the next postjump location defined by $Z_2 = X_{T_2}$ and so on. Thus, associated to the PDMP we have the discrete-time Markov chain $(Z_n, S_n)_{n\in\mathbb{N}}$, given by the postjump locations and the interjump times. A suitable choice of the state space and the local characteristics Φ, λ and Q provides stochastic models covering a great number of problems of operations research as described in [5, Section 33].

We are interested in the approximation of expectations of the form

$$\boldsymbol{E}_x\left[\int_0^{T_N} l(X_t)\,dt + \sum_{j=1}^{N} c(X_{T_j^-})\mathbb{1}_{\{X_{T_j^-}\in\partial E\}}\right]$$

where $(X_t)_{t\geq 0}$ is a PDMP and l and c are some nonnegative, real-valued, bounded functions and ∂E is the boundary of the domain. Such expectations are discussed by M. H. A. Davis in [5, Chapter 3]. They often appear as cost or reward functions in optimization problems. The first term is referred to as the running cost while the second may be called the boundary jump cost. Besides, they are quite general since Davis shows how a "wide variety of apparently different functionals" can be obtained from the above specific form. For example, this wide variety includes quantities such as a mean exit time and even, for any fixed $t \geq 0$, the distribution of X_t (that is, $\boldsymbol{E}_x[\mathbb{1}_F(X_t)]$ where F is a measurable set).

There are surprisingly few works in the literature devoted to the actual computation of such expectations, using other means than direct Monte Carlo simulations. Davis showed that these expectations satisfy integrodifferential equations. However, the set of partial differential equations that is obtained is unusual. Roughly speaking, these differential equations are basically transport equations with a nonconstant velocity and they are coupled by the boundary conditions and by some integral terms involving kernels that are derived from the properties of the underlying stochastic process. The main difficulty comes from the fact that the domains on which the equations have to be solved vary from one equation to another making their numerical resolution highly problem specific. Another similar approach has been recently investigated in [4; 7]. It is based on a discretization of the Chapman Kolmogorov equations satisfied by the distribution of the process $(X_t)_{t\geq 0}$. The

authors propose an approximation of such expectations based on finite volume methods. Unfortunately, their method is only valid if there are no jumps at the boundary. Our approach is completely different and does not rely on differential equations, but on the fact that such expectations can be computed by iterating an integral operator G. This operator only involves the embedded Markov chain $(Z_n, S_n)_{n\in\mathbb{N}}$ and conditional expectations. It is therefore natural to propose a computational method based on the quantization of this Markov chain, following the same idea as [6].

There exists an extensive literature on quantization methods for random variables and processes. The interested reader may for instance consult [8], [9] and the references within. Quantization methods have been developed recently in numerical probability or optimal stochastic control with applications in finance (see [1; 2; 9], for instance). The quantization of a random variable X consists in finding a finite grid such that the projection $\widehat{X}$ of X on this grid minimizes some L^p norm of the difference $X-\widehat{X}$. Roughly speaking, such a grid will have more points in the areas of high density of X. As explained for instance in [9, Section 3], under some Lipschitz-continuity conditions, bounds for the rate of convergence of functionals of the quantized process towards the original process are available.

In the present work, we develop a numerical method to compute expectations of functionals of the above form where the cost functions l and c satisfy some Lipschitz-continuity conditions. We first recall the results presented by Davis according to whom, the above expectation may be computed by iterating an operator denoted by G. Consequently, it appears natural to follow the idea developed in [6] namely to express the operator G in terms of the underlying discrete-time Markov chain $(Z_n, S_n)_{n\in\mathbb{N}}$ and to replace it by its quantized approximation. Moreover, in order to prove the convergence of our algorithm, we replace the indicator function $\mathbb{1}_{\{X_{T_j^-}\in\partial E\}}$ contained within the functional by some Lipschitz continuous approximation. Bounds for the rate of convergence are then obtained. However, and this is the main contribution of this paper, we then tackle two important aspects that had not been investigated in [6].

The first aspect consists in allowing c and l to be time-dependent functions, although still Lipschitz continuous, so that we may compute expectations of the form

$$\boldsymbol{E}_x\left[\int_0^{T_N} l(X_t,t)\,dt+\sum_{j=1}^{N} c(X_{T_j^-},T_j)\mathbb{1}_{\{X_{T_j^-}\in\partial E\}}\right].$$

This important generalization has huge applicative consequences. For instance, it allows discounted cost or reward functions such as $l(x,t)=e^{-\delta t}l(x)$ and $c(x,t)=e^{-\delta t}c(x)$ where δ is some interest rate. To compute the above expectation, our strategy consists in considering, as suggested by Davis in [5], the time-augmented

process $\widetilde{X}_t = (X_t, t)$. Therefore, a natural way to deal with the time-dependent problem is to apply our previous approximation scheme to the time-augmented process $(\widetilde{X}_t)_{t\geq 0}$. However, it is far from obvious, that the assumptions required by our numerical method still hold for this new PDMP $(\widetilde{X}_t)_{t\geq 0}$.

The second important generalization is to consider the deterministic time horizon problem. Indeed, it seems crucial, regarding the applications, to be able to approximate

$$\boldsymbol{E}_x\left[\int_0^{t_f} l(X_t,t)\,dt + \sum_{T_f\leq t_f} c(X_{T_f^-},T_f)\mathbb{1}_{\{X_{T_f^-}\in\partial E\}}\right]$$

for some fixed $t_f > 0$ regardless of how many jumps occur before this deterministic time. To compute this quantity, we start by choosing a time N such that $\boldsymbol{P}(T_N < t_f)$ be small so that the previous expectation boils down to

$$\boldsymbol{E}_x\left[\int_0^{T_N} l(X_t,t)\mathbb{1}_{\{t\leq t_f\}}dt + \sum_{j=1}^{N} c(X_{T_j^-},T_j)\mathbb{1}_{\{X_{T_j^-}\in\partial E\}}\mathbb{1}_{\{T_j\leq t_f\}}\right].$$

At first sight, this functional seems to be of the previous form. Yet, one must recall that Lipschitz continuity conditions have been made concerning the cost functions so that the indicator functions $\mathbb{1}_{\{\cdot\leq t_f\}}$ prevent a direct application of the earlier results. We deal with the two indicator functions in two different ways. On the one hand, we prove that it is possible to relax the regularity condition on the running cost function so that our algorithm still converges in spite of the first indicator function. On the other hand, since the same reasoning cannot be applied to the indicator function within the boundary jump cost term, we bound it between two Lipschitz continuous functions. This provides bounds for the expectation of the deterministic time horizon functional.

An important advantage of our method is that it is flexible. Indeed, as pointed out in [1], a quantization based method is "obstacle free" which means, in our case, that it produces, once and for all, a discretization of the process independently of the functions l and c since the quantization grids merely depend on the dynamics of the process. They are only computed once, stored off-line and may therefore serve many purposes. Once they have been obtained, we are able to approximate very easily and quickly any of the expectations described earlier. This flexibility is definitely an important advantage of our scheme over standard methods such as Monte Carlo simulations since, with such methods, we would have to run the whole algorithm for each expectation we want to compute. This point is illustrated in Section 6 where we easily solve an optimization problem that would be very laboriously handled by Monte Carlo simulations.

The paper is organized as follows. We first recall, in Section 2, the definition of a PDMP and state our assumptions. In Section 3, we introduce the recursive method to compute the expectation. Section 4 presents the approximation scheme and a bound for the rate of convergence. The main contribution of the paper lies in Section 5, which contains generalizations to time-dependent parameters and deterministic time-horizon problems. The paper is illustrated by a numerical example in Section 6; a conclusion (Section 7) is followed by some appendixes containing technical results.

2. Definitions and assumptions

For all metric space E, we denote by $\mathcal{B}(E)$ its Borel σ-field and $B(E)$ the set of real-valued, bounded and measurable functions defined on E. For $a, b \in \mathbb{R}$, set $a \wedge b = \min(a, b)$, $a \vee b = \max(a, b)$, and $a^+ = a \vee 0$.

Definition of a PDMP. In this first section, let us define a piecewise deterministic Markov process and introduce some general assumptions. Let M be a finite set called the set of the modes that will represent the different regimes of evolution of the PDMP. For each $m \in M$, the process evolves in E_m, an open subset of $\mathbb{R}^d$. Let

$$E = \{(m, \zeta), m \in M, \zeta \in E_m\}.$$

This is the state space of the process $(X_t)_{t\in\mathbb{R}^+} = (m_t, \zeta_t)_{t\in\mathbb{R}^+}$. Let ∂E be its boundary and $\overline{E}$ its closure and for any subset Y of E, Y^c denotes its complement.

A PDMP is defined by its local characteristics $(\Phi_m, \lambda_m, Q_m)_{m\in M}$.

- For each $m \in M$, $\Phi_m : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$ is a continuous function called the flow in mode m. For all $t \in \mathbb{R}$, $\Phi_m(\cdot, t)$ is an homeomorphism and $t \to \Phi_m(\cdot, t)$ is a semigroup; i.e., for all $\zeta \in \mathbb{R}^d$, $\Phi_m(\zeta, t + s) = \Phi_m(\Phi_m(\zeta, s), t)$. For all $x = (m, \zeta) \in E$, define the deterministic exit time from E:

  $$t^*(x) = \inf\{t > 0 \text{ such that } \Phi_m(\zeta, t) \in \partial E_m\}.$$

  We use here and throughout the convention $\inf \varnothing = +\infty$.

- For all $m \in M$, the jump rate $\lambda_m : \overline{E}_m \to \mathbb{R}^+$ is measurable, and for all $(m, \zeta) \in E$, there exists $\epsilon > 0$ such that

  $$\int_0^\epsilon \lambda_m(\Phi_m(\zeta, t))\, dt < +\infty.$$

- For all $m \in M$, Q_m is a Markov kernel on $(\mathcal{B}(\overline{E}), \overline{E}_m)$ that satisfies

  $$Q_m(\zeta, \{(m, \zeta)\}^c) = 1 \quad \text{for all } \zeta \in \overline{E}_m.$$

From these characteristics, it can be shown (see [5]) that there exists a filtered probability space $(\Omega, \mathscr{F}, \mathscr{F}_t, (\boldsymbol{P}_x)_{x\in E})$ on which a process $(X_t)_{t\in\mathbb{R}^+}$ is defined. Its motion, starting from a point $x \in E$, may be constructed as follows. Let T_1 be a nonnegative random variable with survival function

$$\boldsymbol{P}_x(T_1 > t) = \begin{cases} e^{-\Lambda(x,t)} & \text{if } 0 \le t < t^*(x), \\ 0 & \text{if } t \ge t^*(x), \end{cases}$$

where for $x = (m, \zeta) \in E$ and $t \in [0, t^*(x)]$,

$$\Lambda(x,t) = \int_0^t \lambda_m(\Phi_m(\zeta, s))\, ds.$$

One then chooses an E-valued random variable Z_1 according to the distribution $Q_m(\Phi_m(\zeta, T_1), \cdot)$. The trajectory of X_t for $t \le T_1$ is

$$X_t = \begin{cases} (m, \Phi_m(\zeta, t)) & \text{if } t < T_1, \\ Z_1 & \text{if } t = T_1. \end{cases}$$

Starting from the point $X_{T_1} = Z_1$, one then selects in a similar way $S_2 = T_2 - T_1$ the time between T_1 and the next jump, Z_2 the next postjump location and so on. Davis shows, in [5], that the process so defined is a strong Markov process $(X_t)_{t\ge 0}$ with jump times $(T_n)_{n\in\mathbb{N}}$ (with $T_0 = 0$). The process $(\Theta_n)_{n\in\mathbb{N}} = (Z_n, S_n)_{n\in\mathbb{N}}$ where $Z_n = X_{T_n}$ is the postjump location and $S_n = T_n - T_{n-1}$ (with $S_0 = 0$) is the n-th interjump time is clearly a discrete-time Markov chain.

The following assumption about the jump-times is standard (see [5, Section 24], for example):

Assumption 2.1. For all $(x,t) \in E \times \mathbb{R}^+$, $\boldsymbol{E}_x\left[\sum_k \mathbb{1}_{\{T_k<t\}}\right] < +\infty$.

It implies in particular that T_k goes to infinity a.s. when k goes to infinity.

Notation and assumptions. For notational convenience, any function h defined on E will be identified with its component functions h_m defined on E_m. Thus, one may write

$$h(x) = h_m(\zeta) \text{ when } x = (m, \zeta) \in E.$$

We also define a generalized flow $\Phi : E \times \mathbb{R}^+ \to E$ such that

$$\Phi(x,t) = (m, \Phi_m(\zeta, t)) \text{ when } x = (m, \zeta) \in E.$$

Define on E the following distance, for $x = (m, \zeta)$ and $x' = (m', \zeta') \in E$:

$$|x - x'| = \begin{cases} +\infty & \text{if } m \ne m', \\ |\zeta - \zeta'| & \text{otherwise.} \end{cases} \tag{1}$$

For any function w in $B(\overline{E})$, introduce the notation

$$Qw(x) = \int_E w(y)Q(x,dy), \qquad C_w = \sup_{x\in\overline{E}} |w(x)|,$$

and for any Lipschitz continuous function w in $B(E)$, denote by $[w]^E$, or if there is no ambiguity by $[w]$, its Lipschitz constant:

$$[w]^E = \sup_{x\neq y\in E} \frac{|w(x)-w(y)|}{|x-y|},$$

with the convention $\dfrac{1}{\infty} = 0$.

Remark 2.2. For $w \in B(\overline{E})$ and from the definition of the distance on E, one has $[w] = \max_{m\in M}[w_m]$.

Definition 2.3. Denote by $\boldsymbol{L}_c(E)$ the set of functions $w \in B(E)$ that are Lipschitz continuous along the flow; i.e., the real-valued, bounded, measurable functions defined on E and satisfying the following conditions:

- For all $x \in E$, the map $w(\Phi(x,\cdot)) : [0, t^*(x)) \to \mathbb{R}$ is continuous, and the limit $\lim_{t\to t^*(x)} w(\Phi(x,t))$ exists and is denoted by $w\big(\Phi(x,t^*(x))\big)$.
- There exists $[w]_1^E \in \mathbb{R}^+$ such that for all $x, y \in E$ and $t \in [0, t^*(x) \wedge t^*(y)]$, one has
$$\big|w(\Phi(x,t)) - w(\Phi(y,t))\big| \leq [w]_1^E |x-y|.$$
- There exists $[w]_2^E \in \mathbb{R}^+$ such that for all $x \in E$ and $t, u \in [0, t^*(x)]$, one has
$$\big|w(\Phi(x,t)) - w(\Phi(x,u))\big| \leq [w]_2^E |t-u|.$$
- There exists $[w]_*^E \in \mathbb{R}^+$ such that for all $x, y \in E$, one has
$$\big|w(\Phi(x,t^*(x))) - w(\Phi(y,t^*(y)))\big| \leq [w]_*^E |x-y|.$$

Denote by $\boldsymbol{L}_c(\partial E)$ the set of real-valued, bounded, measurable functions defined on ∂E satisfying the following condition:

- There exists $[w]_*^{\partial E} \in \mathbb{R}^+$ such that for all $x, y \in E$, one has
$$\big|w(\Phi(x,t^*(x))) - w(\Phi(y,t^*(y)))\big| \leq [w]_*^{\partial E} |x-y|.$$

Remark 2.4. When there is no ambiguity, we will use the notation $[w]_i$ instead of $[w]_i^E$ for $i \in \{1, 2, *\}$ and $[w]_*$ instead of $[w]_*^{\partial E}$.

Remark 2.5. In Definition 2.3, we used the generalized flow for notational convenience. For instance, the definition of $[w]_1$ is equivalent to the following: for all

$m \in M$, there exists $[w_m]_1 \in \mathbb{R}^+$ such that for all $\zeta, \zeta' \in E_m$ and $t \in [0, t^*(m,\zeta) \wedge t^*(m,\zeta')]$, one has

$$\left|w_m(\Phi_m(\zeta,t)) - w_m(\Phi_m(\zeta',t))\right| \leq [w_m]_1 |\zeta - \zeta'|.$$

Let $[w]_1 = \max_{m\in M}[w_m]_1$.

Definition 2.6. For all $u \geq 0$, denote by $\boldsymbol{L}_c^u(E)$ the set of functions $w \in B(E)$ Lipschitz continuous along the flow until time u; i.e., the real-valued, bounded, measurable functions defined on E and satisfying the following conditions:

- For all $x \in E$, the map $w(\Phi(x,\cdot)) : [0, t^*(x) \wedge u) \to \mathbb{R}$ is continuous. If $t^*(x) \leq u$, then $\lim_{t\to t^*(x)} w(\Phi(x,t))$ exists and is denoted by $w\big(\Phi(x,t^*(x))\big)$.
- There exists $[w]_1^{E,u} \in \mathbb{R}^+$ such that for all $x, y \in E$ and $t \in [0, t^*(x) \wedge t^*(y) \wedge u]$, one has
$$\left|w(\Phi(x,t)) - w(\Phi(y,t))\right| \leq [w]_1^{E,u} |x - y|.$$
- There exists $[w]_2^{E,u} \in \mathbb{R}^+$ such that for all $x \in E$ and $t, t' \in [0, t^*(x) \wedge u]$, one has
$$\left|w(\Phi(x,t)) - w(\Phi(x,t'))\right| \leq [w]_2^{E,u} |t - t'|.$$
- There exists $[w]_*^{E,u} \in \mathbb{R}^+$ such that for all $x, y \in E$, if $t^*(x) \leq u$ and $t^*(y) \leq u$, one has
$$\left|w(\Phi(x,t^*(x))) - w(\Phi(y,t^*(y)))\right| \leq [w]_*^{E,u} |x - y|.$$

Remark 2.7. For all $u \leq u'$, one has $\boldsymbol{L}_c^{u'}(E) \subset \boldsymbol{L}_c^u(E)$ with $[w]_i^{E,u} \leq [w]_i^{E,u'}$ where $i \in \{1, 2, *\}$.

Remark 2.8. Definitions 2.3 and 2.6 correspond respectively to the Lipschitz and local Lipschitz continuity along the flow that is, along the trajectories of the process. They can be replaced by (local) Lipschitz assumptions on the flow Φ, t^* and w in the classical sense.

We will require the following assumptions.

Assumption 2.9. The jump rate λ is bounded and there exists $[\lambda]_1 \in \mathbb{R}^+$ such that for all $x, y \in E$ and $t \in [0, t^*(x) \wedge t^*(y)]$, one has

$$\left|\lambda(\Phi(x,t)) - \lambda(\Phi(y,t))\right| \leq [\lambda]_1 |x - y|.$$

Assumption 2.10. The deterministic exit time from E, denoted by t^*, is assumed to be bounded and Lipschitz continuous on E.

Remark 2.11. Since the deterministic exit time t^* is bounded by C_{t^*}, one may notice that $\boldsymbol{L}_c^u(E)$ for $u \geq C_{t^*}$ is no other than $\boldsymbol{L}_c(E)$.

Remark 2.12. In most practical applications, the physical properties of the system ensure that either t^* is bounded, or the problem has a natural finite deterministic time horizon t_f. In the latter case, there is no loss of generality in considering that t^* is bounded by this deterministic time horizon. This leads to replacing C_{t^*} by t_f. An example of such a situation is presented in an industrial example in Section 6.2.

Assumption 2.13. The Markov kernel Q is Lipschitz in the following sense: there exists $[Q] \in \mathbb{R}^+$ such that for all $u \geq 0$ and for all function $w \in \boldsymbol{L}_c^u(E)$, one has

(1) for all $x, y \in E$ and $t \in [0, t^*(x) \wedge t^*(y) \wedge u)$,

$$\big|Qw(\Phi(x,t)) - Qw(\Phi(y,t))\big| \leq [Q][w]_1^{E,u}|x-y|.$$

(2) for all $x, y \in E$ such that $t^*(x) \vee t^*(y) \leq u$,

$$\big|Qw(\Phi(x,t^*(x))) - Qw(\Phi(y,t^*(y)))\big| \leq [Q]\big([w]_*^{E,u} + [w]_1^{E,u}\big)|x-y|.$$

Remark 2.14. Assumption 2.13 is slightly more restrictive that its counterpart in [6] (Assumption 2.5), because of the introduction of the state space $\boldsymbol{L}_c^u(E)$. This is to ensure that the time-augmented process still satisfies a similar assumption; see Section 5.1.

3. Expectation

From now on, we will assume that $Z_0 = x$ a.s. for some $x \in E$. For all fixed $N \in \mathbb{N}^*$, we intend to numerically approximate the quantity

$$J_N(l,c)(x) = \boldsymbol{E}_x\left[\int_0^{T_N} l(X_t)\,dt + \sum_{j=1}^{N} c(X_{T_j^-})\mathbb{1}_{\{X_{T_j^-}\in\partial E\}}\right], \tag{2}$$

where $l \in B(E)$, $c \in B(\partial E)$ and X_{t-} is the left limit of X_t. Thus, $X_{T_j^-}$ is the j-th prejump location. Since the boundary jumps occur exactly at the deterministic exit times from E, one has,

$$J_N(l,c)(x) = \boldsymbol{E}_x\left[\int_0^{T_N} l(X_t)\,dt + \sum_{j=1}^{N} c\big(\Phi(Z_{j-1}, t^*(Z_{j-1}))\big)\mathbb{1}_{\{S_j = t^*(Z_{j-1})\}}\right].$$

In many applications, $J_N(l,c)(x)$ appears as a cost or a reward function. The first term, that depends on l, is called the running cost and the second one, that depends on c, is the boundary jump cost.

The rest of this section is devoted to formulating the expectation above in a way that will allow us to derive a numerical computation method. The Lipschitz continuity property will be a crucial point when it comes to proving the convergence of our approximation scheme. For this reason, the first step of our approximation is to replace the indicator function in $J_N(l,c)(x)$ by a Lipschitz continuous

function. Then we present a recursive method yielding the required expectation. This recursive formulation will be the basis of our numerical method.

3.1. ***Lipschitz continuity.*** We introduce a regularity assumption on l and c.

Assumption 3.1. We assume that $l \in \boldsymbol{L}_c(E)$ and $c \in \boldsymbol{L}_c(\partial E)$.

Moreover, we replace the indicator function in $J_N(l,c)(x)$ by a Lipschitz continuous function δ^A, with $A > 0$. Let then

$$J_N^A(l,c)(x) = \boldsymbol{E}_x\left[\int_0^{T_N} l(X_t)\,dt + \sum_{j=1}^{N} c\big(\Phi(Z_{j-1}, t^*(Z_{j-1}))\big)\delta^A(Z_{j-1}, S_j)\right],$$

where δ^A is a triangular approximation of the indicator function. It is defined on $E \times \mathbb{R}$ by

$$\delta^A(x,t) = \begin{cases} A\Big(t - \Big(t^*(x) - \dfrac{1}{A}\Big)\Big) & \text{for } t \in \Big[t^*(x) - \dfrac{1}{A};\ t^*(x)\Big], \\ -A\Big(t - \Big(t^*(x) + \dfrac{1}{A}\Big)\Big) & \text{for } t \in \Big[t^*(x);\ t^*(x) + \dfrac{1}{A}\Big], \\ 0 & \text{otherwise.} \end{cases}$$

For all $x \in E$, the function $\delta^A(x,t)$ goes to $\mathbb{1}_{\{t=t^*(x)\}}$ when A goes to infinity. The following proposition proves the convergence of $J_N^A(l,c)(x)$ towards $J_N(l,c)(x)$ with an error bound.

Proposition 3.2. *For all $x \in E$, $A > 0$, $N \in \mathbb{N}^*$, $l \in \boldsymbol{L}_c(E)$ and $c \in \boldsymbol{L}_c(\partial E)$, one has*

$$\big|J_N^A(l,c)(x) - J_N(l,c)(x)\big| \le \frac{NC_cC_\lambda}{A}.$$

Proof. For all $x \in E$, one has

$$\begin{aligned} &\big|J_N^A(l,c)(x) - J_N(l,c)(x)\big| \\ &\quad = \left|\boldsymbol{E}_x\left[\sum_{j=1}^{N} c\big(\Phi(Z_{j-1}, t^*(Z_{j-1}))\big)\big(\delta^A(Z_{j-1}, S_j) - \mathbb{1}_{\{S_j = t^*(Z_{j-1})\}}\big)\right]\right| \\ &\quad \le C_c \sum_{j=1}^{N} \boldsymbol{E}_x\big[|\delta^A(Z_{j-1}, S_j) - \mathbb{1}_{\{S_j = t^*(Z_{j-1})\}}|\big] \\ &\quad \le C_c \sum_{j=1}^{N} \boldsymbol{E}_x\big[\boldsymbol{E}\big[|\delta^A(Z_{j-1}, S_j) - \mathbb{1}_{\{S_j = t^*(Z_{j-1})\}}| \;\big|\; Z_{j-1}\big]\big]. \end{aligned}$$

We recall that the conditional law of S_j with respect to Z_{j-1} has density

$$s \to \lambda\big(\Phi(Z_{j-1}, s)\big)e^{-\Lambda(Z_{j-1}, s)}$$

on $[0; t^*(Z_{j-1}))$ and puts the weight $e^{-\Lambda(Z_{j-1},t^*(Z_{j-1}))}$ on the point $t^*(Z_{j-1})$. We also recall that λ is bounded thanks to Assumption 2.9. Finally, one has

$$\begin{aligned}
&\left|J_N^A(l,c)(x) - J_N(l,c)(x)\right| \\
&\qquad \le C_c \sum_{j=1}^{N} \boldsymbol{E}_x\left[\int_{t^*(Z_{j-1})-\frac{1}{A}}^{t^*(Z_{j-1})} \delta^A(Z_{j-1},s)\lambda\big(\Phi(Z_{j-1},s)\big)e^{-\Lambda(Z_{j-1},s)}\,ds\right] \\
&\qquad \le \frac{NC_cC_\lambda}{A}.
\end{aligned}$$

Hence the result. □

Consequently to this proposition, we consider, from now on, the approximation of $J_N^A(l,c)(x)$ for some fixed A, large enough to ensure that the previous error is as small as required. The suitable choice of A will be discussed in Section 4.2.

3.2. ***Recursive formulation.*** Davis shows in [5, Section 32] that the expectation $J_N^A(l,c)(x)$ we are interested in is obtained by merely iterating an operator that we will denote by G. The rest of this section is dedicated to presenting this method from which we will derive our approximation scheme in Section 4.

Definition 3.3. Introduce the functions L, C and F defined for all $x \in E$ and $t \in [0; t^*(x)]$ by

$$\begin{aligned}
L(x,t) &= \int_0^t l\big(\Phi(x,s)\big)\,ds, \\
C(x,t) &= c\big(\Phi(x,t^*(x))\big)\delta^A(x,t), \\
F(x,t) &= L(x,t) + C(x,t),
\end{aligned}$$

along with the operator $G\colon B(E) \to B(E)$ given by

$$Gw(x) = \boldsymbol{E}_x\big[F(x,S_1) + w(Z_1)\big].$$

Definition 3.4. Define the sequence of functions $(v_k)_{0\le k\le N}$ in $B(E)$ by

$$v_N(x) = 0, \quad v_k(x) = Gv_{k+1}(x).$$

Davis then shows in [5, Equation 32.33] that, for all $k \in \{0,\dots,N\}$,

$$v_{N-k}(x) = \boldsymbol{E}_x\left[\int_0^{T_k} l(X_t)\,dt + \sum_{j=1}^{k} c\big(\Phi(Z_{j-1},t^*(Z_{j-1}))\big)\delta^A(Z_{j-1},S_j)\right].$$

Thus, the quantity $J_N^A(l,c)(x)$ we intend to approximate is none other than $v_0(x)$.

Notice that, thanks to the Markov property of the chain $(Z_n, S_n)_{n\in\mathbb{N}}$, one has for all $k \in \{0, \dots, N-1\}$,

$$Gw(x) = \boldsymbol{E}\big[F(Z_k, S_{k+1}) + w(Z_{k+1})\big|Z_k = x\big]. \tag{3}$$

Hence, for all $k \in \{0, \dots, N\}$, let $V_k = v_k(Z_k)$ so that one has

$$V_N = 0, \quad V_k = \boldsymbol{E}\big[F(Z_k, S_{k+1}) + V_{k+1}\big|Z_k\big].$$

This backward recursion provides the required quantity

$$V_0 = J_N^A(l, c)(x).$$

Hence, we need to approximate the sequence of random variables $(V_k)_{0\le k\le N}$. This sequence satisfies a recursion that only depends on the chain $(Z_k, S_k)_{0\le k\le N}$. Therefore, it appears natural to propose an approximation scheme based on a discretization of this chain $(Z_k, S_k)_{0\le k\le N}$, called quantization, similarly to the ideas developed in [6] and [3].

4. Approximation scheme

Let us now turn to the approximation scheme itself. We explained in the previous section how the expectation we are interested in stems from the iteration of the operator G that only depends on the discrete-time Markov chain $(Z_k, S_k)_{0\le k\le N}$. The first step of our numerical method is therefore to discretize this chain in order to approximate the operator G.

4.1. ***Quantization of the chain*** $(Z_k, S_k)_{k\le N}$**.** Our approximation method is based on the quantization of the underlying discrete time Markov chain $(\Theta_k)_{k\le N} = (Z_k, S_k)_{k\le N}$. This quantization consists in finding an optimally designed discretization of the process to provide for each step k the best possible approximation of Θ_k by a random variable $\widehat{\Theta}_k$ which state space has a finite and fixed number of points. Here, *optimal* means that the distance between Θ_k and $\widehat{\Theta}_k$ in a suitably chosen L^p norm is minimal. For details on the quantization methods, we mainly refer to [9] but the interested reader can also consult [1], [2] and the references therein.

More precisely, consider X an $\mathbb{R}^q$-valued random variable such that $\|X\|_p < \infty$ and let K be a fixed integer. The optimal L_p-quantization of the random variable X consists in finding the best possible L_p-approximation of X by a random vector $\widehat{X} \in \{x^1, \dots, x^K\}$ taking at most K values: This procedure consists of two steps:

(1) Find a finite weighted grid $\Gamma \subset \mathbb{R}^q$ with $\Gamma = \{x^1, \dots, x^K\}$.

(2) Set $\widehat{X} = \widehat{X}^\Gamma$ where $\widehat{X}^\Gamma = \mathrm{proj}_\Gamma(X)$ with proj_Γ denotes the closest neighbor projection on Γ.

The asymptotic properties of the L_p-quantization are given by the following result; see [9], for example.

Theorem 4.1. *If* $\mathbb{E}[|X|^{p+\eta}] < +\infty$ *for some* $\eta > 0$ *then one has*

$$\lim_{K\to\infty} K^{p/q} \min_{|\Gamma|\leq K} \|X - \widehat{X}^{\Gamma}\|_p^p = J_{p,q} \int |h|^{q/(q+p)}(u)\, du,$$

where the law of X *is* $P_X(du) = h(u)\lambda_q(du) + \nu$ *with* $\nu \perp \lambda_q$, $J_{p,q}$ *a constant and* λ_q *the Lebesgue measure in* $\mathbb{R}^q$.

Remark that X needs to have finite moments up to the order $p + \eta$ to ensure the above convergence. In this work, we used the CLVQ quantization algorithm described in [1], Section 3.

There exists a similar procedure for the optimal quantization of a Markov chain $\{X_k\}_{k\in\mathbb{N}}$. There are two approaches to provide the quantized approximation of a Markov chain. The first one, based on the quantization at each time k of the random variable X_k is called the *marginal quantization*. The second one that enhances the preservation of the Markov property is called *Markovian quantization*. Remark that for the latter, the quantized Markov process is not homogeneous. These two methods are described in details in [9, Section 3]. In this work, we used the marginal quantization approach for simplicity reasons.

The quantization algorithm provides for each time step $0 \leq k \leq N$ a finite grid Γ_k of $E \times \mathbb{R}^+$ as well as the transition matrices $(\widehat{Q}_k)_{0\leq k\leq N-1}$ from Γ_k to Γ_{k+1}. Let $p \geq 1$ such that for all $k \leq N$, Z_k and S_k have finite moments at least up to order p and let proj_{Γ_k} be the closest-neighbor projection from $E \times \mathbb{R}^+$ onto Γ_k (for the distance associated to norm p). The quantized process $(\widehat{\Theta}_k)_{k\leq N} = (\widehat{Z}_k, \widehat{S}_k)_{k\leq N}$ takes values for each k in the finite grid Γ_k of $E \times \mathbb{R}^+$ and is defined by

$$(\widehat{Z}_k, \widehat{S}_k) = \text{proj}_{\Gamma_k}(Z_k, S_k). \tag{4}$$

Moreover, we also denote by Γ_k^Z and Γ_k^S, respectively, the projections of Γ_k on E and $\mathbb{R}^+$.

Some important remarks must be made concerning the quantization. On the one hand, the optimal quantization has nice convergence properties stated by Theorem 4.1. Indeed, the L^p-quantization error $\|\Theta_k - \widehat{\Theta}_k\|_p$ goes to zero when the number of points in the grids goes to infinity. However, on the other hand, the Markov property is not maintained by the algorithm and the quantized process is generally not Markovian. Although the quantized process can be easily transformed into a Markov chain (see [9]), this chain will not be homogeneous. It must be pointed out that the quantized process $(\widehat{\Theta}_k)_{k\in\mathbb{N}}$ depends on the starting point Θ_0 of the process.

In practice, we begin with the computation of the quantization grids which merely requires to be able to simulate the process. This step is quite time-consuming, especially when the number of points in the quantization grids is large. However, the grids are only computed once and for all and may be stored off-line. What is more, they only depend on the dynamics of the process, not on the cost functions l and c. Hence, the same grids may be used to compute different expectations of functionals as long as they are related to the same process. Our schemes are then based on the following simple idea: we replace the process by its quantized approximation within the operator G. The approximation is thus obtained in a very simple way since the quantized process has finite state space.

4.2. ***Approximation of the expectation and rate of convergence.*** We now use the quantization of the process $(\Theta_k)_{k\leq N} = (Z_k, S_k)_{k\leq N}$. In order to approximate the random variables $(V_k)_{k\leq N}$, we introduce a quantized version of the operator G. Notice that the quantized process is no longer an homogeneous Markov chain so that we have different operators for each time step k. Their definitions naturally stem from a remark made in the previous section: recall that for all $k \in \{1, \ldots, N\}$ and $x \in E$,

$$Gw(x) = \boldsymbol{E}\big[F(Z_{k-1}, S_k) + w(Z_k) \,\big|\, Z_{k-1} = x\big].$$

Definition 4.2. For all $k \in \{1, \ldots, N\}$, $w \in B(\Gamma_k^Z)$ and $z \in \Gamma_{k-1}^Z$, let

$$\widehat{G}_k w(z) = \boldsymbol{E}\big[F(z, \widehat{S}_k) + w(\widehat{Z}_k) \,\big|\, \widehat{Z}_{k-1} = z\big].$$

Introduce also the functions $(\hat{v}_k)_{0\leq k\leq N}$ by

$$\begin{aligned} \hat{v}_N(z) &= 0 && \text{for all } z \in \Gamma_N^Z, \\ \hat{v}_k(z) &= \widehat{G}_{k+1}\hat{v}_{k+1}(z) && \text{for all } k \in \{0, \ldots, N-1\} \text{ and } z \in \Gamma_k^Z. \end{aligned}$$

Finally, for all $k \in \{0, \ldots, N\}$, let

$$\widehat{V}_k = \hat{v}_k(\widehat{Z}_k).$$

Remark 4.3. The conditional expectation in $\widehat{G}_k w(z)$ is a finite sum. Thus, the numerical computation of the sequence $(\widehat{V}_k)_k$ will be easily performed as soon as the quantized process $(\widehat{\Theta}_k)_{k\leq N}$ has been obtained.

Remark 4.4. We have assumed that $Z_0 = x$ a.s. Thus, the quantization algorithm provides that $\widehat{Z}_0 = x$ a.s. too. Consequently, the random variable $\widehat{V}_0 = \hat{v}_0(\widehat{Z}_0)$ is, in fact, deterministic.

The following theorem states the convergence of $\widehat{V}_0$ towards $V_0 = J_N^A(l, c)(x)$ and provides a bound for the rate of convergence.

Theorem 4.5. *For all $k \in \{0, \dots, N\}$, one has $v_k \in \boldsymbol{L}_c(E)$. Moreover, the approximation error satisfies*

$$|J_N(l,c)(x) - \widehat{V}_0| \leq \varepsilon_N(l,c,X,A),$$

where

$$\varepsilon_N(l,c,X,A) = \sum_{k=0}^{N-1} \Big(2[v_{k+1}]\|Z_{k+1} - \widehat{Z}_{k+1}\|_p + \big(2[v_k] + [F]_1\big)\|Z_k - \widehat{Z}_k\|_p + [F]_2\|S_{k+1} - \widehat{S}_{k+1}\|_p\Big) + \frac{NC_cC_\lambda}{A}$$

with

$$\begin{aligned}
[F]_1 &= C_{t^*}[l]_1 + [c]_* + A[t]_*C_c,\\
[F]_2 &= C_l + AC_c.\\
C_{v_n} &\leq n\big(C_{t^*}C_l + C_c\big),\\
[v_n]_1 &\leq e^{C_{t^*}C_\lambda}\big(K(A, v_{n-1}) + nC_{t^*}[\lambda]_1(C_{t^*}C_l + C_c)\big) + C_{t^*}[l]_1,\\
[v_n]_2 &\leq e^{C_{t^*}C_\lambda}\big(C_{t^*}C_lC_\lambda + 2C_l + C_\lambda C_c + (2n-1)C_\lambda(C_{t^*}C_l + C_c)\big) + C_l,\\
[v_n]_* &\leq [v_n]_1 + [t^*][v_n]_2,\\
[v_n] &\leq K(A, v_{n-1}),
\end{aligned}$$

and for all $w \in \boldsymbol{L}_c(E)$, $K(A,w) = E_1 + E_2A + E_3[w]_1 + E_4C_w + [Q][w]_*$, where

$$\begin{aligned}
E_1 &= 2[l]_1C_{t^*} + C_l\big([t^*] + 2C_{t^*}^2[\lambda]_1\big) + [c]_*(1 + C_{t^*}C_\lambda)\\
&\quad + C_c\big(2[\lambda]_1C_{t^*} + C_\lambda C_{t^*}^2[\lambda]_1 + 2[t^*]C_\lambda\big),\\
E_2 &= C_cC_{t^*}C_\lambda[t^*],\\
E_3 &= (1 + C_{t^*}C_\lambda)[Q],\\
E_4 &= 2C_\lambda[t^*] + C_{t^*}[\lambda]_1(2 + C_{t^*}C_\lambda).
\end{aligned}$$

The choice of A. Proposition 3.2 suggests that A should be as large as possible. However, the constants $[F]_1$, $[F]_2$ and $[v_n]$ that appear in the bound of the approximation error proposed by the above Theorem 4.5 grow linearly with A. Thus, in order to control this error, it is necessary that the order of magnitude of the quantization error $\|\Theta_k - \widehat{\Theta}_k\|_p$ be at most $1/A$.

The convergence of the approximation scheme can be derived from Theorem 4.5. Indeed, on the one hand, one must remind that $V_0 = J_N^A(l,c)(x)$ is the expectation we intended to approximate and on the other hand, $\|\Theta_k - \widehat{\Theta}_k\|_p$ may become arbitrarily small when the number of points in the quantization grids goes to infinity (see [9], for example). An outline of the proof is presented in Appendix C.

5. Time-dependent functionals

We now turn to the main contribution of this paper and present two generalizations of the previous problem. On the one hand, we will consider time-dependent functionals of the form

$$\boldsymbol{E}_x\left[\int_0^{T_N} l(X_t,t)\,dt + \sum_{j=1}^{N} c(X_{T_j^-},T_j)\mathbb{1}_{\{X_{T_j^-}\in\partial E\}}\right]$$

where l and c are Lipschitz continuous functions. On the other hand, we wish to replace the random time horizon T_N by a deterministic one, denoted by t_f:

$$\boldsymbol{E}_x\left[\int_0^{t_f} l(X_t,t)\,dt + \sum_{T_j\le t_f} c(X_{T_j^-},T_j)\mathbb{1}_{\{X_{T_j^-}\in\partial E\}}\right].$$

We will reason as follows. As suggested by Davis in [5], we will introduce a transformation $(\widetilde{X}_t)_{t\ge 0}$ of the initial process $(X_t)_{t\ge 0}$ by including the time variable into the state space: $(\widetilde{X}_t) = (X_t, t)$. Indeed, we will see that both the expectation of the time-dependent functional and the one with deterministic time horizon are no other than expectations of time invariant functionals for the time-augmented process $(\widetilde{X}_t)_{t\ge 0}$. We therefore intend to apply the previously exposed approximation scheme to this new PDMP. However, it is far from obvious that the Lipschitz continuity assumptions 2.9, 2.13 and 2.10 still hold for this new process.

Thus, the rest of this section is organized as follows. First, we recall the precise definition of the time-augmented process and prove that it satisfies the Lipschitz continuity assumptions required by our approximation scheme. Then, we will see that the time-dependent functional case corresponds to a time invariant functional for the new transformed process and may therefore be obtained thanks to the earlier method. Finally, we consider the deterministic time horizon problem that features an additional hurdle namely the presence of non-Lipschitz continuous indicator functions.

5.1. ***The time-augmented process.*** Davis suggests, in [5, Section 31], that the case of the time-dependent functionals may be treated by introducing the time variable within the state space. Thus, it will be possible to apply our previous numerical method to the time-augmented process. However, and this is what we discuss in this section, it is necessary to check whether the Lipschitz continuity assumptions still hold. We first recall the definition of the time-augmented process given by Davis.

Definition 5.1. Introduce the new state space

$$\widetilde{E} = E \times \mathbb{R}^+$$

equipped with the norm defined by: for all $\xi = (x,t)$, $\xi' = (x',t') \in \widetilde{E}$, let

$$|\xi - \xi'| = |x - x'| + |t - t'| \tag{5}$$

where the norm on E is given by (1). On this state space, we define the process

$$\widetilde{X}_t = (X_t, t).$$

The local characteristics of the PDMP $(\widetilde{X}_t)_{t\geq 0}$, denoted by $(\tilde{\lambda}, \widetilde{Q}, \widetilde{\Phi})$, are given for all $\xi = (x,t) \in \widetilde{E}$ by

$$\begin{cases} \tilde{\lambda}(\xi) = \lambda(x), \\ \widetilde{\Phi}(\xi, s) = \big(\Phi(x,s), t+s\big) & \text{for } s \leq t^*(x), \\ \widetilde{Q}\big(\xi, A \times \{t\}\big) = Q(x,A) & \text{for all } A \in \mathcal{B}(E). \end{cases}$$

Moreover, we naturally define for all $\xi = (x,t) \in \widetilde{E}$

$$\tilde{t}^*(\xi) = \inf\{s > 0 \text{ such that } \widetilde{\Phi}(\xi, s) \in \partial \widetilde{E}\} = t^*(x)$$

Clearly, Assumptions 2.9 and 2.10 still hold with $[\tilde{\lambda}]_1 = [\lambda]_1$ and $[\tilde{t^*}] = [t^*]$. However, proving Assumption 2.13 is more intricate. We start with the following lemma.

Lemma 5.2. *Let $u, t \geq 0$ and $w \in \boldsymbol{L}_c^u(\widetilde{E})$. Denote by w_t the function of $B(E)$ defined by $w_t = w(\cdot, t)$. One has then $w_t \in \boldsymbol{L}_c^{t\wedge u}(E)$ with*

$$\begin{aligned} [w_t]_1^{E,t\wedge u} &\leq [w]_1^{\widetilde{E},u}, \\ [w_t]_2^{E,t\wedge u} &\leq [w]_1^{\widetilde{E},u} + [w]_2^{\widetilde{E},u}, \\ [w_t]_*^{E,t\wedge u} &\leq (1 + [t^*])[w]_*^{\widetilde{E},u}. \end{aligned}$$

Proof. Let $u, t \geq 0$ and $w \in \boldsymbol{L}_c^u(\widetilde{E})$. For $x, x' \in E$ and $s \leq t^*(x) \wedge t^*(x') \wedge t \wedge u$, one has

$$\big|w_t(\Phi(x,s)) - w_t(\Phi(x',s))\big| = \big|w\big(\widetilde{\Phi}((x,t-s),s)\big) - w\big(\widetilde{\Phi}((x',t-s),s)\big)\big|.$$

We now use the fact that $w \in \boldsymbol{L}_c^u(\widetilde{E})$ which yields since $s \leq u$

$$\big|w_t(\Phi(x,s)) - w_t(\Phi(x',s))\big| \leq [w]_1^{\widetilde{E},u}\big|(x,t-s) - (x',t-s)\big| = [w]_1^{\widetilde{E},u}\big|x - x'\big|.$$

Hence, $[w_t]_1^{E,t\wedge u} \leq [w]_1^{\widetilde{E},u}$, and similarly one obtains $[w_t]_2^{E,t\wedge u} \leq [w]_1^{\widetilde{E},u} + [w]_2^{\widetilde{E},u}$.

On the other hand, for $x, x' \in E$ such that $t^*(x) \vee t^*(x') \leq t \wedge u$, one has

$$\begin{aligned}
&\left|w_t(\Phi(x,t^*(x))) - w_t(\Phi(x',t^*(x')))\right| \\
&\quad = \left|w\big(\widetilde{\Phi}((x,t-t^*(x)),t^*(x))\big) - w\big(\widetilde{\Phi}((x',t-t^*(x')),t^*(x'))\big)\right| \\
&\quad = \Big|w\big(\widetilde{\Phi}\big((x,t-t^*(x)),\tilde{t}^*(x,t-t^*(x))\big)\big) \\
&\qquad\qquad - w\big(\widetilde{\Phi}\big((x',t-t^*(x')),\tilde{t}^*(x',t-t^*(x'))\big)\big)\Big|;
\end{aligned}$$

moreover since $w \in \boldsymbol{L}_c^u(\widetilde{E})$ and $\tilde{t}^*(x,t-t^*(x)) \vee \tilde{t}^*(x',t-t^*(x')) \leq u$ one has

$$\left|w_t(\Phi(x,t^*(x))) - w_t(\Phi(x',t^*(x')))\right| \leq [w]_*^{\widetilde{E},u}\left|(x,t-t^*(x)) - (x',t-t^*(x'))\right|.$$

We conclude thanks to the Lipschitz continuity assumption 2.10 on t^*, which yields $\left|(x,t-t^*(x)) - (x',t-t^*(x'))\right| \leq (1+[t^*])|x-x'|$. One obtains $[w_t]_*^{E,t\wedge u} \leq [w]_*^{\widetilde{E},u}\big(1+[t^*]\big)$ and $w_t \in \boldsymbol{L}_c^{t\wedge u}(E)$. □

The next proposition proves that Assumption 2.13 holds for the time-augmented process $(\widetilde{X})_{t\geq 0}$.

Proposition 5.3. *Let $w \in \boldsymbol{L}_c^u(\widetilde{E})$.*

(1) *For all $\xi, \xi' \in \widetilde{E}$ and $s \in [0, \tilde{t}^*(\xi) \wedge \tilde{t}^*(\xi') \wedge u]$,*

$$\left|\widetilde{Q}w\big(\widetilde{\Phi}(\xi,s)\big) - \widetilde{Q}w\big(\widetilde{\Phi}(\xi',s)\big)\right| \leq ([Q]\vee 1)[w]_1^{\widetilde{E},u}|\xi-\xi'|.$$

(2) *For all $\xi, \xi' \in \widetilde{E}$ such that $\tilde{t}^*(\xi) \vee \tilde{t}^*(\xi') \leq u$,*

$$\begin{aligned}
&\left|\widetilde{Q}w\big(\widetilde{\Phi}(\xi,\tilde{t}^*(\xi))\big) - \widetilde{Q}w\big(\widetilde{\Phi}(\xi',\tilde{t}^*(\xi'))\big)\right| \\
&\qquad\qquad \leq ([Q]\vee 1)(1+[t^*])\big([w]_*^{\widetilde{E},u} + [w]_1^{\widetilde{E},u}\big)|\xi-\xi'|.
\end{aligned}$$

In other words, Assumption 2.13 is satisfied with $[\widetilde{Q}] = ([Q]\vee 1)(1+[t^*])$.

Proof. As in the previous lemma, for all $t \geq 0$, we will denote by w_t the function of $B(E)$ defined by $w_t = w(\cdot,t)$. For $\xi = (x,t) \in \widetilde{E}$ and $w \in \boldsymbol{L}_c^u(\widetilde{E})$, one has, by the definition of $\widetilde{Q}$,

$$\widetilde{Q}w(\xi) = \int_{\xi'\in\widetilde{E}} w(\xi')\widetilde{Q}\big((x,t),d\xi'\big) = \int_{z\in E} w(z,t)Q\big(x,dz\big) = Qw_t(x). \quad (6)$$

We may now check the regularity assumption on $\widetilde{Q}$. Let $\xi = (x,t)$ and $\xi' = (x',t') \in \widetilde{E}$. Let $s \in [0; \tilde{t}^*(\xi) \wedge \tilde{t}^*(\xi') \wedge u]$. Thanks to the definition of $\widetilde{\Phi}$ and (6) one has

$$\begin{aligned}
\left|\widetilde{Q}w(\widetilde{\Phi}(\xi,s)) - \widetilde{Q}w(\widetilde{\Phi}(\xi',s))\right| &= \left|\widetilde{Q}w(\Phi(x,s),t+s) - \widetilde{Q}w(\Phi(x',s),t'+s)\right| \\
&= \left|Qw_{t+s}(\Phi(x,s)) - Qw_{t'+s}(\Phi(x',s))\right|.
\end{aligned}$$

We split this into the sum of two differences:

$$\begin{aligned}&\big|Qw_{t+s}(\Phi(x,s))-Qw_{t'+s}(\Phi(x',s))\big|\\&\qquad\le\big|Qw_{t+s}(\Phi(x,s))-Qw_{t+s}(\Phi(x',s))\big|+\big|Q(w_{t+s}-w_{t'+s})(\Phi(x',s))\big|.\end{aligned}$$

On the one hand, we recall that thanks to Lemma 5.2, $w_{t+s}\in \boldsymbol{L}_c^{(t+s)\wedge u}(E)$, so that, since $s\le(t+s)\wedge u$, we may use the Lipschitz continuity assumption 2.13 on Q and the first term is bounded as follows:

$$\big|Qw_{t+s}(\Phi(x,s))-Qw_{t+s}(\Phi(x',s))\big|\le[Q][w_{t+s}]_1^{E,(t+s)\wedge u}|x-x'|.$$

Lemma 5.2 also provides $[w_{t+s}]_1^{E,(t+s)\wedge u}\le[w]_1^{\widetilde{E},u}$. On the other hand, and more basically, the second term in the equation above satisfies

$$\big|Q(w_{t+s}-w_{t'+s})(\Phi(x',s))\big|\le[w]_1^{\widetilde{E},u}|t-t'|.$$

We obtain

$$\big|\widetilde{Q}w\big(\widetilde{\Phi}(\xi,s)\big)-\widetilde{Q}w\big(\widetilde{\Phi}(\xi',s)\big)\big|\le([Q]\vee 1)[w]_1^{\widetilde{E},u}|\xi-\xi'|.$$

We reason similarly to bound $\big|\widetilde{Q}w\big(\widetilde{\Phi}(\xi,\tilde{t}^*(\xi))\big)-\widetilde{Q}w\big(\widetilde{\Phi}(\xi',\tilde{t}^*(\xi'))\big)\big|$, where $\xi=(x,t)$ and $\xi'=(x',t')\in\widetilde{E}$ are such that $\tilde{t}^*(\xi)\vee\tilde{t}^*(\xi')\le u$. Equation (6) yields

$$\begin{aligned}&\big|\widetilde{Q}w\big(\widetilde{\Phi}(\xi,\tilde{t}^*(\xi))\big)-\widetilde{Q}w\big(\widetilde{\Phi}(\xi',\tilde{t}^*(\xi'))\big)\big|\\&\qquad=\big|Qw_{t+t^*(x)}\big(\Phi(x,t^*(x))\big)-Qw_{t'+t^*(x')}\big(\Phi(x',t^*(x'))\big)\big|,\end{aligned}$$

which we now split as follows:

$$\begin{aligned}&\big|Qw_{t+t^*(x)}\big(\Phi(x,t^*(x))\big)-Qw_{t'+t^*(x')}\big(\Phi(x',t^*(x'))\big)\big|\\&\qquad\le\big|Qw_{t+t^*(x)}\big(\Phi(x,t^*(x))\big)-Qw_{t+t^*(x)}\big(\Phi(x',t^*(x'))\big)\big|\\&\qquad\qquad+\big|(Qw_{t+t^*(x)}-Qw_{t'+t^*(x')})\big(\Phi(x',t^*(x'))\big)\big|.\end{aligned}$$

Thanks to Lemma 5.2, $w_{t+t^*(x)}\in\boldsymbol{L}_c^{(t+t^*(x))\wedge u}(E)$. We assume, without loss of generality, that $t^*(x)\ge t^*(x')$, so $t^*(x)\vee t^*(x')\le(t+t^*(x))\wedge u$. Therefore, the first term in the above equation is bounded, thanks to the Lipschitz continuity assumption 2.13 on Q and Lemma 5.2, by

$$[Q]\big((1+[t^*])[w]_*^{\widetilde{E},u}+[w]_1^{\widetilde{E},u}\big)|x-x'|.$$

It is more straightfoward to obtain a bound for the second term, of the form

$$[w]_1^{\widetilde{E},u}\big|t-t'+t^*(x)-t^*(x')\big|\le[w]_1^{\widetilde{E},u}\big(|t-t'|+[t^*]|x-x'|\big).$$

We obtain

$$\begin{aligned}
&\left|\widetilde{Q}w\big(\widetilde{\Phi}(\xi,\tilde{t}^*(\xi))\big)-\widetilde{Q}w\big(\widetilde{\Phi}(\xi',\tilde{t}^*(\xi'))\big)\right| \\
&\quad\le [Q](1+[t^*])[w]_*^{\widetilde{E},u}|x-x'|+[w]_1^{\widetilde{E},u}\big([Q]|x-x'|+|t-t'|+[t^*]|x-x'|\big) \\
&\quad\le \big([Q]\vee 1\big)\big(1+[t^*]\big)\big([w]_*^{\widetilde{E},u}+[w]_1^{\widetilde{E},u}\big)|\xi-\xi'|.
\end{aligned}$$

Hence the result. □

Consequently, we may apply our numerical method to the time-augmented process $(\widetilde{X}_t)_{t\ge 0}$. In other words, for $l\in \boldsymbol{L}_c(\widetilde{E})$, $c\in \boldsymbol{L}_c(\partial\widetilde{E})$ and $\xi\in\widetilde{E}$, our approximation scheme may be used to compute

$$\widetilde{J}_N(l,c)(\xi)=\boldsymbol{E}_\xi\left[\int_0^{T_N} l(\widetilde{X}_t)\,dt+\sum_{j=1}^{N} c(\widetilde{X}_{T_j^-})\mathbb{1}_{\{\widetilde{X}_{T_j^-}\in\partial\widetilde{E}\}}\right]. \tag{7}$$

We will now see that the time-dependent functional and the deterministic time horizon problems boil down to computing such quantities $\widetilde{J}_N(l,c)(\xi)$ for suitably chosen functions l and c.

5.2. ***Lipschitz continuous cost functions.*** We first consider the time-dependent functional problem with Lipschitz continuous cost functions. Thus, let then $l\in \boldsymbol{L}_c(\widetilde{E})$, $c\in \boldsymbol{L}_c(\partial\widetilde{E})$ and $x\in E$, we wish to compute

$$\boldsymbol{E}_x\left[\int_0^{T_N} l(X_t,t)\,dt+\sum_{j=1}^{N} c(X_{T_j^-},T_j)\mathbb{1}_{\{X_{T_j^-}\in\partial E\}}\right].$$

It is straightforward to show that this quantity may be expressed using the time-augmented process starting from the point $\xi_0=(x,0)$. Indeed, one has

$$\widetilde{J}_N(l,c)(\xi_0)=\boldsymbol{E}_x\left[\int_0^{T_N} l(X_t,t)\,dt+\sum_{j=1}^{N} c(X_{T_j^-},T_j)\mathbb{1}_{\{X_{T_j^-}\in\partial E\}}\right],$$

where $\widetilde{J}_N(l,c)(\xi_0)$ is given by (7). Although they are time-dependent, the cost functions l and c are seen, in the left-hand side term, as time invariant functions of the time-augmented process. The expectation of the time-dependent functional is therefore obtained by computing the expectation of a time invariant functional for the transformed PDMP thanks to the approximation scheme described in Section 4. This is what expresses the following theorem, which proof stems from the previous discussion.

Theorem 5.4. *Let $l\in \boldsymbol{L}_c(\widetilde{E})$ and $c\in \boldsymbol{L}_c(\partial\widetilde{E})$ and apply the approximation scheme described in Section 4 to the time-augmented process $(\widetilde{X}_t)_{t\ge 0}$, one has then*

$$\left|\boldsymbol{E}_x\left[\int_0^{T_N} l(X_t,t)\,dt+\sum_{j=1}^{N} c(X_{T_j^-},T_j)\mathbb{1}_{\{X_{T_j^-}\in\partial E\}}\right]-\hat{V}_0\right|\le \varepsilon_N(l,c,\widetilde{X},A),$$

where we denote by $\varepsilon_N(l,c,\widetilde{X},A)$ *the bound of the approximation error provided by Theorem 4.5 when our approximation scheme is applied with cost functions* l *and* c *to the time-augmented process* $(\widetilde{X}_t)_{t\ge 0}$.

Remark 5.5. The quantity $\varepsilon_N(l,c,\widetilde{X},A)$ is computed with respect to the process $(\widetilde{X}_t)_{t\ge 0}$ instead of $(X_t)_{t\ge 0}$, as presented in Theorem 4.5, so that

$$\begin{aligned}\varepsilon_N(l,c,\widetilde{X},A)=\sum_{k=0}^{N-1}\Big(&2[v_{k+1}]^{\widetilde{E}}\|\widetilde{Z}_{k+1}-\widehat{\widetilde{Z}}_{k+1}\|_p\\&+\big(2[v_k]^{\widetilde{E}}+[F]_1'+[F]_1''A\big)\|\widetilde{Z}_k-\widehat{\widetilde{Z}}_k\|_p\\&+\big([F]_2'+A[F]_2''\big)\|\widetilde{S}_{k+1}-\widehat{\widetilde{S}}_{k+1}\|_p\Big)+\frac{NC_cC_\lambda}{A}.\end{aligned}$$

where $(\widetilde{Z}_k,\widetilde{S}_k)_{k\in\mathbb{N}}$ denotes the sequence of the postjump locations and the interjump times of the time-augmented process $(\widetilde{X}_t)_{t\ge 0}$, and where

$$\begin{aligned}[F]_1'&=C_{t^*}[l]_1^{\widetilde{E}}+[c]_*^{\widetilde{E}},\\ [F]_1''&=[t^*]C_c,\\ [F]_2'&=C_l,\\ [F]_2''&=C_c,\\ C_{v_n}&\le n\big(C_{t^*}C_l+C_c\big),\\ [v_n]_1^{\widetilde{E}}&\le e^{C_{t^*}C_\lambda}\big(\widetilde{K}(A,v_{n-1})+nC_{t^*}[\lambda]_1(C_{t^*}C_l+C_c)\big)+C_{t^*}[l]_1^{\widetilde{E}},\\ [v_n]_2^{\widetilde{E}}&\le e^{C_{t^*}C_\lambda}\big(C_{t^*}C_lC_\lambda+2C_l+C_\lambda C_c+(2n-1)C_\lambda(C_{t^*}C_l+C_c)\big)+C_l,\\ [v_n]_*^{\widetilde{E}}&\le [v_n]_1^{\widetilde{E}}+[t^*][v_n]_2^{\widetilde{E}},\\ [v_n]^{\widetilde{E}}&\le \widetilde{K}(A,v_{n-1}),\end{aligned}$$

and for all $w\in\boldsymbol{L}_c(E)$ we have

$$\widetilde{K}(A,w)=\widetilde{E}_1+E_2A+\widetilde{E}_3[w]_1^{\widetilde{E}}+E_4C_w+[\widetilde{Q}][w]_*^{\widetilde{E}},$$

where

$$\begin{aligned}[\widetilde{Q}]&=([Q]\vee 1)(1+[t^*]),\\ \widetilde{E}_1&=2[l]_1^{\widetilde{E}}C_{t^*}+C_l\big([t^*]+2C_{t^*}^2[\lambda]_1\big)+[c]_*^{\widetilde{E}}(1+C_{t^*}C_\lambda)\\&\quad+C_c\big(2[\lambda]_1C_{t^*}+C_\lambda C_{t^*}^2[\lambda]_1+2[t^*]C_\lambda\big),\\ E_2&=C_cC_{t^*}C_\lambda[t^*],\end{aligned}$$

$$\widetilde{E}_3 = (1 + C_{t^*} C_\lambda)[\widetilde{Q}],$$
$$E_4 = 2C_\lambda [t^*] + C_{t^*} [\lambda]_1 (2 + C_{t^*} C_\lambda).$$

5.3. ***Deterministic time horizon.*** In the context of applications, it seems relevant to consider a deterministic time horizon t_f. For instance, one may want to estimate a mean cost over a given period no matter how many jumps occur during this period. Actually, we will choose a time horizon of the form $t_f \wedge T_N$ with N large enough to ensure the N-th jump will occur after time t_f with high probability: in other words, that $\boldsymbol{P}_x(T_N < t_f)$ be close to zero. For a discussion concerning the choice of such N, and in particular a theoretical bound of the probability $\boldsymbol{P}_x(T_N < t_f)$, we refer to [3]. Simply notice that in practice, this probability may be estimated through Monte Carlo simulations. We thus intend to approximate the following quantity for $l \in \boldsymbol{L}_c(\widetilde{E})$, $c \in \boldsymbol{L}_c(\partial \widetilde{E})$ and $x \in E$:

$$\boldsymbol{E}_x\left[\int_0^{T_N \wedge t_f} l(X_t, t)\, dt + \sum_{T_j \le t_f} c(X_{T_j^-}, T_j) \mathbb{1}_{\{X_{T_j^-} \in \partial E\}} \right]$$
$$= \boldsymbol{E}_x\left[\int_0^{T_N} l(X_t, t) \mathbb{1}_{\{t \le t_f\}} dt + \sum_{j=1}^{N} c(X_{T_j^-}, T_j) \mathbb{1}_{\{X_{T_j^-} \in \partial E\}} \mathbb{1}_{\{T_j \le t_f\}} \right].$$

The natural approach would consist in killing the process at time t_f as Davis suggests in [5, Section 31], and applying our method to the new process. However, the killed process will not necessarily fulfill our Lipschitz continuity assumptions because of the discontinuity introduced at time t_f.

A second idea would then be to use the previous results, to consider the time-augmented process, and to define $\tilde{l}(x,t) = l(x,t)\mathbb{1}_{\{t \le t_f\}}$ and $\tilde{c}(x,t) = c(x,t)\mathbb{1}_{\{t \le t_f\}}$. However, a similar problem appears. Indeed, such functions $\tilde{l}$ and $\tilde{c}$ are not Lipschitz continuous and our numerical method requires this assumption. In the rest of this section, we will see how to overcome this drawback. On the one hand, we prove that the Lipschitz continuity condition on l may be relaxed so that our numerical method may be used directly to approximate $\widetilde{J}_N(\tilde{l}, c)$ for any $c \in \boldsymbol{L}_c(\partial \widetilde{E})$. On the other hand, in the general case, we will deal with the non-Lipschitz continuity of $\tilde{c}$ by bounding it between two Lipschitz continuous functions.

5.3.1. *Direct estimation of the running cost term.* Let us explain how the Lipschitz continuity condition on the running cost function may be relaxed so that Theorem 4.5, stating the convergence of our approximation scheme, remains true when the running cost function is $\tilde{l}(x,t) = l(x,t)\mathbb{1}_{\{t \le t_f\}}$ with $l \in \boldsymbol{L}_c(\widetilde{E})$ and the boundary jump cost function is $c \in \boldsymbol{L}_c(\partial \widetilde{E})$ (although with slightly different constants in the bound of the convergence rate). Indeed, the running cost function $\tilde{l}$ appears inside an integral that will have a regularizing effect allowing us to derive

the required Lipschitz property of the functional in spite of the discontinuity of $\tilde{l}$. Details are provided in Appendix B.

Consequently, our approximation scheme may be used to compute $\widetilde{J}_N(\tilde{l},c)(\xi)$ for any $\xi \in \widetilde{E}$. We recall that $\widetilde{J}_N$ is defined by (7) and that for all $x \in E$, one has

$$\widetilde{J}_N(\tilde{l},c)(x,0) = \boldsymbol{E}_x\left[\int_0^{T_N \wedge t_f} l(X_t,t)\,dt + \sum_{j=1}^{N} c(X_{T_j^-},T_j)\mathbb{1}_{\{X_{T_j^-}\in\partial E\}}\right].$$

We now turn to the indicator function $\mathbb{1}_{\{T_j \le t_f\}}$ required within the boundary jump cost term.

5.3.2. *Bounds of the boundary jump cost term.* We explained how the Lipschitz continuity condition on l may be relaxed. However, when it comes to c, this condition cannot be avoided and our numerical method cannot be used directly with $\tilde{c}(x,t) = c(x,t)\mathbb{1}_{\{t\le t_f\}}$. We overcome this drawback by using Lipschitz continuous approximations of the indicator function. Indeed, for $B > 0$, we introduce the real-valued functions $\underline{u}_B$ and $\bar{u}_B$ defined on $\mathbb{R}$ by

$$\underline{u}_B(t) = \begin{cases} 1 & \text{if } t < t_f - 1/B, \\ -B(t-t_f) & \text{if } t_f - 1/B \le t < t_f, \\ 0 & \text{if } t_f \le t, \end{cases}$$

$$\bar{u}_B(t) = \begin{cases} 1 & \text{if } t < t_f, \\ -B(t-t_f)+1 & \text{if } t_f \le t < t_f + 1/B, \\ 0 & \text{if } t_f + 1/B \le t. \end{cases}$$

The following lemma is straightforward.

Lemma 5.6. *For all* $t \ge 0$, $\lim_{B\to+\infty} \underline{u}_B(t) = \mathbb{1}_{[0;t_f)}(t)$ *and* $\lim_{B\to+\infty} \bar{u}_B(t) = \mathbb{1}_{[0;t_f]}(t)$. *Furthermore, for all* $B > 0$, $\underline{u}_B$ *and* $\bar{u}_B$ *are Lipschitz continuous with Lipschitz constant* B. *Moreover,* $|\underline{u}_B - \mathbb{1}_{[0;t_f]}| \le 1$, $|\bar{u}_B - \mathbb{1}_{[0;t_f]}| \le 1$ *and*

$$\underline{u}_B \le \mathbb{1}_{[0;t_f]} \le \bar{u}_B.$$

Thus, define for $l \in \boldsymbol{L}_c(\widetilde{E})$

$$\tilde{l}(x,t) = l(x,t)\mathbb{1}_{\{t\le t_f\}} \tag{8}$$

and for $c \in \boldsymbol{L}_c(\partial\widetilde{E})$ and for all $B > 0$,

$$\underline{c}_B(x,t) = c(x,t)\underline{u}_B(t) \qquad \text{and} \qquad \bar{c}_B(x,t) = c(x,t)\bar{u}_B(t). \tag{9}$$

We now check that these functions satisfy our Lipschitz continuity conditions.

Proposition 5.7. *The functions* $\underline{c}_B$ *and* $\bar{c}_B$ *belong to* $\boldsymbol{L}_c(\partial\widetilde{E})$ *with* $[\underline{c}_B]_*, [\bar{c}_B]_* \le [c]_* + BC_c(1 \vee [t^*])$.

Proof. We prove the result for $\underline{c}_B$, the other case being similar. For all $\xi = (x,t), \xi' = (x',t') \in \widetilde{E}$, one has

$$\begin{aligned}
&\left|\underline{c}_B\big(\widetilde{\Phi}(\xi, t^*(\xi))\big) - \underline{c}_B\big(\widetilde{\Phi}(\xi', t^*(\xi'))\big)\right| \\
&\qquad = \left|c\big(\widetilde{\Phi}(\xi, \tilde{t^*}(\xi))\big)\underline{u}_B(t + \tilde{t^*}(\xi)) - c\big(\widetilde{\Phi}(\xi', \tilde{t^*}(\xi'))\big)\underline{u}_B(t' + \tilde{t^*}(\xi'))\right| \\
&\qquad \le [c]_*|\xi - \xi'| + C_c\left|\underline{u}_B(t + \tilde{t^*}(\xi)) - \underline{u}_B(t' + \tilde{t^*}(\xi'))\right| \\
&\qquad \le [c]_*|\xi - \xi'| + C_c B\big(|t - t'| + [\tilde{t^*}]|x - x'|\big) \\
&\qquad \le \big([c]_* + C_c B(1 \vee [t^*])\big)|\xi - \xi'|.
\end{aligned}$$

Hence the result. □

Therefore, the functions $\underline{c}_B$ and $\bar{c}_B$ are acceptable boundary jump cost functions and we may bound the deterministic horizon expectation by

$$\begin{aligned}
\widetilde{J}_N(\tilde{l}, \underline{c}_B)(x,0) &\le \boldsymbol{E}_x\left[\int_0^{T_N} l(X_t)\mathbb{1}_{\{t \le t_f\}} dt + \sum_{j=1}^{N} c(X_{T_j^-})\mathbb{1}_{\{X_{T_j^-} \in \partial E\}}\mathbb{1}_{\{T_j \le t_f\}}\right] \\
&\le \widetilde{J}_N(\tilde{l}, \bar{c}_B)(x,0).
\end{aligned}$$

The following proposition provides the convergence of the bounds.

Proposition 5.8. *For all $x \in E$, one has*

$$\begin{aligned}
&\lim_{B \to +\infty} \widetilde{J}_N(\tilde{l}, \underline{c}_B)(x,0) \\
&\qquad = \lim_{B \to +\infty} \widetilde{J}_N(\tilde{l}, \bar{c}_B)(x,0) \\
&\qquad = \boldsymbol{E}_x\left[\int_0^{T_N \wedge t_f} l(X_t, t)\, dt + \sum_{j=1}^{N} c(X_{T_j^-}, T_j)\mathbb{1}_{\{X_{T_j^-} \in \partial E\}}\mathbb{1}_{\{T_j \le t_f\}}\right].
\end{aligned}$$

Convergence holds for every $t_f > 0$ in the case of $\widetilde{J}_N(\tilde{l}, \bar{c}_B)(x,0)$ but only for almost every $t_f > 0$ with respect to the Lebesgue measure on $\mathbb{R}$ in the case of $\widetilde{J}_N(\tilde{l}, \underline{c}_B)(x,0)$.

Proof. Let $x \in E$. We first consider $\widetilde{J}_N(\tilde{l}, \bar{c}_B)(x,0)$.

$$\begin{aligned}
&\left|\boldsymbol{E}_x\left[\sum_{j=1}^{N} c(X_{T_j^-}, T_j)\mathbb{1}_{\{X_{T_j^-} \in \partial E\}}\mathbb{1}_{\{T_j \le t_f\}} - \sum_{j=1}^{N} \bar{c}_B(X_{T_j^-}, T_j)\mathbb{1}_{\{X_{T_j^-} \in \partial E\}}\right]\right| \\
&\qquad \le \boldsymbol{E}_x\left[\sum_{j=1}^{N} \left|c(X_{T_j^-}, T_j)\right|\left|\mathbb{1}_{\{T_j \le t_f\}} - \bar{u}_B(T_j)\right|\right] \\
&\qquad \le C_c \boldsymbol{E}_x\left[\sum_{j=1}^{N} \mathbb{1}_{\{t_f < T_j \le t_f + \frac{1}{B}\}}\right] \le C_c \sum_{j=1}^{N} \left(\varphi_j\left(t_f + \frac{1}{B}\right) - \varphi_j(t_f)\right),
\end{aligned}$$

where φ_j is the distribution function of T_j. For all $j \leq N$, the summand in this last expression goes to 0 as $B \to +\infty$, since φ_j is right-continuous; this shows the required convergence.

We now turn to the case of $\widetilde{J}_N(\tilde{l}, \underline{c}_B)(x, 0)$. Similar computations yield

$$\left| \boldsymbol{E}_x \left[\sum_{j=1}^{N} c(X_{T_j^-}, T_j) \mathbb{1}_{\{X_{T_j^-} \in \partial E\}} \mathbb{1}_{\{T_j \leq t_f\}} - \sum_{j=1}^{N} \underline{c}_B(X_{T_j^-}, T_j) \mathbb{1}_{\{X_{T_j^-} \in \partial E\}} \right] \right| \leq C_c \sum_{j=1}^{N} \left(\varphi_j(t_f) - \varphi_j\left(t_f - \frac{1}{B}\right) \right).$$

One cannot conclude as in the previous case, since φ_j need not be left-continuous. We therefore assume that t_f is not an atom of any of the laws of the random variables T_j. Then, for all $j \leq N$, the summand on the right-hand side tends to 0 as $B \to +\infty$, and the result follows. Indeed, the set of the atoms of T_j is at most countable, so the convergence holds for almost every t_f with respect to the Lebesgue measure on $\mathbb{R}$. □

5.3.3. *Bounds in the general case.* The previous results show that the deterministic horizon expectation may be bounded by applying our numerical method with $\tilde{l}$ and successively $\underline{c}_B$ and $\bar{c}_B$. In other words, we have shown:

Theorem 5.9. *Let $l \in \boldsymbol{L}_c(\widetilde{E})$ and $c \in \boldsymbol{L}_c(\partial \widetilde{E})$. Let $(\underline{V}_{k,B})_{0 \leq k \leq N}$ (respectively $(\overline{V}_{k,B})_{0 \leq k \leq N}$) be the sequence of random variables $(V_k)_{0 \leq k \leq N}$ described in Section 4 when applying our approximation scheme to the time-augmented process $(\widetilde{X}_t)_{t \geq 0}$ with cost functions $\tilde{l}$ and $\underline{c}_B$ (respectively $\bar{c}_B$) defined by (8) and (9). The bounds of the approximation error provided by Theorem 4.5 are respectively denoted by*

$$\varepsilon_N(l, \underline{c}_B, \widetilde{X}, A, B) \quad \text{and} \quad \varepsilon_N(l, \bar{c}_B, \widetilde{X}, A, B).$$

One has then

$$\begin{aligned} \underline{V}_{0,B} - \varepsilon_N(l, \underline{c}_B, \widetilde{X}, A, B) &\leq \boldsymbol{E}_x \left[\int_0^{T_N \wedge t_f} l(X_t, t)\, dt + \sum_{j=1}^{N} c(X_{T_j^-}, T_j) \mathbb{1}_{\{X_{T_j^-} \in \partial E\}} \mathbb{1}_{\{T_j \leq t_f\}} \right] \\ &\leq \overline{V}_{0,B} + \varepsilon_N(l, \bar{c}_B, \widetilde{X}, A, B). \end{aligned}$$

Remark 5.10. In the previous theorem, the quantity $\varepsilon_N(l, \underline{c}_B, \widetilde{X}, A, B)$ (and similarly $\varepsilon_N(l, \bar{c}_B, \widetilde{X}, A, B)$) is computed with respect to the process $(\widetilde{X}_t)_{t \geq 0}$ instead

of $(X_t)_{t\geq 0}$ as presented in Theorem 4.5 so that one has

$$\begin{aligned}
&\varepsilon_N(l, \underline{c}_B, \widetilde{X}, A, B) \\
&\quad = \sum_{k=0}^{N-1} \Big(2[v_{k+1}]^{\widetilde{E}} \|\widetilde{Z}_{k+1} - \widehat{\widetilde{Z}}_{k+1}\|_p \\
&\qquad + \big(2[v_k]^{\widetilde{E}} + [F]_1' + [F]_1'' A + [F]_1''' B\big) \|\widetilde{Z}_k - \widehat{\widetilde{Z}}_k\|_p \\
&\qquad + \big([F]_2' + [F]_2'' A\big) \|\widetilde{S}_{k+1} - \widehat{\widetilde{S}}_{k+1}\|_p \Big) + \frac{N C_c C_\lambda}{A}.
\end{aligned}$$

where $(\widetilde{Z}_k, \widetilde{S}_k)_{k\in\mathbb{N}}$ denotes the sequence of the postjump locations and the interjump times of the time-augmented process $(\widetilde{X}_t)_{t\geq 0}$ and with

$$\begin{aligned}
[F]_1''' &= C_c(1 \vee [t^*]), \\
[v_n]_1^{\widetilde{E}} &\leq e^{C_{t^*} C_\lambda} \big(\widetilde{K}(A, B, v_{n-1}) + n C_{t^*} [\lambda]_1 (C_{t^*} C_l + C_c)\big) + C_{t^*} [l]_1^{\widetilde{E}}, \\
[v_n]^{\widetilde{E}} &\leq \widetilde{K}(A, B, v_{n-1}),
\end{aligned}$$

and for all $w \in \boldsymbol{L}_c(E)$ we have

$$\widetilde{K}(A, B, w) = E_1' + E_2'' B + E_2 A + \widetilde{E}_3 [w]_1^{\widetilde{E}} + E_4 C_w + [\widetilde{Q}][w]_*^{\widetilde{E}},$$

where

$$\begin{aligned}
E_1' &= 2[l]_1^{\widetilde{E}} C_{t^*} + C_l \big([t^*] + 2C_{t^*}^2 [\lambda]_1\big) + [c]_*^{\widetilde{E}} (1 + C_{t^*} C_\lambda) \\
&\qquad + C_c \big(2[\lambda]_1 C_{t^*} + C_\lambda C_{t^*}^2 [\lambda]_1 + 2[t^*] C_\lambda\big), \\
E_1'' &= C_c (1 \vee [t^*])(1 + C_{t^*} C_\lambda)
\end{aligned}$$

The other constants remain unchanged; see Remark 5.5 for their expressions.

Furthermore, it is important to stress the fact that applying twice our numerical method does not increase significantly the computing time. Indeed, the computation of the quantization grids is, by far, the most costly step. These grids, that only depend on the dynamics of the process, may be stored off-line and used for the approximation of both bounds.

The choice of B. We now discuss the choice of the parameter B, the discussion is quite similar to the one concerning the choice of A in Section 4.2. Proposition 5.8 suggests that B should be chosen as large as possible. However, choosing a large value for B will lead to large Lipschitz constants that will decrease the sharpness of the bounds $\varepsilon_N(l, \underline{c}_B, \widetilde{X})$ and $\varepsilon_N(l, \bar{c}_B, \widetilde{X})$ for the approximation error provided by Theorem 4.5. Indeed, it is easy to check that $[v_n]$ grows linearly with B (see the precise expressions of the Lipschitz constants above). Thus, in order to control the error proposed by Theorem 4.5, it is necessary that the order of magnitude of the quantization error $\|\Theta_n - \widehat{\Theta}_n\|_p$ be at most $1/B$.

6. Numerical results

6.1. ***A repair workshop model.*** We now present a repair workshop model adapted from [5, Section 21].

In a factory, a machine produces goods which daily value is $r(x)$, where $x \in [0; 1]$ represents a parameter of evolution of the machine, a setting chosen by the operator. For instance, x may be some load or some pace imposed on the machine. This machine, initially working, may break down with an age-dependent hazard rate $\lambda(t)$ and is then sent to the workshop for repair. Besides, the factory's management has decided that, whenever the machine has worked for a whole year without requiring repair, it is sent to the workshop for maintenance. The daily cost of maintenance is $q(x)$, while the daily cost of a repairs is $p(x)$, with reasonably $p(x) > q(x)$. We assume that after a repair or maintenance, both lasting a fixed time s, the machine is totally repaired and is not worn down.

We therefore consider three modes: the machine is working ($m = 1$), being repaired ($m = 2$), or undergoing maintenance ($m = 3$). The state of the process at time t will be denoted by $X_t = (m_t, \zeta_t, t)$, where ζ_t is the time since the last change of mode. (This component is required since the hazard rate λ is age-dependent.) The state space is $E = \big(\{1\}\times[0; 365]\times\mathbb{R}^+\big) \cup \big(\{2\}\times[0; s]\times\mathbb{R}^+\big) \cup \big(\{3\}\times[0; s]\times\mathbb{R}^+\big)$. In each mode, the flow is $\Phi_m\big((\zeta, t), u\big) = (\zeta + u, t + u)$. Concerning the transition kernel, one sees from the previous discussion that, for instance, from the point $(1, \zeta, t)$, the process can jump to the point $(2, 0, t)$ if $\zeta < 365$ and the jump is forced to $(3, 0, t)$ if $\zeta = 365$. Figure 1 presents the state space and an example of trajectory of the process.

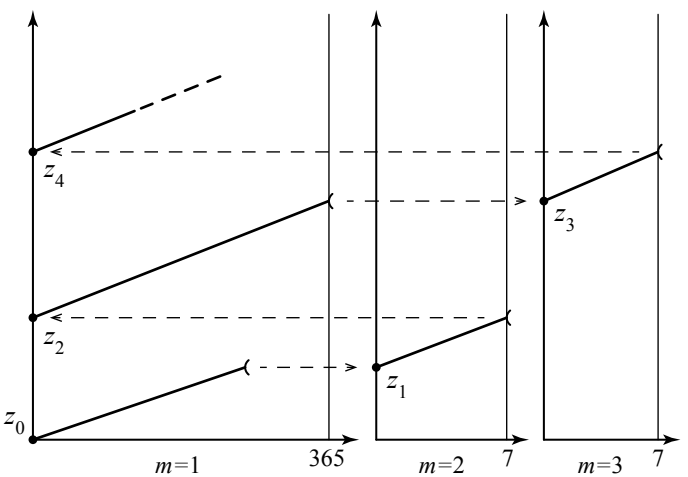

Figure 1. An example trajectory. The process starts from the point Z_0 in mode $m = 1$ (machine in service). The machine may be sent to the workshop for repairs ($m = 2$) or for maintenance ($m = 3$).

Our aim is to find the value of the setting x that maximizes the expected total benefits $B(x)$, that is, the discounted value (for an interest rate ρ) of production minus maintenance and reparation costs over a period $t_f = 5$ years:

$$B^* = \sup_{x\in[0;1]} B(x),$$

where

$$B(x) = \boldsymbol{E}_{(1,0,0)}\left[\int_0^{t_f} e^{-\rho t}\left(r(x)\mathbb{1}_{\{m_t=1\}} - p(x)\mathbb{1}_{\{m_t=2\}} - q(x)\mathbb{1}_{\{m_t=3\}}\right) dt\right].$$

We will use the following values $r(x) = x$, $p(x) = 100x^2$, $q(x) = 5$, $s = 7$ days, $\rho = \frac{0.03}{365}$ and λ represents a Weibull distribution with parameters $\alpha = 2$ et $\beta = 600$.

Our assumptions clearly hold so that we may run our numerical method. We first need to find $N \in \mathbb{N}$ such that $\boldsymbol{P}_{(1,0,0)}(T_N < t_f)$ be small. Monte Carlo simulations lead to the value $N = 18$. For a fixed $x \in [0; 1]$, we will therefore compute $\widetilde{J}_N(\tilde{l}, 0)(1, 0, 0)$ where $\tilde{l}(m, \zeta, t) = e^{-\rho t}\big(r(x)\mathbb{1}_{\{m=1\}} - p(x)\mathbb{1}_{\{m=2\}} - q(x)\mathbb{1}_{\{m=3\}}\big)\mathbb{1}_{\{t\le t_f\}}$. Finally, notice that we could have chosen r, p and q slightly more generally by allowing them to be time-dependent.

It is important to stress the fact that, once the Markov chain associated to the process is quantized, we will be able to compute the approximation of $B(x)$ almost instantly for any $x \in [0; 1]$ because the same grids are used for every computation. Thanks to this flexibility, we are able to draw the function $x \to B(x)$ and, thus, to solve the above optimization problem very easily. This is a very important advantage of our method. Indeed, if we computed $B(x)$ through standard methods such as Monte Carlo simulations, we would have to repeat the whole algorithm again and again for each value of x and solving the optimization problem would be intractable.

The following figure represents the approximation of the function B computed on a constant step grid of $[0; 1]$ with step 10^{-2}. This leads to the solution of the earlier optimization problem. Indeed, we obtain $B^* = B(x^*) = 537.84$ where $x^* = 0.78$ is the value of the setting x that maximizes the benefits of the factory.

Now let $x = 0.78$. The following table presents the values of $\widehat{V}_N$, which are the approximations of $B(x)$, for different number of points in the quantization grids. A reference value $B_{\text{Monte Carlo}} = 537.69$ is obtained via the Monte Carlo method (10^8 simulations).

Points in the quantization grids	$\widehat{V}_N$	relative error to 537.69
20 points	542.14	0.83%
50 points	539.57	0.35%
100 points	538.24	0.10%
500 points	537.84	0.03%

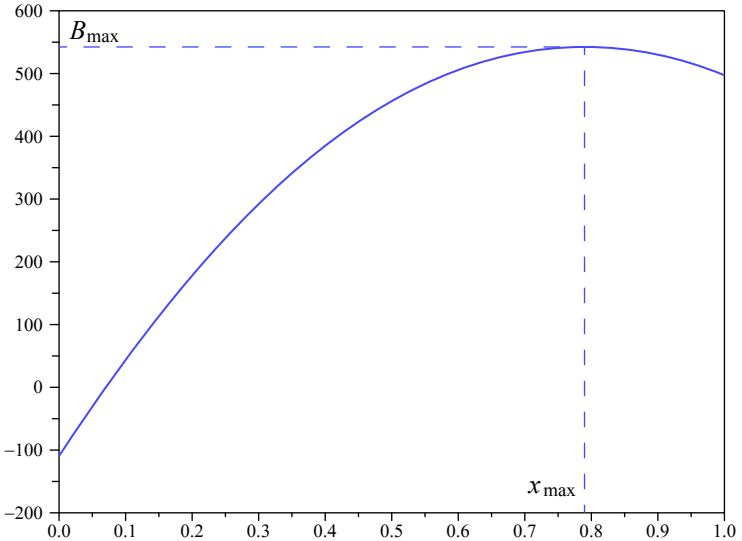

Figure 2. The function B drawn with 500 points in the quantization grids.

From a computational time point of view, we have already explained that the computation of large quantization grids is, by far, the most costly step since it may take up to several hours whereas the approximation of the expectation that follows is then almost instantaneous. However, we may notice, in the above table, that grids containing only 50 points yield a quite accurate result with merely 0.35% error. Such grids only require a few minutes to be designed.

Remark 6.1. We already noticed that the same grids may serve several purposes. For instance, we may also have been interested in the computation of the mean time spent by the machine in the workshop by taking $l(m, \zeta, t) = \mathbb{1}_{\{m \in \{2;3\}\}}$.

6.2. ***A corrosion model.*** We consider here a corrosion model for an aluminum metallic structure. This example was provided by Astrium. It concerns a small structure within a strategic ballistic missile. The missile is stored successively in three different environments which are more or less corrosive. It is made to have potentially large storage durations. The requirement for security is very strong. The mechanical stress exerted on the structure depends in part on its thickness. A loss of thickness will cause an overconstraint and therefore increase the risk of rupture. It is thus crucial to study the evolution of the thickness of the structure over time.

Let us describe more precisely the usage profile of the missile. It is stored successively in three different environments: the workshop ($m = 1$), the submarine in operation ($m = 2$) and the submarine in dry-dock ($m = 3$). This is because the

structure must be equipped and used in a given order. Then it goes back to the workshop and so on. The missile stays in each environment during a random duration with exponential distribution. Its parameter λ_m depends on the environment. The degradation law for the thickness loss then depends on the environment through two parameters, a deterministic transition period η_m and a random corrosion rate ρ uniformly distributed within a given range. Typically, the workshop and dry-dock are the most corrosive environments but the time spent in operation is more important. The randomness of the corrosion rate accounts for small variations and uncertainties in the corrosiveness of each environment.

In each environment $m \in \{1; 2; 3\}$, the thickness loss d_m evolves in time as

$$d_m(\rho, s) = \rho\big(s + \eta_m(e^{-s/(2\eta_m)} - 1)\big). \tag{10}$$

Here are the numerical values of the parameters of the corrosion model:

	environment 1	environment 2	environment 3
λ_m (h^{-1})	$(17520)^{-1}$	$(131400)^{-1}$	$(8760)^{-1}$
η_m (h)	30000	200000	40000
ρ (mm/h)	$[10^{-6}, 10^{-5}]$	$[10^{-7}, 10^{-6}]$	$[10^{-6}, 10^{-5}]$

Initially, the structure is in environment $m = 1$ and the thickness loss is null. One draws the corrosion rate ρ_0 uniformly distributed in the interval $[10^{-6}, 10^{-5}]$ and the time of the first change of environment T_1 exponentially distributed with parameter $\lambda_1 = (17520)^{-1}$ hours^{-1}. The corrosion starts according to (10) so that, for all $0 \le t \le T_1$, the loss of thickness is $d_1(\rho_0, t)$. The structure then moves to environment 2 and the process restarts similarly: a new corrosion rate ρ_{T_1} is drawn according to an uniform law on $[10^{-7}, 10^{-6}]$, the time of the second jump T_2 is drawn so that $T_2 - T_1$ is exponentially distributed with parameter $\lambda_2 = (131400)^{-1}$ hours^{-1} and for $T_1 \le t \le T_2$, the loss of thickness is $d_1(\rho_0, T_1) + d_2(\rho_{T_1}, t - T_1)$ and so on.

At each change of environment, a new corrosion rate ρ is drawn according to a uniform law on the corresponding interval. The thickness loss, however, evolves continuously.

We are interested in computing the mean loss of thickness in environment 2 until a given time $t_f = 18$ years.

Modeling by PDMP.

The state space E. The loss of thickness will be modeled by a PDMP whose modes are the different environments. Let then $M = \{1, 2, 3\}$. The PDMP $(X_t)_{t \ge 0}$ will contain the following components: the mode $m \in M$, the loss of thickness d, the time since the last jump s (this is to ensure that the Markov property is satisfied),

the corrosion rate ρ and the time t (since we consider the time-augmented process). Clearly, one has always $s \le t$, so we can reasonably consider the state space

$$E = \big\{(m,d,s,\rho,t) \in M \times \mathbb{R}^+ \times \mathbb{R}^+ \times [10^{-7};10^{-5}] \times \mathbb{R}^+ \text{ such that } s \le t\big\}.$$

The flow Φ. The flow is given for all $u \ge 0$ by

$$\Phi\left(\begin{pmatrix} m \\ d \\ s \\ \rho \\ t \end{pmatrix}, u\right) = \begin{pmatrix} m \\ d + d_m(\rho, s+u) - d_m(\rho,s) \\ s+u \\ \rho \\ t+u \end{pmatrix}.$$

The transition kernel Q. Let us now study the jumps of this process. When the process jumps from a point $x = (m,d,s,\rho,t) \in E$, m becomes $m+1$ modulo 3 (denoted $m+1[3]$), d and t remain unchanged, s becomes 0. Only ρ is randomly drawn, according to a uniform law on an interval $[\rho_{\min};\rho_{\max}]$ that depends on the new mode. One has then for $w \in B(E)$, $x = (m,d,s,\rho,t) \in E$, and $u \ge 0$,

$$\begin{aligned} Qw\big(\Phi\left(\begin{pmatrix} m \\ d \\ s \\ \rho \\ t \end{pmatrix}, u\right)\big) &= Qw \begin{pmatrix} cm \\ d + d_m(\rho, s+u) - d_m(\rho,s) \\ s+u \\ \rho \\ t+u \end{pmatrix} \\ &= \frac{1}{\rho_{\max} - \rho_{\min}} \int_{\rho_{\min}}^{\rho_{\max}} w \begin{pmatrix} m+1[3] \\ d + d_m(\rho, s+u) - d_m(\rho,s) \\ 0 \\ \tilde{\rho} \\ t+u \end{pmatrix} d\tilde{\rho}. \end{aligned} \tag{11}$$

The cost function l. The function $l \in B(E)$ will be the cost function to compute the mean loss of thickness in mode 2. It is defined as follows: for all $x = (m,d,s,\rho,t) \in E$ and $u \ge 0$,

$$l(\Phi(x,u)) = \rho\big(1 - \tfrac{1}{2} e^{-(s+u)/(2\eta_m)}\big)\mathbb{1}_{\{m=2\}} = \frac{d}{du}\big(d_m(\rho, s+u)\big)\mathbb{1}_{\{m=2\}}. \tag{12}$$

One then defines $\tilde{l}(\Phi(x,u)) = l(\Phi(x,u))\mathbb{1}_{\{t+u \le t_f\}}$, so that

$$\begin{aligned} L(x,u) &= \int_0^u \tilde{l}\big(\Phi(x,u')\big)\, du' = \int_0^{u \wedge (t_f - t)^+} l\big(\Phi(x,u')\big)\, du' \\ &= \big(d_m(\rho, s + u \wedge (t_f - t)^+) - d_m(\rho,s)\big)\mathbb{1}_{\{m=2\}}; \end{aligned}$$

that is indeed the thickness lost in mode 2 from the point $x = (m, d, s, \rho, t)$ during a time $u \wedge (t_f - t)^+$.

The assumptions. Assumptions 2.1 and 2.9 are clearly satisfied. It is easy to check, from (12), that $l \in \boldsymbol{L}_c(E)$, so Assumption 3.1 holds.

We now turn to Assumption 2.13 and we will see that, although it does not hold for any function $w \in \boldsymbol{L}_c^v(E)$, it holds for a sufficiently big subclass of functions. We first need to make a remark. Recall that for all $x = (m, d, s, \rho, t) \in E$ and for all $k \in \{0, \dots, N\}$, one has $v_{N-k}(x) = \boldsymbol{E}_x bigl[\int_0^{T_k} l\big(\Phi(x,u)\big)\mathbb{1}_{\{t+u\le t_f\}} du\bigr]$. Therefore, for all $k \in \{0, \dots, N\}$ the function v_k as well as the function $\tilde{l}$ satisfy

$$x = (m, d, s, \rho, t) \in E \text{ and } t \ge t_f \implies w(x) = 0. \tag{13}$$

The next step consists in proving that Assumption 2.13, although it is not satisfied for any function $w \in \boldsymbol{L}_c^v(E)$, holds for any function $w \in \boldsymbol{L}_c^v(E)$ that also satisfies condition (13). This is done in Lemma 6.2 and it is sufficient because in the proof of the theorem that ensures the convergence of our approximation scheme, Assumption 2.13 is only used with the functions $(v_k)_{k\in\{0,\dots,N\}}$ that do satisfy condition (13).

Lemma 6.2. *There exists $[Q] \in \mathbb{R}^+$ such that for all $v \ge 0$ and $w \in \boldsymbol{L}_c^v(E)$ that satisfies condition (13), one has for all $x, x' \in E$ and $0 \le u \le v$,*

$$\big|Qw\big(\Phi(x,u)\big) - Qw\big(\Phi(x',u)\big)\big| \le [Q][w]_1^{E,v}|x - x'|.$$

Proof. Let $x = (m, d, s, \rho, t)$ and $x' = (m', d', s', \rho', t') \in E$ with for instance $t \le t'$. First we may choose $m = m'$; otherwise, $|x - x'| = +\infty$ and there is nothing to prove. Now, we are facing three different cases:

- If $t_f \le t + u \le t' + u$, then one has $Qw\big(\Phi(x,u)\big) = Qw\big(\Phi(x',u)\big) = 0$ because w satisfies condition (13) and there is nothing to prove.
- If $t + u \le t_f \le t' + u$, notice that

$$Qw\big(\Phi(x',u)\big) = Qw\big(\Phi((m', d', s', \rho', t_f), u)\big) = 0$$

(this stems from condition (13)), so that we are reduced to the following case.

- We assume from now on that $t + u \le t' + u \le t_f$. We now intend to bound $\big|Qw\big(\Phi(x,u)\big) - Qw\big(\Phi(x',u)\big)\big|$. It is clear from (11) that we only need to prove that the function $(\rho, s) \to d_m(\rho, s)$, defined by (10), is Lipschitz continuous with respect to both its variables on the set $[10^{-7}; 10^{-5}] \times [0; t_f]$. Indeed, we have $s \le t$ and $s' \le t'$ so that $s, s', s + u, s' + u \le t_f$. Standard computations yield

$$|d_m(\rho, s) - d_m(\rho', s')| \le s|\rho - \rho'| + \tfrac{3}{2}\rho'|s - s'| \le t_f|\rho - \rho'| + \tfrac{3}{2}10^{-5}|s - s'|.$$

Hence the result. □

Assumption 2.10 is not satisfied because in our corrosion model, one has $t^*(x) = +\infty$ for all $x \in E$. Besides, we may notice that the previous proof would have been more straightforward if t^* had been bounded. Indeed in that case, we would have had $s, s', s+u, s'+u \leq C_{t^*}$ and the introduction of condition (13) would have been unnecessary. Nevertheless, we have been able to overcome the drawback of having t^* unbounded by noticing that somehow the deterministic time horizon t_f plays the part of the missing C_{t^*}. This is the meaning of condition (13): roughly speaking, we do not consider what happens beyond t_f.

More generally, we will now see that in our deterministic time horizon problem, the boundedness of t^* may be dropped and our results remain true replacing C_{t^*} by t_f. This is clear in the case of Proposition A.2 because the function $\tilde{l}$ satisfies the condition (13). Proposition A.7 remains also true replacing C_{t^*} by t_f. Indeed, on the one hand, it is clear that $L(x,u) \leq t_f C_l$. On the other hand, when computing $|v_n(\Phi(x,u)) - v_n(\Phi(x',u'))|$, we are facing three cases, as in the proof of Lemma 6.2:

- If $t_f \leq u \leq u'$, one has

$$v_n(\Phi(x,u)) = v_n(\Phi(x',u')) = 0,$$

by condition (13).

- If $u \leq t_f \leq u'$, one has

$$\big|v_n(\Phi(x,u)) - v_n(\Phi(x',u'))\big| = \big|v_n(\Phi(x,u)) - v_n(\Phi(x',t_f))\big|,$$

since $v_n(\Phi(x',u')) = v_n(\Phi(x',t_f)) = 0$ (condition (13) once again), so that we are reduced to the next case.

- If $u \leq u' \leq t_f$, the computations remain unchanged and t_f replaces C_{t^*} as a bound for u and u'.

Numerical results. The table below presents the values of the loss of thickness in environment 2 obtained through our approximation scheme with quantization grids of varying fineness, as well as the relative deviation with respect to the Monte Carlo value of 0.036755, obtained with 10^8 simulations.

Quantization grids	$\widehat{V}_0$	error	Quantization grids	$\widehat{V}_0$	error
20 points	0.038386	4.43%	2000 points	0.037041	0.77%
50 points	0.037804	2.85%	4000 points	0.037007	0.69%
100 points	0.037525	2.09%	6000 points	0.036973	0.57%
200 points	0.037421	1.81%	8000 points	0.036944	0.49%
500 points	0.037264	1.38%	10000 points	0.036911	0.40%
1000 points	0.037160	1.10%	12000 points	0.036897	0.36%

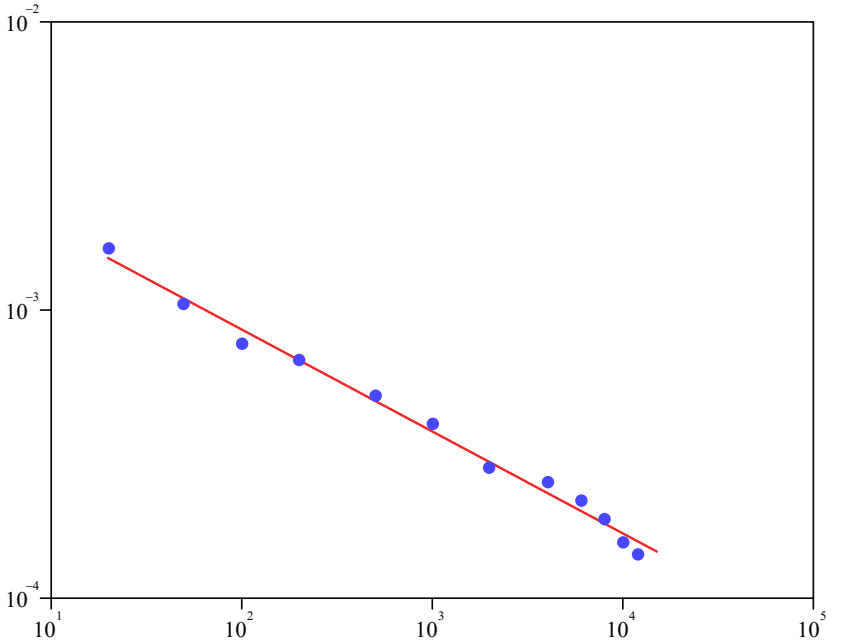

Figure 3. Log-log plot of error when approximating the loss of thickness in environment 2 versus number of points in the quantization grids. The empirical convergence rate, estimated through a regression model, is -0.35.

Figure 3 presents respectively the empirical convergence rate. The convergence rate, estimated through a regression model is -0.35. This is roughly the same order of magnitude as the rate of convergence of the optimal quantizer (see for instance [9]) since here the dimension is 3 (indeed, m is deterministic and $s = 0$ immediately after a jump so that we only quantize the variables ρ, d and t).

Finally, we show here the CPU time to compute the expectations from the quantization grids (computations are run with Matlab R2010b on a MacBook Pro 2.66 GHz i7 processor). The CPU time for 10^8 Monte Carlo simulations was approximately 16 000 s. It can be seen that, once the quantization grids are obtained, our approximation scheme performs very fast.

Quantization grids	CPU time (s)	Quantization grids	CPU time (s)
20 points	0.0059	2000 points	1.5
50 points	0.0085	4000 points	5.6
100 points	0.014	6000 points	13
200 points	0.034	8000 points	24
500 points	0.12	10000 points	35
1000 points	0.37	12000 points	54

7. Conclusion

We have presented an efficient and easy to implement numerical method to approximate expectations of functionals of piecewise deterministic Markov processes. We proved the convergence of our algorithm with bounds for the rate of convergence.

Although our method concerns time invariant functionals, we proved that we are able to tackle time-dependent problems such as Lipschitz continuous time-dependent functionals or deterministic time horizon expectations. Indeed, we proved that, thanks to the introduction of the time-augmented process, time-dependent problems may be seen, paradoxically, as special cases of the time invariant situation.

Our method is easy to implement because it merely requires to be able to simulate the process. Furthermore, although the computation of the quantization grids may be quite time-consuming, it may be performed preliminarily because the grids only depend on the dynamics of the process and not on the cost functions l and c. Therefore, they may be stored off-line and serve several purposes. As illustrated by the examples presented in Section 6, storing the grids provides to our approximation scheme efficiency and flexibility. Indeed, the computation of the expectation can be performed very quickly once the grids are available. Thus, if one decides for instance to modify the functional, the same grids may be used so that the new result is obtained very quickly. This flexibility is an important advantage over standard Monte Carlo simulations.

Appendix A. Lipschitz continuity of F, G and v_n

The first lemma and the first proposition of this section present mainly the Lipschitz continuity of the functions δ^A and F. They are stated without proof because they are quite straightforward.

Lemma A.1. *The function δ^A is Lipschitz continuous with respect to both its variables; i.e., for all $x, y \in E$ and $u, t \in \mathbb{R}$, one has*

$$|\delta^A(x,t) - \delta^A(y,t)| \leq A[t^*]|x-y|,$$
$$|\delta^A(x,t) - \delta^A(x,u)| \leq A|t-u|,$$

Moreover, one has for all $x \in E$ and $t, s \geq 0$ such that $t+s \leq t^*(x)$,

$$\delta^A(\Phi(x,s),t) = \delta^A(x,t+s).$$

Proposition A.2. *The function F introduced in Definition 3.3, is Lipschitz continuous with respect to both its variables. For all $x, y \in E$ and $u, v \in [0; t^*(x) \wedge t^*(y)]$, one has*

$$|F(x,u) - F(y,v)| \leq [F]_1|x-y| + [F]_2|u-v|,$$

with

$$[F]_1 = C_{t^*}[l]_1 + [c]_* + A[t^*]C_c, \quad [F]_2 = C_l + AC_c.$$

The next two lemmas are adapted from [6], the second one being a special case of Lemma A.1 there. Thus, they are stated without proof.

Lemma A.3. *For $h \in \boldsymbol{L}_c(E)$, $(x, y) \in E^2$, and $t \le t^*(x) \wedge t^*(y)$*

$$\left| \int_t^{t^*(x)} h(\Phi(x,s)) e^{-\Lambda(x,s)} ds - \int_t^{t^*(y)} h(\Phi(y,s)) e^{-\Lambda(y,s)} ds \right| \le \big(C_{t^*}[h]_1 + (C_{t^*}^2[\lambda]_1 + [t^*])C_h\big)|x-y|.$$

Lemma A.4. *For $h \in \boldsymbol{L}_c(\partial E) \cup \boldsymbol{L}_c(E)$ and $x, y \in E$, one has*

$$\left| e^{-\Lambda(x,t^*(x))} h\big(\Phi(x,t^*(x))\big) - e^{-\Lambda(y,t^*(y))} h\big(\Phi(y,t^*(y))\big) \right| \le \big([h]_* + C_h\big(C_{t^*}[\lambda]_1 + [t^*]C_\lambda\big)\big)|x-y|.$$

The following notation will be convenient later on. For $w \in \boldsymbol{L}_c(E)$, $x \in E$ and $t \in [0; t^*(x)]$, we define

$$\begin{aligned} G_t w(x) &= E_x\big[\big(F(x,S_1) + w(Z_1)\big)\mathbb{1}_{\{S_1 \ge t\}}\big] \\ &= E_x\big[\big(L(x,S_1) + C(x,S_1) + w(Z_1)\big)\mathbb{1}_{\{S_1 \ge t\}}\big]. \end{aligned}$$

In particular, $G_0 = G$. Since we know the law of (Z_1, S_1), it can be shown that

$$G_t w(x) = \Upsilon_1(x) + \Upsilon_2(x) + \Upsilon_3(x) + \Upsilon_4(x) + \Upsilon_5(x), \tag{14}$$

with

$$\begin{aligned} \Upsilon_1(x) &= e^{-\Lambda(x,t)} \int_0^t l \circ \Phi(x,s)\, ds, \\ \Upsilon_2(x) &= \int_t^{t^*(x)} l \circ \Phi(x,s) e^{-\Lambda(x,s)} ds. \\ \Upsilon_3(x) &= c \circ \Phi(x, t^*(x)) \int_t^{t^*(x)} \delta^A(x,s) \lambda \circ \Phi(x,s) e^{-\Lambda(x,s)} ds, \\ \Upsilon_4(x) &= \int_t^{t^*(x)} \big(\lambda Q w\big) \circ \Phi(x,s) e^{-\Lambda(x,s)} ds, \\ \Upsilon_5(x) &= e^{-\Lambda(x,t^*(x))} \big(Qw + c\big) \circ \Phi(x, t^*(x)). \end{aligned}$$

Proposition A.5. *For $w \in \boldsymbol{L}_c(E)$, $(x, y) \in E^2$ and $t \in [0; t^*(x) \wedge t^*(y)]$, one has*

$$\big|G_t w(x) - G_t w(y)\big| \le K(A, w)|x-y|,$$

where $K(A,w) = E_1 + E_2 A + E_3[w]_1 + E_4 C_w + [Q][w]_*$, *with*

$$\begin{aligned}
E_1 &= 2[l]_1 C_{t^*} + C_l\big([t^*] + 2C_{t^*}^2[\lambda]_1\big) + [c]_*(1 + C_{t^*}C_\lambda) \\
&\qquad + C_c\big(2[\lambda]_1 C_{t^*} + C_\lambda C_{t^*}^2[\lambda]_1 + 2[t^*]C_\lambda\big), \\
E_2 &= C_c C_{t^*} C_\lambda [t^*], \\
E_3 &= (1 + C_{t^*}C_\lambda)[Q], \\
E_4 &= 2C_\lambda[t^*] + C_{t^*}[\lambda]_1(2 + C_{t^*}C_\lambda).
\end{aligned}$$

Proof. Let $w \in \boldsymbol{L}_c(E)$, $(x,y) \in E^2$ and $t \in [0; t^*(x) \wedge t^*(y)]$. In view of (14), we naturally split $|G_t w(x) - G_y w(y)|$ into the sum of five differences.

The first one, $|\Upsilon_1(x) - \Upsilon_1(y)|$, is bounded by

$$\begin{aligned}
|\Upsilon_1(x) - \Upsilon_1(y)| &\le C_{t^*} C_l \big|e^{-\Lambda(x,t)} - e^{-\Lambda(y,t)}\big| + \int_0^t \big(l\circ\Phi(x,s) - l\circ\Phi(y,s)\big)\,ds \\
&\le \big(C_{t^*}^2 C_l[\lambda]_1 + C_{t^*}[l]_1\big)|x-y|.
\end{aligned}$$

The differences $|\Upsilon_2(x) - \Upsilon_2(y)|$ and $|\Upsilon_4(x) - \Upsilon_4(y)|$ can be bounded thanks to Lemma A.3, with successively $h = l$ and $h = \lambda Q w$. Notice that $C_{\lambda Q w} \le C_\lambda C_w$ and $[\lambda Q w]_1 \le C_\lambda [Q][w]_1 + C_w[\lambda]_1$.

For the difference of the Υ_5 terms, we use Lemma A.4 with $h = Qw + c$. Notice that $C_{Qw+c} \le C_w + C_c$ and that $[Qw + c]_* \le [Q]\big([w]_* + [w]_1\big) + [c]_*$.

Finally, to bound $|\Upsilon_3(x) - \Upsilon_3(y)|$, we assume without loss of generality that $t^*(x) \le t^*(y)$ and we have

$$\begin{aligned}
&|\Upsilon_3(x) - \Upsilon_3(y)| \\
&\le C_c \int_t^{t^*(x)} \big|\delta^A(x,s)\lambda\circ\Phi(x,s)e^{-\Lambda(x,s)} - \delta^A(y,s)\lambda\circ\Phi(y,s)e^{-\Lambda(y,s)}\big|\,ds \\
&+ C_c \int_{t^*(x)}^{t^*(y)} \big|\delta^A(y,s)\lambda\circ\Phi(y,s)e^{-\Lambda(y,s)}\big|\,ds + [c]_* C_{t^*} C_\lambda |x-y| \\
&\le C_c \int_t^{t^*(x)} \big(C_\lambda\big|\delta^A(x,s) - \delta^A(y,s)\big| + [\lambda]_1|x-y| + C_\lambda\big|e^{-\Lambda(x,s)} - e^{-\Lambda(y,s)}\big|\big)\,ds \\
&\qquad + C_c[t^*]C_\lambda|x-y| + [c]_* C_{t^*} C_\lambda |x-y| \\
&\le \big(C_c C_{t^*}\big(C_\lambda A[t^*] + [\lambda]_1 + C_\lambda C_{t^*}[\lambda]_1\big) + C_c[t^*]C_\lambda + [c]_* C_{t^*} C_\lambda\big)|x-y|.
\end{aligned}$$

The result follows. □

The next lemma is stated without proof, as it is very close to [5, Lemma 51.7].

Lemma A.6. *For all* $x \in E$ *and* $t \in [0; t^*(x)]$, *one has*

$$v_n(\Phi(x,t)) = e^{\Lambda(x,t)} G_t v_{n-1}(x) - \int_0^t l\circ\Phi(x,s)\,ds.$$

Proposition A.7. *For all $n \in \{0, 1, \dots, N\}$, one has $v_n \in \boldsymbol{L}_c(E)$ and*

$$
\begin{aligned}
C_{v_n} &\leq n\big(C_{t^*} C_l + C_c\big), \\
[v_n]_1 &\leq e^{C_{t^*} C_\lambda}\big(K(A, v_{n-1}) + n C_{t^*}[\lambda]_1 (C_{t^*} C_l + C_c)\big) + C_{t^*}[l]_1, \\
[v_n]_2 &\leq e^{C_{t^*} C_\lambda}\big(C_{t^*} C_l C_\lambda + 2C_l + C_\lambda C_c + (2n-1) C_\lambda (C_{t^*} C_l + C_c)\big) + C_l, \\
[v_n]_* &\leq [v_n]_1 + [t^*][v_n]_2, \\
[v_n] &\leq K(A, v_{n-1}),
\end{aligned}
$$

Proof. Recall that for $x \in E$, one has from Definition 3.3

$$
v_n(x) = G v_{n-1}(x) = \boldsymbol{E}_x\left[L(x, S_1)\right] + \boldsymbol{E}_x\left[C(x, S_1)\right] + \boldsymbol{E}_x[v_{n-1}(Z_1)].
$$

Thus, $C_{v_n} \leq C_{t^*} C_l + C_c + C_{v_{n-1}} \leq n\big(c_{t^*} C_l + C_c\big)$ by induction.

Let us now turn to $[v_n]_1$. Lemma A.6 yields

$$
\begin{aligned}
&|v_n(\Phi(x,t)) - v_n(\Phi(y,t))| \\
&\quad \leq |e^{\Lambda(x,t)} G_t v_{n-1}(x) - e^{\Lambda(y,t)} G_t v_{n-1}(y)| + \int_0^t \left|l \circ \Phi(x,s) - l \circ \Phi(y,s)\right| ds \\
&\quad \leq e^{\Lambda(x,t)}\left|G_t v_{n-1}(x) - G_t v_{n-1}(y)\right| + \left|G_t v_{n-1}(y)\right|\left|e^{\Lambda(x,t)} - e^{\Lambda(y,t)}\right| \\
&\qquad\qquad + C_{t^*}[l]_1 |x - y|.
\end{aligned}
$$

The result follows using Proposition A.5 and noticing that

$$
\begin{aligned}
\Lambda(x,t) &\leq C_{t^*} C_\lambda, \\
|G_t v_{n-1}(y)| &\leq C_{t^*} C_l + C_c + C_{v_{n-1}} \leq n(C_{t^*} C_l + C_c), \\
|e^{\Lambda(x,t)} - e^{\Lambda(y,t)}| &\leq e^{C_{t^*} C_\lambda} C_{t^*}[\lambda]_1 |x - y|.
\end{aligned}
$$

We now turn to $[v_n]_2$. For $x \in E$ and $s, t \in [0, t^*(x)]$ with $s \leq t$, one has

$$
\begin{aligned}
\left|v_n(\Phi(x,t)) - v_n(\Phi(x,s))\right| &\leq e^{\Lambda(x,t)}\left|G_t v_{n-1}(x) - G_s v_{n-1}(x)\right| \\
&\quad + \left|G_s v_{n-1}(x)\right|\left|e^{\Lambda(x,t)} - e^{\Lambda(x,s)}\right| + C_l |t - s|.
\end{aligned}
$$

Moreover, from (14), one has

$$
\begin{aligned}
&\left|G_t v_{n-1}(x) - G_s v_{n-1}(x)\right| \\
&\quad \le E_x\left[\left|F(x, S_1) + v_{n-1}(Z_1)\right| \mathbb{1}_{\{s \le S_1 < t\}}\right] \\
&\quad \le \left| e^{-\Lambda(x,t)} \int_0^t l(\Phi(x,u))\, du - e^{-\Lambda(x,s)} \int_0^s l(\Phi(x,u))\, du \right| \\
&\qquad\qquad + \int_s^t \left| l(\Phi(x,u)) e^{-\Lambda(x,u)} \right| du \\
&\qquad\qquad + \left| c \circ \Phi(x, t^*(x)) \right| \int_s^t \left| \delta^{\mathcal{A}}(x,u) \lambda \circ \Phi(x,u) e^{-\Lambda(x,u)} \right| du \\
&\qquad\qquad + \int_s^t \left| (\lambda Q v_{n-1}) \circ \Phi(x,u) e^{-\Lambda(x,u)} \right| du \\
&\quad \le \left(C_{t^*} C_l \left| e^{-\Lambda(x,t)} - e^{-\Lambda(x,s)} \right| + C_l |t-s| \right) + (C_l |t-s|) \\
&\qquad\qquad + (C_c C_\lambda |t-s|) + (C_\lambda C_{v_{n-1}} |t-s|)
\end{aligned}
$$

and

$$
|e^{\Lambda(x,t)} - e^{\Lambda(x,s)}| \le e^{C_{t^*} C_\lambda} C_\lambda |t-s|.
$$

Finally, the bound for $[v_n]$ is a direct consequence from Proposition A.5. □

Appendix B. Relaxed assumption on the running cost function

In this section, we consider the approximation applied to the time-augmented process so that the local characteristics are $\tilde{\Phi}$, $\tilde{\lambda}$ and $\tilde{Q}$ defined in Section 5.1. Moreover, we consider a function $l \in \boldsymbol{L}_c(\tilde{E})$ and we define $\tilde{l} \in B(\tilde{E})$ by

$$
\text{for all } \xi = (x,t) \in \tilde{E},\ \tilde{l}(\xi) = l(x,t) \mathbb{1}_{\{t \le t_f\}}.
$$

We intend to prove that the convergence of our approximation scheme, stated by Theorem 4.5, remains true if we choose $\tilde{l}$ as the running cost function even though it does not fulfill the required Lipschitz conditions, i.e., $\tilde{l} \notin \boldsymbol{L}_c(\tilde{E})$. Indeed, the Lipschitz continuity of l is used four times in the proof of the theorem, once in Proposition A.2, twice in Proposition A.5 (when bounding the difference of the Υ_1 terms and the one of the Υ_2 ones) and once in Proposition A.7 (when bounding $[v_n]_1$). In each case, the Lipschitz continuity of the running cost function l is used to bound a term of the form

$$
\int_s^{s'} \left| \tilde{l} \circ \tilde{\Phi}(\xi, u) - \tilde{l} \circ \tilde{\Phi}(\xi', u) \right| du \tag{15}
$$

for $\xi, \xi' \in \widetilde{E}$ and $s, s' \in [0; \tilde{t}^*(\xi) \wedge \tilde{t}^*(\xi')]$, or of the form

$$\int_s^{\tilde{t}^*(\xi)\wedge\tilde{t}^*(\xi')} \left|\tilde{l} \circ \widetilde{\Phi}(\xi, u) e^{-\widetilde{\Lambda}(\xi,u)} - \tilde{l} \circ \widetilde{\Phi}(\xi', u) e^{-\widetilde{\Lambda}(\xi',u)}\right| du \tag{16}$$

for $\xi, \xi' \in \widetilde{E}$ and $s \in [0; \tilde{t}^*(\xi) \wedge \tilde{t}^*(\xi')]$ and where we use the natural notation $\widetilde{\Lambda}(\xi, u) = \int_0^u \tilde{\lambda}(\widetilde{\Phi}(\xi, v))\, dv$. Concerning this second form, Equation (16), notice that

$$\begin{aligned}
&\int_s^{\tilde{t}^*(\xi)\wedge\tilde{t}^*(\xi')} \left|\tilde{l} \circ \widetilde{\Phi}(\xi, u) e^{-\widetilde{\Lambda}(\xi,u)} - \tilde{l} \circ \widetilde{\Phi}(\xi', u) e^{-\widetilde{\Lambda}(\xi',u)}\right| du \\
&\quad \le \int_s^{\tilde{t}^*(\xi)\wedge\tilde{t}^*(\xi')} \left|\tilde{l} \circ \widetilde{\Phi}(\xi, u) - \tilde{l} \circ \widetilde{\Phi}(\xi', u)\right| du \\
&\qquad\qquad + C_l \int_s^{\tilde{t}^*(\xi)\wedge\tilde{t}^*(\xi')} \left|e^{-\widetilde{\Lambda}(\xi,u)} - e^{-\widetilde{\Lambda}(\xi',u)}\right| du \\
&\quad \le \int_s^{\tilde{t}^*(\xi)\wedge\tilde{t}^*(\xi')} \left|\tilde{l} \circ \widetilde{\Phi}(\xi, u) - \tilde{l} \circ \widetilde{\Phi}(\xi', u)\right| du + C_l C_{t^*}^2 [\lambda]_1 |\xi - \xi'|,
\end{aligned}$$

so that, to ensure that Theorem 4.5 remains true with $\tilde{l}$ as the running cost function, it is sufficient to be able to bound terms of the form (15). This is done in the following lemma.

Lemma B.1. *For $\xi = (x, t), \xi' = (x', t') \in \widetilde{E}$ and $s \in [0; \tilde{t}^*(\xi) \wedge \tilde{t}^*(\xi')]$, one has*

$$\int_0^s \left|\tilde{l} \circ \widetilde{\Phi}(\xi, u) - \tilde{l} \circ \widetilde{\Phi}(\xi', u)\right| du \le (C_{t^*}[l]_1 + C_l)|\xi - \xi'|.$$

Proof. Let $\xi = (x, t), \xi' = (x', t') \in \widetilde{E}$ and $s \in [0; \tilde{t}^*(\xi) \wedge \tilde{t}^*(\xi')]$. One has

$$\begin{aligned}
&\int_0^s \left|\tilde{l} \circ \widetilde{\Phi}(\xi, u) - \tilde{l} \circ \widetilde{\Phi}(\xi', u)\right| du \\
&\quad \le \int_0^s \left|l \circ \widetilde{\Phi}(\xi, u) \mathbb{1}_{\{t+u\le t_f\}} - l \circ \widetilde{\Phi}(\xi', u) \mathbb{1}_{\{t'+u\le t_f\}}\right| du \\
&\quad \le \int_0^s \left|l \circ \widetilde{\Phi}(\xi, u) - l \circ \widetilde{\Phi}(\xi', u)\right| du + C_l \int_0^s \left|\mathbb{1}_{\{t+u\le t_f\}} - \mathbb{1}_{\{t'+u\le t_f\}}\right| du
\end{aligned}$$

The left-hand side term is bounded by $C_{t^*}[l]_1|\xi - \xi'|$ since $l \in \boldsymbol{L}_c(\widetilde{E})$. For the right-hand side term, assume without loss of generality that $t \le t'$, one has

$$\left|\mathbb{1}_{\{t+u\le t_f\}} - \mathbb{1}_{\{t'+u\le t_f\}}\right| = \left|\mathbb{1}_{\{t-t_f\le u\}} - \mathbb{1}_{\{t'-t_f\le u\}}\right| = \mathbb{1}_{\{t-t_f\le u<t'-t_f\}},$$

so that the right-hand side term is bounded by $C_l|t - t'| \le C_l|\xi - \xi'|$. The result follows. □

Theorem 4.5 remains true if we choose $\tilde{l}$ as the running cost function. One only needs to slightly modify the Lipschitz constants given in propositions A.2, A.5 and A.7. The terms $C_{t^*}[l]_1$ have to be replaced by $C_{t^*}[l]_1 + C_l$.

Appendix C. Proof of Theorem 4.5

The Lipschitz continuity of the functions v_k is proved by Proposition A.7. Now let $A > 0$ and notice that

$$|J_N(l,c)(x) - \widehat{V}_0| \le |J_N(l,c)(x) - V_0| + |V_0 - \widehat{V}_0|.$$

Proposition 3.2 says that $|J_N(l,c)(x) - V_0| \le NC_cC_\lambda/A$ since $V_0 = J_N^A(l,c)(x)$. We now have to bound $|V_0 - \widehat{V}_0|$.

Some of the arguments of the proof are similar to the ones used in Theorem 5.1 from [6], thus we will not develop the details of the proof. Recall that $\|V_N - \widehat{V}_N\|_p = 0$ and let $k \in \{0, \dots, N-1\}$. In order to bound the approximation error, let us split it into three terms $\|V_k - \widehat{V}_k\|_p \le \Xi_1 + \Xi_2 + \Xi_3$, where

$$\begin{aligned}
\Xi_1 &= \|v_k(Z_k) - v_k(\widehat{Z}_k)\|_p, \\
\Xi_2 &= \|Gv_{k+1}(\widehat{Z}_k) - \widehat{G}_{k+1}v_{k+1}(\widehat{Z}_k)\|_p, \\
\Xi_3 &= \|\widehat{G}_{k+1}v_{k+1}(\widehat{Z}_k) - \widehat{G}_{k+1}\hat{v}_{k+1}(\widehat{Z}_k)\|_p.
\end{aligned}$$

The theorem is then a direct consequence from the three following lemmas, stated without proof, that provide bounds for each of these three terms.

Lemma C.1. *The first term, Ξ_1, is bounded by*

$$\|v_k(Z_k) - v_k(\widehat{Z}_k)\|_p \le [v_k]\|Z_k - \widehat{Z}_k\|_p.$$

Lemma C.2. *The second term, Ξ_2, is bounded by*

$$\begin{aligned}
&\left\|Gv_{k+1}(\widehat{Z}_k) - \widehat{G}_{k+1}v_{k+1}(\widehat{Z}_k)\right\|_p \\
&\quad \le [v_{k+1}]\|Z_{k+1} - \widehat{Z}_{k+1}\|_p + \big([v_k] + [F]_1\big)\|Z_k - \widehat{Z}_k\|_p + [F]_2\|S_{k+1} - \widehat{S}_{k+1}\|_p.
\end{aligned}$$

Lemma C.3. *The third term, Ξ_3, is bounded by*

$$\begin{aligned}
&\left\|\widehat{G}_{k+1}v_{k+1}(\widehat{Z}_k) - \widehat{G}_{k+1}\hat{v}_{k+1}(\widehat{Z}_k)\right\|_p \\
&\qquad\qquad \le [v_{k+1}]\|Z_{k+1} - \widehat{Z}_{k+1}\|_p + \|V_{k+1} - \widehat{V}_{k+1}\|_p.
\end{aligned}$$

Acknowledgements

This work was supported by ARPEGE program of the French National Agency of Research (ANR), project "FAUTOCOES", number ANR-09-SEGI-004. Besides, the authors gratefully acknowledge Astrium for its financial support.

References

[1] V. Bally and G. Pagès, *A quantization algorithm for solving multi-dimensional discrete-time optimal stopping problems*, Bernoulli **9** (2003), no. 6, 1003–1049. MR 2005f:60096

[2] V. Bally, G. Pagès, and J. Printems, *A quantization tree method for pricing and hedging multidimensional American options*, Math. Finance **15** (2005), no. 1, 119–168. MR 2005k:91142

[3] A. Brandejsky, B. de Saporta, and F. Dufour, *Numerical methods for the exit time of a piecewise-deterministic markov process*, Adv. Appl. Probab. **44** (2012), no. 1. arXiv 1012.2659

[4] C. Cocozza-Thivent, R. Eymard, and S. Mercier, *A finite-volume scheme for dynamic reliability models*, IMA J. Numer. Anal. **26** (2006), no. 3, 446–471. MR 2008h:60370 Zbl 1109.65011

[5] M. H. A. Davis, *Markov models and optimization*, Monographs on Statistics and Applied Probability, no. 49, Chapman & Hall, London, 1993. MR 96b:90002 Zbl 0780.60002

[6] B. de Saporta, F. Dufour, and K. Gonzalez, *Numerical method for optimal stopping of piecewise deterministic Markov processes*, Ann. Appl. Probab. **20** (2010), no. 5, 1607–1637. MR 2011g:60079 Zbl 1197.93162

[7] R. Eymard, S. Mercier, and A. Prignet, *An implicit finite volume scheme for a scalar hyperbolic problem with measure data related to piecewise deterministic Markov processes*, J. Comput. Appl. Math. **222** (2008), no. 2, 293–323. MR 2009j:65194 Zbl 1158.65008

[8] R. M. Gray and D. L. Neuhoff, *Quantization*, IEEE Trans. Inform. Theory **44** (1998), no. 6, 2325–2383. MR 99i:94029 Zbl 1016.94016

[9] G. Pagès, H. Pham, and J. Printems, *Optimal quantization methods and applications to numerical problems in finance*, Handbook of computational and numerical methods in finance, Birkhäuser, Boston, 2004, pp. 253–297. MR 2083055 Zbl 1138.91467

Received May 6, 2011. Revised November 8, 2011.

ADRIEN BRANDEJSKY: adrien.brandejsky@math.u-bordeaux1.fr
INRIA Bordeaux Sud-Ouest, team CQFD, F-33400 Talence, France

and

Université de Bordeaux, IMB, UMR 5251, F-33400 Talence, France

and

CNRS, IMB, UMR 5251, F-33400 Talence, France

BENOÎTE DE SAPORTA: saporta@math.u-bordeaux1.fr
Université de Bordeaux, Gretha, UMR 5113, F-33400 Talence, France

and

CNRS, Gretha, UMR 5113, IMB, UMR 5251, F-33400 Talence, France

and

INRIA Bordeaux Sud-Ouest, team CQFD, F-33400 Talence, France

FRANÇOIS DUFOUR: francois.dufour@math.u-bordeaux1.fr
Université de Bordeaux, IMB, UMR 5251, F-33400 Talence, France

and

CNRS, IMB, UMR 5251, F-33400 Talence, France

and

INRIA Bordeaux Sud-Ouest, team CQFD, F-33400 Talence, France

Communications in Applied Mathematics and Computational Science

map.berkeley.edu/camcos

See inside back cover or msp.berkeley.edu/camcos for submission instructions.

The subscription price for 2012 is US $75/year for the electronic version, and $105/year for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to Mathematical Sciences Publishers, Department of Mathematics, University of California, Berkeley, CA 94720-3840, USA.

Communications in Applied Mathematics and Computational Science, at Mathematical Sciences Publishers, Department of Mathematics, University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

CAMCoS peer review and production are managed by EditFLOW™ from Mathematical Sciences Publishers.

PUBLISHED BY
mathematical sciences publishers
http://msp.org/
A NON-PROFIT CORPORATION
Typeset in LaTeX

Communications in Applied Mathematics and Computational Science

vol. 7 no. 1 2012