Noname manuscript No.
(will be inserted by the editor)

Optimal control of semi-Markov processes with a backward stochastic differential equations approach

Elena Bandini · Fulvia Confortola

Received: date / Accepted: date

Abstract In the present work we employ backward stochastic differential equations (BSDEs) to study the optimal control problem of semi-Markov processes on a finite horizon, with general state and action spaces. More precisely, we prove that the value function and the optimal control law can be represented by means of the solution of a class of BSDEs driven by a semi-Markov process or, equivalently, by the associated random measure. We also introduce a suitable Hamilton-Jacobi-Bellman (HJB) equation. With respect to the pure jump Markov framework, the HJB equation in the semi-Markov case is characterized by an additional differential term ∂_a. Taking into account the particular structure of semi-Markov processes we rewrite the HJB equation in a suitable integral form which involves a directional derivative operator D related to ∂_a. Then, using a formula of Itô type tailor-made for semi-Markov processes and the operator D, we are able to prove that a BSDE of the above mentioned type provides the unique classical solution to the HJB equation, which identifies the value function of our control problem.

Keywords: Backward stochastic differential equations, optimal control problems, semi-Markov processes, marked point processes.

1 Introduction

The topic of optimal control of jump processes has been treated in many papers due to the many applications in queueing theory and other areas in

Elena Bandini
LUISS Roma, Dipartimento di Economia e Finanza, via Romania 32, 00197 Roma, Italy.
E-mail: elena.bandini@luiss.it

Fulvia Confortola
Politecnico di Milano, Dipartimento di Matematica, piazza Leonardo da Vinci 32, 20133 Milano, Italy.
E-mail: fulvia.confortola@polimi.it

engineering, see [26], [11], [45], [43], [10], and [30] for a recent reference. An increasing number of economic applications has been recently considered, see for instance [29], [4], [6].

Several approaches have been proposed to solve optimal control problems for jump processes. In the case of a Markov jump process, the optimality conditions can be derived via the corresponding infinitesimal generator, see for instance [34], [43], [48]. For a general jump process that is not necessarily Markovian, the so called martingale approach can be used to derive the optimality conditions by exploiting the Doob-Meyer decomposition of martingales, see [10], [11].

Recently in [14] the optimal control problem of a Markovian pure jump process has been solved by means of a suitable class of backward stochastic differential equations (BSDEs), driven by the random measure associated to the process itself. In this work the authors extend to the jump framework the method exhaustively used when the controlled process is a diffusion driven by a Brownian process. The BSDE approach is rather simple and is useful by a computational point of view, since the solution to the backward equation admits numerical approximations, see for instance the recent works [37], [38].

In this paper we aim at applying the BSDE approach to solve optimal control problems for semi-Markov pure jump processes. A semi-Markov process can be seen as a two-dimensional Markov jump process (X, θ), where, roughly speaking, the component X is pure jump and the component θ has also a deterministic motion between jumps. A common approach to tackle this problem is to characterize the value function as a solution to the Hamilton-Jacobi-Bellman (HJB) equation associated with an imbedded discrete-stage Markov decision model, with the stages defined by the jump times T_n of the process. In this case the decision is to find, at each stage, a control function that solves an imbedded deterministic optimal control problem. Usually the control strategy is chosen among the set of open loop policies, that is, stochastic kernels or measurable functions that depend only on the last jump time and post jump location. We can mention [1], [5], [16], [18], [27], [53] as works following this technique. Another important approach for this class of problems, which is sometimes called the infinitesimal approach, is to characterize the optimal value function as the viscosity solution of the corresponding integro-differential HJB equation. A sample of works using this kind of approach for more general jump processes is for instance [20], [19], [23], [24], [54]. Our results follow the infinitesimal approach, and provide sufficient conditions for the existence of a solution for an integro-differential HJB equation associated with the problem. In particular, exploiting the structure of the semi-Markov processes, we are able to prove that the integro-differential HJB has a classical solution. This solution is then shown to be unique and to coincide with the optimal value for the problem. The main novelty of our work consists in the fact that the characterization of the optimality is obtained using a suitable BSDE and its relation to the HJB equation.

Let us briefly describe our framework. Our starting point is a semi-Markov pure jump process X on a general state space K. It is constructed starting from a jump rate function $\lambda(x,a)$ and a jump measure $A \mapsto \bar{q}(x,a,A)$ on K, depending on $x \in K$ and $a \geq 0$. Our approach is to consider a semi-Markov pure jump process as a two dimensional time-homogeneous and strong Markov process $\{(X_s,\theta_s),\ s \geq 0\}$ with its natural filtration $\mathcal{F}$ and a family of probabilities $\mathbb{P}^{x,a}$ for $x \in K$, $a \in [0,\infty)$ such that $\mathbb{P}^{x,a}(X_0 = x, \theta_0 = a) = 1$. If the process starts from (x,a) at time $t = 0$ then the distribution of its first jump time T_1 under $\mathbb{P}^{x,a}$ is described by the formula

$$\mathbb{P}^{x,a}(T_1 > s) = \exp\left(-\int_a^{a+s} \lambda(x,r)\,dr\right), \tag{1}$$

and the conditional probability that the process is in A immediately after a jump at time $T_1 = s$ is

$$\mathbb{P}^{x,a}(X_{T_1} \in A \,|\, T_1 = s) = \bar{q}(x,s,A).$$

X_s is called the state of the process at time s, and θ_s is the duration period in this state up to moment s:

$$\theta_s = \begin{cases} a+s & \text{if } X_p = X_s \quad \forall\, 0 \leqslant p \leqslant s,\ \ p,s \in \mathbb{R}, \\ s - \sup\{\, p:\ 0 \leqslant p \leqslant s,\ X_p \neq X_s\} & \text{otherwise.} \end{cases}$$

We note that X alone is not a Markov process, while the two dimensional process (X,θ) is Markov but not pure jump. We limit ourselves to the case of a semi-Markov process X such that the survivor function of T_1 under $\mathbb{P}^{x,0}$ is absolutely continuous and admits a hazard rate function λ as in (1). The holding times of the process are not necessarily exponentially distributed and can be infinite with positive probability. Our main restriction is that the jump rate function λ is uniformly bounded, which implies that the process X is non explosive.

Roughly speaking, the Markov process (X,θ) presents random jumps and a deterministic motion between jumps. Denoting by T_n the jump times of X, one can consider the marked point process (T_n, X_{T_n}) and the associated random measure

$$p(dt\,dy) = \sum_n \delta_{(T_n, X_{T_n})}(dt\,dy)$$

on $(0,\infty) \times K$, where δ denotes the Dirac measure. The dual predictable projection $\tilde{p}$ of p (shortly, the compensator) has the following explicit expression

$$\tilde{p}(ds\,dy) = \lambda(X_{s-},\theta_{s-})\,\bar{q}(X_{s-},\theta_{s-},dy)\,ds.$$

Since the motion of (X,θ) is piecewise-deterministic, the random behaviour of the process turns out to be completely specified by λ and $\bar{q}$ or, equivalently, by the associated pure jump process, see for more details [9], [19].

In Section 3 we address an optimal intensity-control problem for the semi-Markov process X. This is formulated in a classical way by means of an absolutely continuous change of probability measure of Girsanov type, see e.g. [25], [26], [8], [11]. We define a class $\mathcal{A}$ of admissible control processes $(u_s)_{s\in[0,\,T]}$;

for every fixed $t \in [0, T]$ and $(x, a) \in K \times [0, \infty)$, the cost to be minimized and the corresponding value function are

$$\begin{aligned} J(t, x, a, u(\cdot)) &= \mathbb{E}^{x,a}_{u,t} \left[\int_0^{T-t} l(t+s, X_s, \theta_s, u_s)\, ds + g(X_{T-t}, \theta_{T-t}) \right], \\ v(t, x, a) &= \inf_{u(\cdot) \in \mathcal{A}} J(t, x, a, u(\cdot)), \end{aligned}$$

where g, l are given real functions. Here $\mathbb{E}^{x,a}_{u,t}$ denotes the expectation with respect to another probability $\mathbb{P}^{x;a}_{u,t}$, depending on t and on the control process u and constructed in such a way that the compensator under $\mathbb{P}^{x;a}_{u,t}$ equals $r(t+s, X_{s-}, \theta_{s-}, y, u_s)\, \lambda(X_{s-}, \theta_{s-})\, \bar{q}(X_{s-}, \theta_{s-}, dy)\, ds$, for some function r given in advance as another datum of the control problem. Since the process $(X_s, \theta_s)_{s \geq 0}$ we want to control is time-homogeneous and starts from (x, a) at time $s = 0$, we introduce a temporal translation which allows us to define the cost functional for all $t \in [0, T]$. For more details see Remark 4.

As we have already mentioned, the optimal control problem for semi-Markov processes is a classical topic in the literature. In particular, the case of a finite number of states has been studied in [12], [31], [33], [42], while the case of arbitrary state space is considered for instance in [46], [48], [10]. We remark that, comparing with [48], the controlled processes we deal with have laws absolutely continuous with respect to a given, uncontrolled process; see also a more detailed comment in Remark 5 below. It is worth noting that in [48] are considered only controls which give rise to Markov processes with stationary transition probabilities. In all the above mentioned works the authors use the dynamic programming method, and elements of the martingale theory in the way developed by [21], [49] and [26] for the Brownian motion case. Finally, notice that the semi-Markov processes belong to the larger class of piecewise-deterministic Markov processes (PDMPs) introduced in [19]. Optimal control problems for PDMPs have been considered by several authors, see for instance [20], [51], [22], [40], and the book [16] for a recent reference.

Differently to the related literature, our approach to this control problem is based on the backward stochastic differential equations (BSDEs) theory.

Backward equations driven by random measures have been studied in many papers, among which [50], [3], [47], [39], [52], and more recently in [7], [13], [14], [17], [35], [36]. In the most part of the cases, the stochastic equations are driven by a Wiener process and a Poisson random measure, see, e.g., [50], [3], [47], [39]. In this framework the compensator of the random measure is deterministic. In [7], [17], [35], [36], the authors deal with BSDEs driven by random measures more general than the Poisson one; however, the corresponding compensators are absolutely continuous with respect to a deterministic measure, and the problem can therefore be reduced to the Poisson case by a Girsanov change of probability. BSDEs driven by random measures that are not dominated by some fixed deterministic measure have been recently studied in [15], [13] and [14] respectively in the non-Markov and in the Markov pure jump case. Another well-posedness result for general BSDEs driven by non-dominated random measures has been recently given in [2]; here, unlike the previous literature, the random measure driving the BSDE can admit stochastic jumps.

The BSDE we consider is driven by the compensated random measure $q(dt\,dy) = p(dt\,dy) - \tilde{p}(dt\,dy)$ associated with the two dimensional Markov process (X, θ), where the compensator $\tilde{p}(dt\,dy) = \lambda(X_{s-}, \theta_{s-})\,\bar{q}(X_{s-}, \theta_{s-}, dy)$ ds is a stochastic random measure with a non-dominated intensity as in [14]. In particular, we introduce a family of BSDEs parametrized by $(t, x, a) \in [0, T] \times K \times [0, \infty)$:

$$\begin{aligned} Y^{x,a}_{s,t} + \int_s^{T-t} \int_K Z^{x,a}_{\sigma,t}(y)\, q(d\sigma\, dy) = g(X_{T-t}, \theta_{T-t}) \\ + \int_s^{T-t} f\Big(t+\sigma, X_\sigma, \theta_\sigma, Z^{x,a}_{\sigma,t}(\cdot)\Big)\, d\sigma, \quad s \in [0,\, T-t]. \end{aligned} \tag{2}$$

where the generator is given by the Hamiltonian function f defined for every $s \in [0,\, T]$, $(x, a) \in K \times [0,\, +\infty)$, $z \in L^2(K, \mathcal{K}, \lambda(x, a)\bar{q}(x, a, dy))$, as

$$\begin{aligned} &f(s, x, a, z(\cdot)) \\ &= \inf_{u \in U} \Big\{\, l(s, x, a, u) + \int_K z(y)(r(s, x, a, y, u) - 1)\lambda(x, a)\bar{q}(x, a, dy) \,\Big\}. \end{aligned} \tag{3}$$

Even if the process (X, θ) is not pure jump, its jump mechanism is completely specified by the random measure $\lambda(x, a)\,\bar{q}(x, a, dy)$ appearing in the BSDE. Therefore the existence, uniqueness and continuous dependence on the data for the BSDE (2) can be deduced extending in a straightforward way the results in [14], see Remark 3. Then, under appropriate assumptions, we are able to prove that the optimal control problem has a solution and that the value function and the optimal control can be represented by means of the solution to the BSDE (2). The BSDE approach to optimal control is well-known in the diffusive context, while few results are available in the non-diffusive context, see [15], [14], [13]. In particular, it seems to us be pursued here for the first time in the case of the semi-Markov processes. It allows us to treat in a unified way a large class of control problems, where the state space is general and the running and final cost are not necessarily bounded.

In Section 4 we solve a nonlinear variant of the Kolmogorov equation for the process (X, θ), with the BSDEs approach. The process (X, θ) is time-homogeneous and Markov, but is not a pure jump process. In particular it has the integro-differential infinitesimal generator

$$\tilde{\mathcal{L}}\Phi(x, a) := \partial_a \Phi(x, a) + \int_K [\Phi(y, 0) - \Phi(x, a)]\, \lambda(x, a)\, \bar{q}(x, a, dy),$$

for $(x, a) \in K \times [0, \infty)$. The additional differential term ∂_a forbid to study the associated nonlinear Kolmogorov equation proceeding as in the pure jump Markov processes framework, see [14]. Nevertheless, taking into account the particular structure of semi-Markov processes, we provide a reformulation of the Kolmogorov equation which allows us to consider solutions in a classical sense. Indeed, we notice that the second component of the process $(X_s,\, \theta_s)_{s\geqslant 0}$ is linear in s. This fact suggests to introduce the formal directional derivative operator

$$(Dv)(t,x,a) := \lim_{h\downarrow 0} \frac{v(t+h,x,a+h)-v(t,x,a)}{h},$$

and to consider the following nonlinear Kolmogorov equation

$$\begin{cases} Dv(t,x,a)+\mathcal{L}v(t,x,a)+f(t,x,a,v(t,x,a),v(t,\cdot,0)-v(t,x,a))=0, \\ \qquad\qquad\qquad\qquad\qquad\qquad t\in[0,T],\, x\in K,\, a\in[0,\infty), \\ v(T,x,a)=g(x,a), \end{cases} \tag{4}$$

where

$$\mathcal{L}\Phi(x,a) := \int_K [\Phi(y,0)-\Phi(x,a)]\,\lambda(x,a)\,\bar q(x,a,dy), \qquad (x,a)\in K\times[0,\infty).$$

Then we look for a solution v such that the map $t\mapsto v(t,x,t+c)$ is absolutely continuous on $[0,T]$, for all constants $c\in[-T,\,+\infty)$. The functions f,g in (4) are given. While it is easy to prove well-posedness of (4) under boundedness assumptions, we achieve the purpose of finding a unique solution under much weaker conditions related to the distribution of the process (X,θ): see Theorem 7. To this end we need to define a formula of Itô type, involving the directional derivative operator D, for the composition of the process $(X_s,\,\theta_s)_{s\geqslant 0}$ with functions v smooth enough (see Lemma 3 below).

The solution v of (4) is constructed by means of a family of BSDEs of the form (2). By the results above there exists a unique solution $(Y^{x,a}_{s,t},\, Z^{x,a}_{s,t})_{s\in[0,\,T-t]}$ and the estimates on the BSDEs are used to prove well-posedness of (4). As a by-product we also obtain the representation formulae

$$\begin{aligned} &v(t,x,a)=Y^{x,a}_{0,t},\\ &Y^{x,a}_{s,t}=v(t+s,X_s,\theta_s), \qquad Z^{x,a}_{s,t}(y)=v(t+s,y,0)-v(t+s,X_{s-},\theta_{s-}), \end{aligned}$$

which are sometimes called, at least in the diffusive case, non linear Feynman-Kac formulae.

Finally we can go back to the original control problem and observe that the associated Hamilton-Jacobi-Bellman equation has the form (4) where f is the Hamiltonian function (3). By previous results we are able to identify the HJB solution $v(t,x,a)$, constructed probabilistically via BSDEs, with the value function.

2 Notation, preliminaries and basic assumptions

2.1 Semi-Markov jump processes

We recall the definition of a semi-Markov process, as given, for instance, in [28]. More precisely we will deal with a semi-Markov process with infinite lifetime (i.e. non explosive).

Suppose we are given a measurable space $(K, \mathcal{K})$, a set Ω and two functions $X : \Omega \times [0, \infty) \to K$, $\theta : \Omega \times [0, \infty) \to [0, \infty)$. For every $t \geq 0$, we denote by $\mathcal{F}_t$ the σ-algebra $\sigma((X_s, \theta_s),\ s \in [0, t])$. We suppose that for every $x \in K$ and $a \in [0, \infty)$, a probability $\mathbb{P}^{x,a}$ is given on $(\Omega, \mathcal{F}_{[0,\infty)})$ and the following conditions hold.

1. $\mathcal{K}$ contains all one-point sets. Δ denotes a point not included in K.
2. $\mathbb{P}^{x,a}(X_0 = x, \theta_0 = a) = 1$ for every $x \in K$, $a \in [0, \infty)$.
3. For every s, $p \geqslant 0$ and $A \in \mathcal{K}$ the function $(x,\, a) \mapsto \mathbb{P}^{x,a}(X_s \in A,\, \theta_s \leqslant p)$ is $\mathcal{K} \otimes \mathcal{B}^+$-measurable.
4. For every $0 \leq t \leq s$, $p \geqslant 0$, and $A \in \mathcal{K}$ we have
$$\mathbb{P}^{x,a}(X_s \in A,\, \theta_s \leqslant p \,|\, \mathcal{F}_t) = \mathbb{P}^{X_t,\theta_t}(X_s \in A,\, \theta_s \leqslant p), \quad \mathbb{P}^{x,a}\text{a.s.}$$
5. All the trajectories of the process X have right limits when K is given the discrete topology (the one where all subsets are open). This is equivalent to require that for every $\omega \in \Omega$ and $t \geq 0$ there exists $\delta > 0$ such that $X_s(\omega) = X_t(\omega)$ for $s \in [t, t + \delta]$.
6. All the trajectories of the process θ are continuous from the right piecewise linear functions. For every $\omega \in \Omega$, if $[\alpha, \beta)$ is the interval of linearity of $\theta_\cdot(\omega)$ then $\theta_s(\omega) = \theta_\alpha(\omega) + s - \alpha$ and $X_\alpha(\omega) = X_s(\omega)$; if β is a discontinuity point of $\theta_\cdot(\omega)$ then $\theta_{\beta+}(\omega) = 0$ and $X_\beta(\omega) \neq X_{\beta-}(\omega)$.
7. For every $\omega \in \Omega$ the number of jumps of the trajectory $t \mapsto X_t(\omega)$ is finite on every bounded interval.

X_s is called the *state* of the process at time s, θ_s is the *duration period* in this state up to moment s. Also we call X_s the *phase* and θ_s the *age* or the *time component* of a semi-Markov process. X is a non explosive process because of condition 7. We note, moreover, that the two-dimensional process (X, θ) is a strong Markov process with time-homogeneous transition probabilities because of conditions $2, 3$, and 4. It has right-continuous sample paths because of conditions 1, 5 and 6, and it is not a pure jump Markov process, but only a PDMP (see Section 24 in [19]).

The class of semi-Markov processes we consider in the paper will be described by means of a special form of joint law Q under $\mathbb{P}^{x,a}$ of the first jump time T_1, and the corresponding position X_{T_1}. To proceed formally, we fix $X_0 = x \in K$ and define the first jump time

$$T_1 = \inf\{p > 0 :\ X_p \neq x\},$$

with the convention that $T_1 = +\infty$ if the indicated set is empty.
We introduce $S := K \times [0,\, +\infty)$ an we denote by $\mathcal{S}$ the smallest σ-algebra containing all sets of $\mathcal{K} \otimes \mathcal{B}([0,\, +\infty))$. (Here and in the following $\mathcal{B}(\Lambda)$ denotes the Borel σ-algebra of a topological space Λ). Take an extra point $\Delta \notin K$ and define $X_\infty(\omega) = \Delta$ for all $\omega \in \Omega$, so that $X_{T_1} : \Omega \to K \cup \{\Delta\}$ is well defined. Then on the extended space $S \cup \{(\Delta,\, \infty)\}$ we consider the smallest σ-algebra, denoted by $\mathcal{S}^{\text{enl}}$, containing $\{(\Delta,\, \infty)\}$ and all sets of $\mathcal{K} \otimes \mathcal{B}([0,\, +\infty))$. Then

(X_{T_1}, T_1) is a random variable with values in $(S \cup \{(\Delta, \infty)\}, \mathcal{S}^{\mathrm{enl}})$. Its law under $\mathbb{P}^{x,a}$ will be denoted by $Q(x, a, \cdot)$.

We will assume that Q is constructed from two given functions denoted by λ and $\bar{q}$. More precisely we assume the following.

Hypothesis 1. *There exist two functions*

$$\lambda : S \to [0, \infty) \text{ and } \bar{q} : S \times \mathcal{K} \to [0, 1]$$

such that

(i) $(x, a) \mapsto \lambda(x, a)$ *is* $\mathcal{S}$*-measurable;*
(ii) $\sup_{(x,a)\in S} \lambda(x, a) \leqslant C \in \mathbb{R}^+$*;*
(iii) $(x, a) \mapsto \bar{q}(x, a, A)$ *is* $\mathcal{S}$*-measurable* $\forall A \in \mathcal{K}$*;*
(iv) $A \mapsto \bar{q}(x, a, A)$ *is a probability measure on* $\mathcal{K}$ *for all* $(x, a) \in S$*.*

We define a function H on $K \times [0, \infty]$ by

$$H(x, s) := 1 - e^{-\int_0^s \lambda(x,r)dr}. \tag{5}$$

Given λ and $\bar{q}$, we will require that for the semi-Markov process X we have, for every $(x, a) \in S$ and for $A \in \mathcal{K}$, $0 \le c < d \le \infty$,

$$\begin{aligned} Q(x, a, A \times (c, d)) &= \frac{1}{1 - H(x, a)} \int_c^d \bar{q}(x, s, A) \frac{d}{d\,s} H(x, a + s)\, ds \\ &= \int_c^d \bar{q}(x, s, A)\, \lambda(x, a + s)\, \exp\left(-\int_a^{a+s} \lambda(x, r)\, dr\right) ds, \end{aligned} \tag{6}$$

where Q was described above as the law of (X_{T_1}, T_1) under $\mathbb{P}^{x,a}$.

The existence of a semi-Markov process satisfying (6) is a well known fact, see for instance [48] Theorem 2.1, where it is proved that X is in addition a strong Markov process. The nonexplosive character of X is made possible by Hypothesis 1-(ii).

We note that our data only consist initially in a measurable space $(K, \mathcal{K})$ ($\mathcal{K}$ contains all singleton subsets of K), and in two functions λ, $\bar{q}$ satisfying Hypothesis 1. The semi-Markov process X can be constructed in an arbitrary way provided (6) holds.

Remark 1 1. Note that (6) completely specifies the probability measure $Q(x, a, \cdot)$ on $(S \cup \{(\Delta, \infty)\}, \mathcal{S}^{\mathrm{enl}})$: indeed simple computations show that, for $s \ge 0$,

$$\mathbb{P}^{x,a}(T_1 \in (s, \infty]) = 1 - Q(x, a, K \times (0, s]) = \exp\left(-\int_a^{a+s} \lambda(x, r)\, dr\right), \tag{7}$$

and we clearly have

$$\mathbb{P}^{x,a}(T_1 = \infty) = Q(x, a, \{(\Delta, \infty)\}) = \exp\left(-\int_a^{\infty} \lambda(x, r)\, dr\right).$$

Moreover, the kernel Q is well defined, since by assumption 1-(ii) $H(x, a) < 1$ for all $(x, a) \in S$.

2. The data λ and $\bar{q}$ have themselves a probabilistic interpretation. In fact if in (7) we set $a = 0$ we obtain

$$\mathbb{P}^{x,0}(T_1 > s) = \exp\left(-\int_0^s \lambda(x,r)\,dr\right) = 1 - H(x,s). \tag{8}$$

This means that under $\mathbb{P}^{x,0}$ the law of T_1 is described by the distribution function H, and

$$\lambda(x,a) = \frac{\frac{\partial H}{\partial a}(x,a)}{1 - H(x,a)}.$$

Then $\lambda(x,a)$ is the jump rate of the process X given that it has been in state x for a time a.
Moreover, the probability $\bar{q}(x,s,\cdot)$ can be interpreted as the conditional probability that X_{T_1} is in $A \in \mathcal{K}$ given that $T_1 = s$; more precisely,

$$\mathbb{P}^{x,a}(X_{T_1} \in A, T_1 < \infty \,|\, T_1) = \bar{q}(x, T_1, A)\, 1_{T_1 < \infty}, \qquad \mathbb{P}^{x,a} - a.s.$$

3. In [28] the following observation is made: starting from $T_0 = t$ define inductively $T_{n+1} = \inf\{s > T_n \,:\, X_s \neq X_{T_n}\}$, with the convention that $T_{n+1} = \infty$ if the indicated set is empty; then, under the probability $\mathbb{P}^{x,a}$, the sequence of the successive states of the semi-Markov X is a Markov chain, as in the case of Markov processes. However, while for the latter the duration period in the state depends only on this state and it is necessarily exponentially distributed, in the case of a semi Markov process the duration period depends also on the state into which the process moves and the distribution of the duration period may be arbitrary.
4. In [28] is also proved that the sequence $(X_{T_n}, T_n)_{n\geq 0}$ is a discrete-time Markov process in $(S \cup \{(\varDelta, \infty)\}, \mathcal{S}^{\mathrm{enl}})$ with transition kernel Q, provided we extend the definition of Q making the state $(\varDelta, \infty)$ absorbing, i.e. we define

$$Q(\varDelta, \infty, S) = 0, \qquad Q(\varDelta, \infty, \{(\varDelta, \infty)\}) = 1.$$

Note that $(X_{T_n}, T_n)_{n\geq 0}$ is time-homogeneous. This fact allows for a simple description of the process X. Suppose one starts with a discrete-time Markov process $(\tau_n, \xi_n)_{n\geq 0}$ in S with transition probability kernel Q and a given starting point $(x,a) \in S$ (conceptually, trajectories of such a process are easy to simulate). One can then define a process Y in K setting $Y_t = \sum_{n=0}^{N} \xi_n 1_{[\tau_n, \tau_{n+1})}(t)$, where $N = \sup\{n \geq 0 \,:\, \tau_n \leqslant \infty\}$. Then Y has the same law as the process X under $\mathbb{P}^{x,a}$.
5. We stress that (5) limits ourselves to deal with a class of semi-Markov processes for which the survivor function T_1 under $\mathbb{P}^{x,0}$ admits a hazard rate function λ.

2.2 BSDEs driven by a semi-Markov process

Let be given a measurable space $(K, \mathcal{K})$, a transition measure $\bar{q}$ on K and a given positive function λ, satisfying Hypothesis 1. Let X be the associated semi-Markov process constructed out of them as described in Section 2.1. We fix a deterministic terminal time $T > 0$ and a pair $(x, a) \in S$, and we look at all processes under the probability $\mathbb{P}^{x,a}$. We denote by $\mathcal{F}$ the natural filtration $(\mathcal{F}_t)_{t\in[0,\infty)}$ of X. Conditions 1, 5 and 6 above imply that the filtration $\mathcal{F}$ is right continuous (see [11], Appendix A2, Theorem T26). The predictable σ-algebra (respectively, the progressive σ-algebra) on $\Omega \times [0, \infty)$ is denoted by $\mathcal{P}$ (respectively, by $Prog$). The same symbols also denote the restriction to $\Omega \times [0, T]$.

We define a sequence $(T_n)_{n\geqslant 1}$ of random variables with values in $[0, \infty]$, setting

$$T_0(\omega) = 0, \quad T_{n+1}(\omega) = \inf\{s \geqslant T_n(\omega) : X_s(\omega) \neq X_{T_n}(\omega)\}, \tag{9}$$

with the convention that $T_{n+1}(\omega) = \infty$ if the indicated set is empty. Being X a jump process we have $T_n(\omega) \leqslant T_{n+1}(\omega)$ if $T_{n+1}(\omega) < \infty$, while the non explosion of X means that $T_{n+1}(\omega) \to \infty$ as n goes to infinity. We stress the fact that $(T_n)_{n\geqslant 1}$ coincide by definition with the time jumps of the two dimensional process (X, θ).

For $\omega \in \Omega$ we define a random measure on $([0, \infty) \times K, \mathcal{B}[0, \infty) \otimes \mathcal{K})$ setting

$$p(\omega, C) = \sum_{n\geqslant 1} \mathbb{1}_{\{(T_n(\omega),\, X_{T_n}(\omega))\in C\}}, \qquad C \in \mathcal{B}[0, \infty) \otimes \mathcal{K}. \tag{10}$$

The random measure $\lambda(X_{s-}, \theta_{s-})\, \bar{q}(X_{s-}, \theta_{s-}, dy)\, ds$ is called the compensator, or the dual predictable projection, of $p(ds, dy)$. We are interested in the following family of backward equations driven by the compensated random measure $q(ds\, dy) = p(ds\, dy) - \lambda(X_{s-}, \theta_{s-})\, \bar{q}(X_{s-}, \theta_{s-}, dy)\, ds$ and parametrized by (x, a): $\mathbb{P}^{x,a}$-a.s.,

$$Y_s + \int_s^T \int_K Z_r(y)\, q(dr\, dy) = g(X_T, \theta_T) + \int_s^T f\Big(r, X_r, \theta_r, Y_r, Z_r(\cdot)\Big)\, dr, \tag{11}$$

for $s \in [0, T]$. We consider the following assumptions on the data f and g.

Hypothesis 2. (1) *The final condition* $g : S \to \mathbb{R}$ *is* $\mathcal{S}$*-measurable and* $\mathbb{E}^{x,a}\Big[|g(X_T, \theta_T)|^2\Big] < \infty$.

(2) *The generator* f *is such that*
 (i) *for every* $s \in [0, T]$, $(x, a) \in S$, $r \in \mathbb{R}$, f *is a mapping* $f(s, x, a, r, \cdot) : \mathcal{L}^2(K, \mathcal{K}, \lambda(x, a)\, \bar{q}(x, a, dy)) \to \mathbb{R}$;
 (ii) *for every bounded and* $\mathcal{K}$*-measurable* $z : K \to \mathbb{R}$ *the mapping*

$$(s, x, a, r) \mapsto f(s, x, a, r, z(\cdot)) \tag{12}$$

is $\mathcal{B}([0, T]) \otimes \mathcal{S} \otimes \mathcal{B}(\mathbb{R})$*-measurable;*

(iii) *there exist* $L \geqslant 0$, $L' \geqslant 0$ *such that for every* $s \in [0, T]$, $(x, a) \in S$, $r, r' \in \mathbb{R}$, $z, z' \in \mathcal{L}^2(K, \mathcal{K}, \lambda(x, a)\, \bar{q}(x, a, dy))$ *we have*

$$\begin{aligned}
&|f(s, x, a, r, z(\cdot)) - f(s, x, a, r', z'(\cdot))| \\
&\leqslant L' \, |r - r'| + L \left(\int_K |z(y) - z'(y)|^2 \, \lambda(x, a)\, \bar{q}(x, a, dy) \right)^{1/2}; \qquad (13)
\end{aligned}$$

(iv) *we have*

$$\mathbb{E}^{x,a} \left[\int_0^T |f(s, X_s, \theta_s, 0, 0)|^2 \, ds \right] < \infty. \qquad (14)$$

Remark 2 Assumptions (i), (ii), and (iii) imply the following measurability properties of $f(s, X_s, \theta_s, Y_s, Z_s(\cdot))$:

- if $Z \in \mathcal{L}^2(p)$, then the mapping

$$(\omega, s, y) \mapsto f(s, X_{s-}(\omega), \theta_{s-}(\omega), y, Z_s(\omega, \cdot))$$

  is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R})$-measurable;
- if, in addition, Y is a $Prog$-measurable process, then

$$(\omega, s) \mapsto f(s, X_{s-}(\omega), \theta_{s-}(\omega), Y_s(\omega), Z_s(\omega, \cdot))$$

  is $Prog$-measurable.

We introduce the space $\mathbb{M}^{x,a}$ of the processes (Y, Z) on $[0, T]$ such that Y is real-valued and $Prog$-measurable, $Z : \Omega \times K \to \mathbb{R}$ is $\mathcal{P} \otimes \mathcal{K}$-measurable, and

$$\begin{aligned}
||(Y, Z)||^2_{\mathbb{M}^{x,a}} := &\, \mathbb{E}^{x,a} \left[\int_0^T |Y_s|^2 \, ds \right] \\
&+ \mathbb{E}^{x,a} \left[\int_0^T \int_K |Z_s(y)|^2 \, \lambda(X_s, \theta_s)\, \bar{q}(X_s, \theta_s, dy)\, ds \right] < \infty.
\end{aligned}$$

The space $\mathbb{M}^{x,a}$ endowed with this norm is a Banach space, provided we identify pairs of processes whose difference has norm zero.

Theorem 3 *Suppose that Hypothesis 2 holds for some* $(x, a) \in S$. *Then there exists a unique pair* (Y, Z) *in* $\mathbb{M}^{x,a}$ *which solves the BSDE* (11). *Let moreover* (Y', Z') *be another solution in* $\mathbb{M}^{x,a}$ *to the BSDE* (11) *associated with the driver* f' *and final datum* g'. *Then*

$$\begin{aligned}
&\sup_{s \in [0, T]} \mathbb{E}^{x,a} \left[|Y_s - Y'_s|^2 \right] + \mathbb{E}^{x,a} \left[\int_0^T |Y_s - Y'_s|^2 ds \right] \\
&+ \mathbb{E}^{x,a} \left[\int_0^T \int_K |Z_s(y) - Z'_s(y)|^2 \lambda(X_s, \theta_s)\, \bar{q}(X_s, \theta_s, dy)\, ds \right] \\
&\leqslant C \Bigg\{ \mathbb{E}^{x,a} \left[|g(X_T) - g'(X_T)|^2 \right]
\end{aligned}$$

$$+\,\mathbb{E}^{x,a}\left[\int_0^T |f(s,X_s,\theta_s,Y'_s,Z'_s(\cdot)) - f'(s,X_s,\theta_s,Y'_s,Z'_s(\cdot))|^2 ds\right]\Bigg\}, \qquad (15)$$

where C is a constant depending on T, L, L'.

Remark 3 The construction of a solution to the BSDE (11) is based on the integral representation theorem of marked point process martingales (see, e.g., [19]), and on a fixed-point argument. Similar results of well-posedness for BSDEs driven by random measures can be found in literature, see, in particular, the theorems given in [14], Section 3, and in [7]. Notice that these results can not be a priori straight applied to our framework: in [7] are involved random compensators which are absolutely continuous with respect to a deterministic measure, instead in our case the compensator is a stochastic random measure with a non-dominated intensity; [14] applies to BSDEs driven by a general random measure associated to a pure jump Markov process, while the two dimensional process (X,θ) is Markov but not pure jump. Nevertheless, BSDE (11) is driven by the marked point process associated to the semi-Markov process. Indeed (X,θ) is piecewise deterministic, and its random behaviour is completely specified by the associated underlying pure jump process, or, equivalently, by the associated compensator $\tilde{p}$ (see [19] and [32]). Therefore the well-posedness results for the BSDE obtained in [14] in the pure jump case can be extended to our framework without additional difficulties. In particular, the proof of Theorem 3 turns out to be very similar to those of Theorem 3.4 and Proposition 3.5 in [14], and we do not report it here to alleviate the presentation.

3 Optimal control

3.1 Formulation of the problem

In this section we consider again a measurable space $(K,\mathcal{K})$, a transition measure $\bar{q}$ and a function λ satisfying Hypothesis 1. The data specifying the optimal control problem we will address to are an action (or decision) space U, a running cost function l, a terminal cost function g, a (deterministic, finite) time horizon $T>0$ and another function r specifying the effect of the control process. We define an admissible control process, or simply a control, as a predictable process $(u_s)_{s\in[0,\,T]}$ with values in U. The set of admissible control processes is denoted by $\mathcal{A}$. We will make the following assumptions:

Hypothesis 4. (1) *$(U,\mathcal{U})$ is a measurable space.*

(2) *The function $r:[0,\,T]\times S\times K\times U\to\mathbb{R}$ is $\mathcal{B}([0,\,T])\otimes\mathcal{S}\otimes\mathcal{K}\otimes\mathcal{U}$-measurable and there exists a constant $C_r>1$ such that,*

$$0\leqslant r(t,x,a,y,u)\leqslant C_r, \qquad t\in[0,\,T],\,(x,a)\in S,\,y\in K,\,u\in U. \qquad (16)$$

(3) *The function $g:S\to\mathbb{R}$ is $\mathcal{S}$-measurable, and for all fixed $t\in[0,\,T]$,*

$$\mathbb{E}^{x,a}\left[|g(X_{T-t},\theta_{T-t})|^2\right]<\infty, \qquad \forall (x,a)\in S. \qquad (17)$$

(4) *The function* $l : [0, T] \times S \times U \to \mathbb{R}$ *is* $\mathcal{B}([0\ T]) \otimes \mathcal{S} \otimes \mathcal{U}$*-measurable and there exists* $\alpha > 1$ *such that, for every fixed* $t \in [0, T]$*, for every* $(x, a) \in S$ *and* $u(\cdot) \in \mathcal{A}$*,*

$$
\begin{array}{l}
\inf_{u\in U} l(t,x,a,u) > \infty; \\
\mathbb{E}^{x,a}\left[\int_0^{T-t} \left|\inf_{u\in U} l(t+s,X_s,\theta_s,u)\right|^2\, ds\right] < \infty, \\
\mathbb{E}^{x,a}\left[\int_0^{T-t} |l(t+s,X_s,\theta_s,u_s)|\, ds\right]^{\alpha} < \infty.
\end{array}
\tag{18}
$$

To any $(t, x, a) \in [0, T] \times S$ and any control $u(\cdot) \in \mathcal{A}$ we associate a probability measure $\mathbb{P}^{x,a}_{u,t}$ by a change of measure of Girsanov type, as we now describe. Recalling the definition of the jump times T_n in (9), we define, for every fixed $t \in [0, T]$,

$$
\begin{aligned}
L^t_s = \exp&\left(\int_0^s \int_K (1 - r(t+\sigma, X_\sigma, \theta_\sigma, y, u_\sigma))\, \lambda(X_\sigma, \theta_\sigma)\, \bar{q}(X_\sigma, \theta_\sigma, dy)\, d\sigma\right) \cdot \\
&\cdot \prod_{n\geqslant 1 : T_n \leqslant s} r(t+T_n, X_{T_n-}, \theta_{T_n-}, X_{T_n}, u_{T_n}),
\end{aligned}
$$

for all $s \in [0, T-t]$, with the convention that the last product equals 1 if there are no indices $n \geqslant 1$ satisfying $T_n \leqslant s$. As a consequence of the boundedness assumption on $\bar{q}$ and λ it can be proved, using for instance Lemma 4.2 in [13], or [11] Chapter VIII Theorem T11, that for every fixed $t \in [0, T]$ and for every $\gamma > 1$ we have

$$
\mathbb{E}^{x,a}\left[\left|L^t_{T-t}\right|^{\gamma}\right] < \infty, \qquad \mathbb{E}^{x,a}\left[L^t_{T-t}\right] = 1,
\tag{19}
$$

and therefore the process L^t is a martingale (relative to $\mathbb{P}^{x,a}$ and $\mathcal{F}$). Defining a probability $\mathbb{P}^{x,a}_{u,t}(d\omega) = L^t_{T-t}(\omega)\, \mathbb{P}^{x,a}(d\omega)$, we introduce the cost functional corresponding to $u(\cdot) \in \mathcal{A}$ as

$$
J(t,x,a,u(\cdot)) = \mathbb{E}^{x,a}_{u,t}\left[\int_0^{T-t} l(t+s,X_s,\theta_s,u_s)\, ds + g(X_{T-t},\theta_{T-t})\right],
\tag{20}
$$

where $\mathbb{E}^{x,a}_{u,t}$ denotes the expectation under $\mathbb{P}^{x,a}_{u,t}$. Taking into account (17), (18) and (19), and using Hölder inequality it is easily seen that the cost is finite for every admissible control. The control problem starting at (x, a) at time $s = 0$ with terminal time $s = T - t$ consists of minimizing $J(t, x, a, \cdot)$ over $\mathcal{A}$.

We finally introduce the value function

$$
v(t,x,a) = \inf_{u(\cdot)\in\mathcal{A}} J(t,x,a,u(\cdot)), \qquad t \in [0, T],\ (x,a) \in S.
$$

The previous formulation of the optimal control problem by means of change of probability measure is classical (see e.g. [25], [26], [11]). Some comments may be useful at this point.

Remark 4 1. The particular form of cost functional (20) is due to the fact that the time-homogeneous Markov process $(X_s, \theta_s)_{s \geqslant 0}$ satisfies

$$\mathbb{P}^{x,a}(X_0 = x,\, \theta_0 = a) = 1;$$

the introduction of the temporal translation in the first component allows us to define $J(t, x, a, u(\cdot))$ for all $t \in [0,\, T]$.

2. We recall (see e.g. [11], Appendix A2, Theorem T34) that a process u is $\mathcal{F}$-predictable if and only if it admits the representation

$$u_s(\omega) = \sum_{n \geqslant 0} u_s^{(n)}(\omega)\, \mathbb{1}_{(T_n(\omega), T_{n+1}(\omega)]}(s)$$

where for each $(\omega, s) \mapsto u_s^{(n)}(\omega)$ is $\mathcal{F}_{[0,\, T_n]} \otimes \mathcal{B}(\mathbb{R}^+)$-measurable, with $\mathcal{F}_{[0,\, T_n]} = \sigma(T_i, X_{T_i},\, 0 \leqslant i \leqslant n)$ (see e.g. [11], Appendix A2, Theorem T30). Thus the fact that controls are predictable processes admits the following interpretation: at each time T_n (i.e. immediately after a jump) the controller, having observed the random variables T_i, X_{T_i}, $(0 \leqslant i \leqslant n)$, chooses his current action, and updates her/his decisions only at time T_{n+1}.

3. It can be proved (see [32] Theorem 4.5) that the compensator of $p(ds\, dy)$ under $\mathbb{P}^{x,a}_{u,t}$ is

$$r(t+s, X_{s-}, \theta_{s-}, y, u_s)\, \lambda(X_{s-}, \theta_{s-})\, \bar{q}(X_{s-}, \theta_{s-}, dy)\, ds,$$

whereas the compensator of $p(ds\, dy)$ under $\mathbb{P}^{x,a}$ was

$$\lambda(X_{s-}, \theta_{s-}) \bar{q}(X_{s-}, \theta_{s-}, dy) ds.$$

This explains that the choice of a given control $u(\cdot)$ affects the stochastic system multiplying its compensator by $r(t+s, X_{s-}, \theta_{s-}, y, u_s)$.

4. We call control law an arbitrary measurable function $\underline{\mathrm{u}} : [0,\, T] \times S \to U$. Given a control law one can define an admissible control u setting $u_s = \underline{\mathrm{u}}(s, X_{s-}, \theta_{s-})$. Controls of this form are called feedback controls. For a feedback control the compensator of $p(ds\, dy)$ is

$$r(t+s, X_{s-}, \theta_{s-}, y, \underline{\mathrm{u}}(s, X_{s-}, \theta_{s-}))\, \lambda(X_{s-}, \theta_{s-})\, \bar{q}(X_{s-}, \theta_{s-}, dy)\, ds$$

under $\mathbb{P}^{x,a}_{u,t}$. Thus, the process (X, θ) under the optimal probability is a two-dimensional Markov process corresponding to the transition measure

$$r(t+s, x, a, y, \underline{\mathrm{u}}(s, x, a))\, \lambda(x, a)\, \bar{q}(x, a, dy)$$

instead of $\lambda(x, a)\, \bar{q}(x, a, dy)$. However, even if the optimal control is in the feedback form, the optimal process is not, in general, time-homogeneous since the control law may depend on time. In this case, according to the definition given in Section 2, the process X under the optimal probability is not a semi-Markov process.

Remark 5 Our formulation of the optimal control problem should be compared with another approach (see e.g. [48]). In [48] is given a family of jump measures on K $\{\bar{q}(x,b,\cdot),\, b\in B\}$ with B some index set endowed with a topology. In the so called *strong formulation* a control u is an ordered pair of functions (λ',β) with $\lambda' : S\to\mathbb{R}^+$, $\beta : S\to B$ such that

$$\begin{aligned}
&\lambda' \text{ and } \beta \text{ are } \mathcal{S}-\text{measurable};\\
&\forall x\in K,\, \exists\, t(x)>0 : \int_0^{t(x)} \lambda'(x,r)\,dr<\infty;\\
&\bar{q}(\cdot,\beta,A) \text{ is } \mathcal{B}^+\text{-measurable } \forall A\in\mathcal{K}.
\end{aligned}$$

If $\mathcal{A}$ is the class of controls which satisfies the above conditions, then a control $u=(\lambda',\beta)\in\mathcal{A}$ determines a controlled process X^u in the following manner. Let

$$H^u(x,s) := 1-e^{-\int_0^s \lambda'(x,r)\,dr}, \quad \forall (x,s)\in S,$$

and suppose that $(X_0^u,\theta_0^u)=(x,a)$. Then at time 0, the process starts in state x and remains there a random time $S_1>0$, such that

$$\mathbb{P}^{x,a}\{S_1\leqslant s\} = \frac{H^u(x,a+s)-H^u(x,a)}{1-H^u(x,a)}. \tag{21}$$

At time S_1 the process transitions to the state $X^u_{S_1}$, where

$$\mathbb{P}^{x,a}\left\{X^u_{S_1}\in A|S_1\right\} = \bar{q}(x,\beta(x,S_1),A).$$

The process stays in state $X^u_{S_1}$ for a random time $S_2>0$ such that

$$\mathbb{P}^{x,a}\left\{S_2\leqslant s|S_1,\, X^u_{S_1}\right\} = H^u(X^u_{S_1},s)$$

and then at time S_1+S_2 transitions to $X^u_{S_1+S_2}$, where

$$\mathbb{P}^{x,a}\left\{X^u_{S_1+S_2}\in A|S_1,\, X^u_{S_1},\, S_2\right\} = \bar{q}(X^u_{S_1},\beta(X^u_{S_1},S_2),A).$$

We remark that the process X^u constructed in this way turns out to be semi-Markov.

We also mention that the class of control problems specified by the initial data λ' and β is in general larger that the one we address in this paper. This can be seen noticing that in our framework all the controlled processes have laws which are absolutely continuous with respect to a single uncontrolled process (the one corresponding to $r\equiv 1$) whereas this might not be the case for the rate measures $\lambda'(x,a)\,\bar{q}(x,\beta(x,a),A)$ when $u=(\lambda',\,\beta)$ ranges in the set of all possible control laws.

3.2 BSDEs and the synthesis of the optimal control

We next proceed to solve the optimal control problem formulated above. A basic role is played by the BSDE: for every fixed $t \in [0,\, T]$, $\mathbb{P}^{x,a}$-a.s.

$$\begin{aligned} Y^{x,a}_{s,t} + \int_s^{T-t} \int_K Z^{x,a}_{\sigma,t}(y)\, q(d\sigma\, dy) &= g(X_{T-t}, \theta_{T-t}) \\ &+ \int_s^{T-t} f\Big(t+\sigma, X_\sigma, \theta_\sigma, Z^{x,a}_{\sigma,t}(\cdot)\Big) d\sigma, \quad \forall s \in [0,\, T-t], \end{aligned} \tag{22}$$

with terminal condition given by the terminal cost g and generator given by the Hamiltonian function f defined for every $s \in [0,\, T]$, $(x,a) \in S$, $z \in L^2(K, \mathcal{K},\, \lambda(x,a)\, \bar q(x,a,dy))$, as

$$\begin{aligned} &f(s,x,a,z(\cdot)) \\ &= \inf_{u \in U} \Big\{ l(s,x,a,u) + \int_K z(y)(r(s,x,a,y,u)-1)\lambda(x,a)\bar q(x,a,dy) \Big\}. \end{aligned} \tag{23}$$

In (22) the superscript (x,a) denotes the starting point at time $s=0$ of the process $(X_s,\, \theta_s)_{s \geqslant 0}$, while the dependence of Y and Z on the parameter t is related to the temporal horizon of the considered optimal control problem. For every $t \in [0\; T]$, we look for a process $Y^{x,a}_{s,t}(\omega)$ adapted and cÃ dlÃ g and a process $Z^{x,a}_{s,t}(\omega, y)$ $\mathcal{P} \otimes \mathcal{K}$-measurable satisfying the integrability conditions

$$\mathbb{E}^{x,a}\left[\int_0^{T-t} \left|Y^{x,a}_{s,t}\right|^2 ds\right] < \infty,$$
$$\mathbb{E}^{x,a}\left[\int_0^{T-t} \int_K \left|Z^{x,a}_{s,t}(y)\right|^2 \lambda(X_s,\theta_s)\, \bar q(X_s,\theta_s,dy)\, ds\right] < \infty.$$

Lemma 1 *Under Hypothesis 4, all the assumptions of Hypothesis 2 hold true for the generator f and the terminal condition g in the BSDE* (22).

Proof The only non trivial verification is the Lipschitz condition (13), which follows from the boundedness assumption (16). Indeed, for every $s \in [0,\, T]$, $(x,a) \in S,\ z,\ z' \in L^2(K, \mathcal{K}, \lambda(x,a)\, \bar q(x,a,dy))$,

$$\begin{aligned} &\int_K z(y)(r(s,x,a,y,u)) - 1)\, \lambda(x,a)\, \bar q(x,a,dy) \\ &\leqslant \int_K |z(y) - z'(y)|\, (r(s,x,a,y,u)-1)\, \lambda(x,a)\, \bar q(x,a,dy) \\ &+ \int_K z'(y)(r(s,x,a,y,u)-1)\, \lambda(x,a)\, \bar q(x,a,dy) \\ &\leqslant (C_r+1)\, (\lambda(x,a)\, \bar q(x,a,K))^{1/2} \left(\int_K |z(y)-z'(y)|^2\ \lambda(x,a)\, \bar q(x,a,dy)\right)^{1/2} \end{aligned}$$

$$+\int_K z'(y)(r(s,x,a,y,u)-1)\,\lambda(x,a)\,\bar{q}(x,a,dy),$$

so that, adding $l(s,x,a,u)$ on both sides and taking the infimum over $u \in U$, it follows that

$$f(s,x,a,z) \leqslant L\left(\int_K |z(y)-z'(y)|^2\,\lambda(x,a)\,\bar{q}(x,a,dy)\right)^{1/2} + f(s,x,a,z'), \quad (24)$$

where $L := (C_r+1)\sup_{(x,a)\in S}(\lambda(x,a)\,\bar{q}(x,a,K))^{1/2}$; exchanging z and z' roles we obtain (13).

Taking into account Lemma 1, Theorem 3 implies that for every fixed $t \in [0,T]$, for every $(x,a) \in S$, there exists a unique solution $(Y^{x,a}_{s,t}, Z^{x,a}_{s,t})_{s\in[0,T-t]}$ of (22), and $Y^{x,a}_{0,t}$ is deterministic. Moreover, we have the following result:

Proposition 1 *Assume that Hypotheses 4 hold. Then, for every $t \in [0,T]$, $(x,a) \in S$, and for every $u(\cdot) \in \mathcal{A}$,*

$$Y^{x,a}_{0,t} \leqslant J(t,x,a,u(\cdot)).$$

Proof We consider the BSDE (22) at time $s=0$ and we apply the expected value $\mathbb{E}^{x,a}_{u,t}$ associated to the controlled probability $\mathbb{P}^{x,a}_{u,t}$. Since the $\mathbb{P}^{x,a}_{u,t}$-compensator of $p(dsdy)$ is $r(t+s,X_{s-},\theta_{s-},y,u_s)\,\lambda(X_{s-},\theta_{s-})\,\bar{q}(X_{s-},\theta_{s-},dy)$ ds, we have that

$$\begin{aligned}
&\mathbb{E}^{x,a}_{u,t}\left[\int_0^{T-t}\int_K Z^{x,a}_{s,t}(y)\,q(dsdy)\right]\\
&= \mathbb{E}^{x,a}_{u,t}\left[\int_0^{T-t}\int_K Z^{x,a}_{s,t}(y)\,p(dsdy)\right]\\
&- \mathbb{E}^{x,a}_{u,t}\left[\int_0^{T-t}\int_K Z^{x,a}_{s,t}(y)\,\lambda(X_s,\theta_s)\,\bar{q}(X_s,\theta_s,dy)\,ds\right]\\
&= \mathbb{E}^{x,a}_{u,t}\left[\int_0^{T-t}\int_K Z^{x,a}_{s,t}(y)\,[r(t+s,X_s,\theta_s,y,u_s)-1]\,\lambda(X_s,\theta_s)\,\bar{q}(X_s,\theta_s,dy)\,ds\right].
\end{aligned}$$

Then

$$\begin{aligned}
&Y^{x,a}_{0,t} = \mathbb{E}^{x,a}_{u,t}\,[g(X_{T-t},\theta_{T-t})] + \mathbb{E}^{x,a}_{u,t}\left[\int_0^{T-t} f(t+s,X_s,\theta_s,Z^{x,a}_{s,t}(\cdot))\,ds\right]\\
&- \mathbb{E}^{x,a}_{u,t}\left[\int_0^{T-t}\int_K Z^{x,a}_{s,t}(y)\,[r(t+s,X_s,\theta_s,y,u_s)-1]\,\lambda(X_s,\theta_s)\,\bar{q}(X_s,\theta_s,dy)\,ds\right].
\end{aligned}$$

Adding and subtracting $\mathbb{E}^{x,a}_{u,t}\left[\int_0^{T-t} l(t+s,X_s,\theta_s,u_s)\,ds\right]$ on the right side we obtain the following relation:

$$Y^{x,a}_{0,t} = J(t,x,a,u(\cdot)) \quad (25)$$

$$+\,\mathbb{E}^{x,a}_{u,t}\left[\int_0^{T-t}\left[f(t+s,X_s,\theta_s,Z^{x,a}_{s,t}(\cdot))-l(t+s,X_s,\theta_s,u_s)\right]\,ds\right]$$
$$-\,\mathbb{E}^{x,a}_{u,t}\left[\int_0^{T-t}\int_K Z^{x,a}_{s,t}(\cdot)\,[r(t+s,X_s,\theta_s,y,u_s)-1]\,\lambda(X_s,\theta_s)\,\bar q(X_s,\theta_s,dy)\,ds\right].$$

By the definition of the Hamiltonian function f, the two last terms are non positive, and it follows that $Y^{x,a}_{0,t}\leqslant J(t,x,a,u(\cdot))$, $\forall u(\cdot)\in\mathcal{A}$.

At this point, let us define the following, possibly empty, sets:

$$\begin{aligned}\varGamma(s,x,a,z(\cdot))=\Big\{\,&u\in U: f(s,x,a,z(\cdot))=l(s,x,a,u)\\ &+\int_K z(y)\,(r(s,x,a,y,u)-1)\,\lambda(x,a)\,\bar q(x,a,dy)\Big\}\end{aligned}\tag{26}$$

for $s\in[0,\,T]$, $(x,a)\in S$, $z\in L^2(K,\mathcal{K},\lambda(x,a)\,\bar q(x,a,dy))$.

In order to prove the existence of an optimal control we need to require that the infimum in the definition of f is achieved. Namely we assume that

Hypothesis 5. *The sets $\varGamma$ introduced in* (26) *are non empty; moreover, for every fixed $t\in[0,\,T]$ and $(x,a)\in S$, one can find an $\mathcal{F}$-predictable process $u^{*\,t,x,a}(\cdot)$ with values in U satisfying*

$$u^{*\,t,x,a}_s\in\varGamma(t+s,X_{s-},\theta_{s-},Z^{x,a}_{s,t}(\cdot)),\qquad \mathbb{P}^{x,a}\text{-}a.s.\ \forall s\in[0,\,T-t].\tag{27}$$

Theorem 6 *Under Hypothesis 4 and 5 for every fixed $t\in[0,\,T]$ and $(x,a)\in S$, $u^{*\,t,x,a}(\cdot)\in\mathcal{A}$ is an optimal control for the control problem starting from (x,a) at time $s=0$ with terminal value $s=T-t$. Moreover, $Y^{x,a}_{0,t}$ coincides with the value function, i.e. $Y^{x,a}_{0,t}=J(t,x,a,u^{*\,t,x,a}(\cdot))$.*

Proof The result follows immediately from the relation (25) and from the definition of the Hamiltonian function f.

Remark 6 General conditions can be formulated for the existence of a process $u^{*\,t,x,a}(\cdot)$ satisfying (27), hence of an optimal control. This can be done by means of an appropriate selection theorem, see e.g. Proposition 5.9 in [14].

We end this section with an example where the BSDE (22) can be explicitly solved and a closed form solution of an optimal control problem can be found.

Example 1 We consider a fixed time interval $[0,\,T]$ and a state space consisting of three states: $K=\{x_1,x_2,x_3,x_4\}$. We introduce $(T_n,\xi_n)_{n\geqslant 0}$ setting $(T_0,\xi_0)=(0,x_1)$, $(T_n,\xi_n)=(+\infty,x_1)$ if $n\geqslant 3$ and on (T_1,ξ_1) and (T_2,ξ_2) we make the following assumptions: ξ_1 takes values x_2 with probability 1, ξ_2 takes values x_3,x_4 with probability $1/2$. This means that the system starts at time zero in a given state x_1, jumps into state x_2 with probability 1 at the random time T_1 and into state x_3 or x_4 with equal probability at the random time T_2.

It has no jumps after. We take $U = [0, 2]$ and define the function r specifying the effects of the control process as $r(x_1, u) = r(x_2, u) = 1$, $r(x_3, u) = u$, $r(x_4, u) = 2 - u$, $u \in U$. Moreover, the final cost g assumes the value 1 in $(x, a) = (x_4, T - T_2)$ and zero otherwise, and the running cost is defined as $l(s, x, a, u) = \frac{\alpha\, u}{2}\, \lambda(x, a)$, where $\alpha > 0$ is a fixed parameter. The BSDE we want to solve takes the form:

$$\begin{aligned} Y_s + \int_s^T \int_K Z_\sigma(y)\, p(d\sigma\, dy) &= g(X_T,\, \theta_T) \\ + \int_s^T \inf_{u\in[0,\,2]} \Big\{ \frac{\alpha\, u}{2} &+ \int_K Z_\sigma(y)\, r(y, u)\, \bar{q}(X_\sigma, \theta_\sigma, dy) \Big\} \lambda(X_\sigma, \theta_\sigma) d\sigma \end{aligned} \tag{28}$$

that can be written as

$$\begin{aligned} &Y_s + \sum_{n\geqslant 1} Z_{T_n}(X_{T_n})\, \mathbb{1}_{\{s<T_n\leqslant T\}} \\ &= g(X_T,\, \theta_T) + \int_s^T \inf_{u\in[0,\,2]} \Big\{ \frac{\alpha\, u}{2} + Z_\sigma(x_2) \Big\} \lambda(x_1, a+\sigma) \mathbb{1}_{\{0\leqslant\sigma<T_1\wedge T\}}\, d\sigma \\ &+ \int_s^T \inf_{u\in[0,\,2]} \Big\{ \frac{\alpha\, u}{2} + Z_\sigma(x_3)\frac{u}{2} + Z_\sigma(x_4)(1-\frac{u}{2}) \Big\} \lambda(x_2, \sigma - T_1) \mathbb{1}_{\{T_1\leqslant\sigma<T_2\wedge T\}}\, d\sigma. \end{aligned}$$

It is known by [15] that BSDEs of this type admit the following explicit solution $(Y_s, Z_s(\cdot))_{s\in[0,\,T]}$:

$$\begin{aligned} Y_s &= y^0(s)\mathbb{1}_{\{s<T_1\}} + y^1(s, T_1, \xi_1)\, \mathbb{1}_{\{T_1\leqslant s<T_2\}} + y^2(s, T_2, \xi_2, T_1, \xi_1)\, \mathbb{1}_{\{T_2\leqslant s\}} \\ Z_s(y) &= z^0(s, y)\, \mathbb{1}_{\{s\leqslant T_1\}} + z^1(s, y, T_1, \xi_1)\, \mathbb{1}_{\{T_1<s\leqslant T_2\}}, \quad y \in K. \end{aligned}$$

To deduce y^0 and y^1 we reduce the BSDE to a system of two ordinary differential equation. To this end, it suffices to consider the following cases:

- $\omega \in \Omega$ such that $T < T_1(\omega) < T_2(\omega)$: (28) reduces to

$$\begin{aligned} y^0(s) &= \int_s^T \inf_{u\in[0,\,2]} \Big\{ \frac{\alpha\, u}{2} + z^0(\sigma, x_2) \Big\} \lambda(x_1, a+\sigma)\, d\sigma \\ &= \int_s^T z^0(\sigma, x_2)\, \lambda(x_1, a+\sigma)\, d\sigma \\ &= \int_s^T (y^1(\sigma, \sigma, x_2) - y^0(\sigma))\, \lambda(x_1, a+\sigma)\, d\sigma; \end{aligned} \tag{29}$$

- $\omega \in \Omega$ such that $T_1(\omega) < T < T_2(\omega)$, $s > T_1$: (28) reduces to

$$\begin{aligned} &y^1(s, T_1, \xi_1) \\ &= \int_s^T \inf_{u\in[0,\,2]} \Big\{ \frac{\alpha\, u}{2} + z^1(\sigma, x_3, T_1, \xi_1)\frac{u}{2} + z^1(\sigma, x_4, T_1, \xi_1)(1-\frac{u}{2}) \Big\} \cdot \\ &\qquad\qquad \cdot \lambda(\xi_1, \sigma - T_1)\, d\sigma \end{aligned}$$

$$= \int_s^T [z^1(\sigma, x_4, T_1, \xi_1) \wedge (\alpha + z^1(\sigma, x_3, T_1, \xi_1))]\, \lambda(\xi_1, \sigma - T_1)\, d\sigma$$
$$= \int_s^T [(1 \wedge \alpha) - y^1(\sigma, T_1, \xi_1)]\, \lambda(\xi_1, \sigma - T_1)\, d\sigma. \tag{30}$$

Solving (29) and (30) we obtain

$$y^0(s) = (1 \wedge \alpha) \left(1 - e^{-\int_s^T \lambda(x_1, a+\sigma)\, d\sigma}\right)$$
$$- (1 \wedge \alpha)\, e^{-\int_s^T \lambda(x_1, a+\sigma)\, d\sigma} \int_s^T \lambda(x_1, a+\sigma)\, e^{\int_\sigma^T \lambda(x_1, a+z)\, dz} e^{-\int_\sigma^T \lambda(x_2, z-\sigma)\, dz}\, d\sigma\},$$
$$y^1(s, T_1, \xi_1) = (1 \wedge \alpha) \left(1 - e^{-\int_s^T \lambda(\xi_1, \sigma - T_1)\, d\sigma}\right);$$

moreover,

$$\begin{aligned}
&y^2(s, T_2, \xi_2, T_1, \xi_1) = \mathbb{1}_{\{\xi_2 = x_4\}}, \\
&z^0(s, x_1) = z^0(s, x_3) = z^0(s, x_4) = 0, \\
&z^0(s, x_2) = y^1(s, s, x_2) - y^0(s), \\
&z^1(s, x_1, T_1, \xi_1) = z^1(s, x_2, T_1, \xi_1) = 0, \\
&z^1(s, x_3, T_1, \xi_1) = (1 \wedge \alpha) \left(e^{-\int_s^T \lambda(\xi_1, \sigma - T_1)\, d\sigma} - 1\right), \\
&z^1(s, x_4, T_1, \xi_1) = 1 + z^1(s, x_3, T_1, \xi_1),
\end{aligned}$$

where z^0 and z^1 are obtained respectively from y^2, y^1 and y^1, y^0 by subtraction.

The optimal cost is then given by $Y_0 = y^0(0)$. The optimal control is obtained during the computation of the Hamiltonian function: it is the process $u_s = 2\, \mathbb{1}_{(T_1, T_2]}(s)$ if $\alpha \leqslant 1$, and the process $u_s = 0$ if $\alpha \geqslant 1$ (both are optimal if $\alpha = 1$).

4 Nonlinear variant of Kolmogorov equation

Throughout this section we still assume that a semi-Markov process X is given. It is constructed as in Section 2.1 by the rate function λ and the measure $\bar{q}$ on K, and (X, θ) is the associated time-homogeneous Markov process. We assume that λ and $\bar{q}$ satisfy Hypothesis 1.

It is our purpose to present here some nonlinear variants of the classical backward Kolmogorov equation associated to the Markov process (X, θ) and to show that their solution can be represented probabilistically by means of an appropriate BSDE of the type considered above.

We will suppose that two functions f and g are given, satisfying Hypothesis 2, and that moreover g verifies, for every fixed $t \in [0, T]$,

$$\mathbb{E}^{x,a} \left[|g(X_{T-t}, \theta_{T-t})|^2 \right] < \infty. \tag{31}$$

We define the operator

$$\mathcal{L}\psi(x,a) := \int_K [\psi(y,0) - \psi(x,a)]\,\lambda(x,a)\,\bar{q}(x,a,dy), \qquad (x,a) \in S, \tag{32}$$

for every measurable function $\psi : S \to \mathbb{R}$ for which the integral is well defined. The equation

$$v(t,x,a) = g(x, a+T-t) + \int_t^T \mathcal{L}v(s,x,a+s-t)\,ds \tag{33}$$
$$+ \int_t^T f(s,x,a+s-t, v(s,x,a+s-t), v(s,\cdot,0) - v(s,x,a+s-t))\,ds,$$

for $t \in [0,\,T]$, $(x,a) \in S$, with unknown function $v : [0,\,T] \times S \to \mathbb{R}$ will be called the nonlinear Kolmogorov equation. Equivalently, one requires that for every $x \in K$ and for all constant $c \in [-T,\,+\infty)$,

$$t \mapsto v(t,x,t+c) \text{ is absolutely continuous on } [0,T], \tag{34}$$

and

$$\begin{cases} Dv(t,x,a) + \mathcal{L}v(t,x,a) + f(t,x,a,v(t,x,a), v(t,\cdot,0) - v(t,x,a)) = 0 \\ v(T,x,a) = g(x,a), \end{cases} \tag{35}$$

where D denotes the formal directional derivative operator

$$(Dv)(t,x,a) := \lim_{h \downarrow 0} \frac{v(t+h,x,a+h) - v(t,x,a)}{h}. \tag{36}$$

In other words, the presence of the directional derivative operator (36) allows us to understand the nonlinear Kolmogorov equation (35) in a classical sense. In particular, the first equality in (35) is understood to hold almost everywhere on $[0,\,T]$ outside of a dt-null set of points which can depend on (x,a).

We have the following result:

Lemma 2 *Suppose that f and g verify Hypothesis 2 and that* (31) *holds; suppose, in addition, that*

$$\sup_{t\in[0,\,T],\,(x,a)\in S} \Big(|g(x,a)| + |f(t,x,a,0,0)| \Big) < \infty. \tag{37}$$

Then the nonlinear Kolmogorov equation (33) *has a unique solution v in the class of measurable bounded functions.*

Proof The result follows as usual from a fixed-point argument, that we only sketch. Let us define a map Γ setting $v = \Gamma(w)$ where

$$v(t,x,a) = g(x,a+T-t) + \int_t^T \mathcal{L}w(s,x,a+s-t)\,ds$$
$$+ \int_t^T f(s,x,a+s-t, w(s,x,a+s-t), w(s,\cdot,0) - w(s,x,a+s-t))\,ds.$$

Using the Lipshitz character of f and Hypothesis 1-ii), one can show that, for some $\beta > 0$ sufficiently large, the above map is a contraction in the space of bounded measurable real functions on $[0,\,T] \times S$ endowed with the supremum norm:

$$||v||_* := \sup_{0 \leqslant t \leqslant T} \sup_{(x,a) \in S} e^{-\beta(T-t)} \, |v(t,x,a)| \, .$$

The unique fixed point of Γ gives the required solution.

Our goal is now to remove the boundedness assumption (37). To this end we need to define a formula of Itô type for the composition of the process $(X_s,\, \theta_s)_{s \geqslant 0}$ with functions v smooth enough defined on $[0,\,T] \times S$. Taking into account the particular form of (33), and the fact that the second component of the process $(X_s,\, \theta_s)_{s \geqslant 0}$ is linear in s, the idea is to use in this formula the directional derivative operator D given by (36).

Lemma 3 (A formula of Itô type) *Let* $v : [0,\,T] \times S \to \mathbb{R}$ *such that*

(i) $\forall\, x \in K$, $\forall\, c \in [-T,\, +\infty)$*, the map* $t \mapsto v(t,x,t{+}c)$ *is absolutely continuous on* $[0,\,T]$*, with directional derivative* D *given by* (36)*;*
(ii) for fixed $t \in [0,\,T]$, $\{v(t{+}s,y,0) - v(t{+}s,X_{s-},\theta_{s-}),\ s \in [0,\,T{-}t],\ y \in K\}$ *belongs to* $\mathcal{L}^1_{loc}(p)$*.*

Then $\mathbb{P}^{x,a}$*-a.s., for every* $t \in [0,\,T]$*,*

$$\begin{aligned} &v(T, X_{T-t}, \theta_{T-t}) - v(t,x,a) \\ &= \int_0^{T-t} Dv(t+s, X_s, \theta_s)\, ds + \int_0^{T-t} \mathcal{L}v(t+s, X_s, \theta_s)\, ds \\ &+ \int_0^{T-t} \int_K (v(t+s,y,0) - v(t+s, X_{s-}, \theta_{s-}))\; q(ds,dy), \end{aligned} \tag{38}$$

where the stochastic integral is a local martingale.

Proof We proceed by reasoning as in the proof of Theorem 26.14 in [19]. We consider a function $v : [0,\,T] \times S \to \mathbb{R}$ satisfying (i) and (ii), and we denote by N_t the number of jumps in the interval $[0,\,t]$:

$$N_t = \sum_{n \geqslant 1} \mathbb{1}_{\{T_n \leqslant t\}}.$$

We have

$$\begin{aligned} v(T, X_T, \theta_T) - v(0,x,a) &= v(T, X_T, \theta_T) - v(T_{N_T}, X_{T_{N_T}}, \theta_{T_{N_T}}) \\ &+ \sum_{n=2}^{N_T} \big\{ v(T_n, X_{T_n}, \theta_{T_n}) - v(T_{n-1}, X_{T_{n-1}}, \theta_{T_{n-1}}) \big\} \\ &+ \ v(T_1, X_{T_1}, \theta_{T_1}) - v(0,x,a). \end{aligned}$$

Noticing that $X_{T_{n-}} = X_{T_{n-1}}$ for all $n \in [1, N_T]$, $X_T = X_{T_{N_T}}$, and that $\theta_{T_n} = 0$ for all $n \in [1, N_T]$, $\theta_{T_{1-}} = a + T_1$, and $\theta_{T_{n-}} = T_n - T_{n-1}$ for all $n \in [2, N_T]$, we have

$$v(T, X_T, \theta_T) - v(0, x, a) = I + II + III,$$

where

$$\begin{aligned}
I &= (v(T_1, X_{T_1}, 0) - v(T_1, X_{T_1-}, \theta_{T_1-})) + (v(T_1, x, a + T_1) - v(0, x, a)) \\
&=: I' + I'', \\
II &= \sum_{n=2}^{N_T} (v(T_n, X_{T_n}, 0) - v(T_n, X_{T_n-}, \theta_{T_n-}) \\
&+ \sum_{n=2}^{N_T} (v(T_n, X_{T_{n-1}}, T_n - T_{n-1}) - v(T_{n-1}, X_{T_{n-1}}, 0))) \\
&=: II' + II'', \\
III &= v(T, X_T, T - T_N) - v(T_N, X_{T_N}, 0).
\end{aligned}$$

Let H denote the $\mathcal{P} \otimes \mathcal{K}$-measurable process

$$H_s(y) = v(s, y, 0) - v(s, X_{s-}, \theta_{s-}),$$

with the convention $X_{0-} = X_0$, $\theta_{0-} = \theta_0$. We have

$$\begin{aligned}
I' + II' &= \sum_{n \geqslant 1 : T_n \leqslant T} (v(T_n, X_{T_n}, 0) - v(T_n, X_{T_{n-}}, \theta_{T_{n-}})) \\
&= \sum_{n \geqslant 1 : T_n \leqslant T} H_{T_n}(X_{T_N}) = \int_0^T \int_K H_s(y)\, p(ds, dy).
\end{aligned}$$

On the other hand, since v satisfies (i) and recalling the definition 36 of the directional derivative operator D,

$$\begin{aligned}
I'' + II'' + III &= \int_0^{T_1} \lim_{h \to 0} \frac{v(0 + hs, x, a + hs) - v(0, x, a)}{h}\, ds \\
&+ \sum_{n \geqslant 2 : T_n \leqslant T} \int_{T_{n-1}}^{T_n} \lim_{h \to 0} \frac{1}{h} [v(T_{n-1} + h(s - T_{n-1}), X_{T_{n-1}}, \theta_{T_{n-1}} + h(s - T_{n-1})) \\
&\qquad - v(T_{n-1}, X_{T_{n-1}}, \theta_{T_{n-1}})]\, ds \\
&+ \int_{T_{N_T}}^T \lim_{h \to 0} \frac{1}{h} [v(T_{N_T} + h(s - T_{N_T}), X_{T_{N_T}}, \theta_{T_{N_T}} + h(s - T_{N_T})) \\
&\qquad - v(T_{N_T}, X_{T_{N_T}}, \theta_{T_{N_T}})]\, ds \\
&= \int_0^T Dv(s, X_s, \theta_s)\, ds.
\end{aligned}$$

Then $\mathbb{P}^{x,a}$-a.s.,

$$\begin{aligned}
&v(T, X_T, \theta_T) - v(0, x, a)\\
&= \int_0^T Dv(s, X_s, \theta_s)\, ds + \int_0^T \int_K (v(s, y, 0) - v(s, X_{s-}, \theta_{s-}))\; p(ds, dy)\\
&= \int_0^T Dv(s, X_s, \theta_s)\, ds + \int_0^T \mathcal{L}v(s, X_s, \theta_s)\, ds\\
&+ \int_0^T \int_K (v(s, y, 0) - v(s, X_{s-}, \theta_{s-}))\; q(ds, dy),
\end{aligned}$$

where the second equality is obtained using the identity $q(dt\, dy) = p(dt\, dy) - \lambda(X_{t-}, \theta_{t-})\, \bar{q}(X_{t-}, \theta_{t-}, dy)\, dt$ together with the definition (32) of the operator $\mathcal{L}$.

Finally, applying a shift in time, i.e. considering for every $t \in [0, T]$ the differential of the process $v(s + t, X_{s-}, \theta_{s-})$ with respect to $s \in [0, T - t]$, the previous formula becomes: $\mathbb{P}^{x,a}$-a.s., for every $t \in [0, T]$,

$$\begin{aligned}
&v(T - t, X_T, \theta_T) - v(t, x, a)\\
&= \int_0^{T-t} Dv(s + t, X_s, \theta_s)\, ds + \int_0^{T-t} \mathcal{L}v(s + t, X_s, \theta_s)\, ds\\
&+ \int_0^{T-t} \int_K (v(s + t, y, 0) - v(s + t, X_{s-}, \theta_{s-}))\; q(ds, dy),
\end{aligned}$$

where the stochastic integral is a local martingale thanks to condition (ii).

We will call (38) the Itô formula for $v(t + s, \cdot, \cdot) \circ (X_s, \theta_s)_{s \in [0,\, T-t]}$. In differential notation we have:

$$\begin{aligned}
dv(t + s, X_{s-}, \theta_{s-}) &= Dv(t + s, X_{s-}, \theta_{s-})\, ds \;+\; \mathcal{L}v(t + s, X_{s-}, \theta_{s-})\, ds\\
&\quad + \int_K (v(t + s, y, 0) - v(t + s, X_{s-}, \theta_{s-}))\; q(ds, dy).
\end{aligned}$$

Remark 7 With respect to the classical Itô formula, we underline that in (38) we have

- the directional derivative operator D instead of the usual time derivative;
- the temporal translation in the first component of v, i.e. we consider the differential of the process $v(t + s, X_{s-}, \theta_{s-})$ with respect to $s \in [0, T - t]$. Indeed, the time-homogeneous Markov process $(X_s, \theta_s)_{s \geqslant 0}$ satisfies

$$\mathbb{P}^{x,a}(X_0 = x,\, \theta_0 = a) = 1,$$

and the temporal translation in the first component allows us to consider $dv(t, X_t, \theta_t)$ for all $t \in [0, T]$.

Let us go back to consider the Kolmogorov equation (33) in a more general setting. More precisely, on the functions f, g we will only ask that they satisfy Hypothesis 2 for every $(x, a) \in S$ and that (31) holds.

Definition 1 We say that a measurable function $v : [0, T] \times S \to \mathbb{R}$ is a solution of the nonlinear Kolmogorov equation (33), if, for every fixed $t \in [0, T]$, $(x, a) \in S$,

1. $\mathbb{E}^{x,a}\left[\int_0^{T-t}\int_K |v(t+s, y, 0) - v(t+s, X_s, \theta_s)|^2\, \lambda(X_s, \theta_s)\, \bar{q}(X_s, \theta_s, dy)\, ds\right] < \infty$;
2. $\mathbb{E}^{x,a}\left[\int_0^{T-t} |v(t+s, X_s, \theta_s)|^2\, ds\right] < \infty$;
3. (33) is satisfied.

Remark 8 Condition 1. is equivalent to the fact that $v(t+s, y, 0) - v(t+s, X_{s-}, \theta_{s-})$ belongs to $\mathcal{L}^2(p)$. Conditions 1. and 2. together are equivalent to the fact that the pair $\{v(t+s, X_s, \theta_s),\ v(t+s, y, 0) - v(t+s, X_{s-}, \theta_{s-});\ s \in [0, T-t],\ y \in K\}$ belongs to the space $\mathbb{M}^{x,a}$; in particular they hold true for every measurable bounded function v.

Remark 9 We need to verify the well-posedness of equation (33) for a function v satisfying the condition 1. and 2. above. We start by noticing that, for every $(x, a) \in S$, $\mathbb{P}^{x,a}$-a.s.,

$$\int_0^T \int_K |v(s, y, 0) - v(s, X_s, \theta_s)|^2\, \lambda(X_s, \theta_s)\, \bar{q}(X_s, \theta_s, dy)\, ds + \int_0^T |v(s, X_s, \theta_s)|^2\, ds < \infty.$$

By the law (7) of the first jump it follows that the set $\{\omega \in \Omega : T_1(\omega) > T\}$ has positive $\mathbb{P}^{x,a}$ probability, and on this set we have $X_{s-}(\omega) = x$, $\theta_{s-}(\omega) = a + s$. Taking such an ω we get

$$\int_0^T \int_K |v(s, y, 0) - v(s, x, a+s)|^2\ \lambda(x, a+s)\, \bar{q}(x, a+s, dy)\, ds + \int_0^T |v(s, x, a+s)|^2\, ds < \infty, \qquad \forall (x, a) \in S.$$

Since $\sup_{(x,a)\in S} \lambda(x, a)\bar{q}(x, a, K) < \infty$ by assumption, Hölder's inequality implies that

$$\begin{aligned}
&\int_0^T |\mathcal{L}(v(s, x, a+s))|\ ds \\
&\leqslant \int_0^T \int_K |v(s, y, 0) - v(s, x, a+s)|\ \lambda(x, a+s)\, \bar{q}(x, a+s, dy)\, ds \\
&\leqslant c \left(\int_0^T \int_K |v(s, y, 0) - v(s, x, a+s)|^2\ \lambda(x, a+s)\, \bar{q}(x, a+s, dy)\, ds\right)^{1/2}
\end{aligned}$$

$$< \infty$$

for some constant c and for all $(x,a) \in S$. Similarly, since

$$\mathbb{E}^{x,a}\left[\int_0^T |f(s, X_s, \theta_s, 0, 0)|^2\, ds\right] < \infty,$$

and arguing again on the jump time T_1, we deduce that

$$\int_0^T |f(s, x, a+s, 0, 0)|^2\, ds < \infty, \qquad \forall (x,a) \in S;$$

finally, from the Lipschitz conditions on f we can conclude that

$$\begin{aligned}
&\int_0^T |f(s, x, a+s, v(s,x,a+s), v(s,\cdot,0) - v(s,x,a+s))|\, ds \\
&\leqslant c_1 \left(\int_0^T |f(s,x,a+s,0,0)|^2\, ds\right)^{1/2} + c_2 \left(\int_0^T |v(s,x,a+s)|^2\, ds\right)^{1/2} \\
&+ c_3 \left(\int_0^T \int_K |v(s,y,0) - v(s,x,a+s)|^2\, \lambda(x,a+s)\, \bar{q}(x,a+s,dy)\, ds\right)^{1/2} \\
&< \infty
\end{aligned}$$

for some constants c_i, $i = 1,2,3$, and for all $(x,a) \in S$. Therefore, all terms occurring in equation (33) are well defined.

For every fixed $t \in [0,\, T]$ and $(x,a) \in S$, we consider now a BSDE of the form

$$\begin{aligned}
Y_{s,t}^{x,a} + \int_s^{T-t} \int_K Z_{r,t}^{x,a}(y)\, q(dr\, dy) &= g(X_{T-t}, \theta_{T-t}) \\
+ \int_s^{T-t} f\Big(t+r, X_{r-}, \theta_{r-}, Y_{r,t}^{x,a}, Z_{r,t}^{x,a}(\cdot)\Big)\, dr&,\ s \in [0,\, T-t]. \qquad (39)
\end{aligned}$$

Then there exists a unique solution $(Y_{s,t}^{x,a}, Z_{s,t}^{x,a}(\cdot))_{s\in[0,\,T-t]}$, in the sense of Theorem 3, and $Y_{0,t}^{x,a}$ is deterministic. We are ready to state the main result of this section.

Theorem 7 *Suppose that f, g satisfy Hypothesis 2 for every $(x,a) \in S$ and that* (31) *holds. Then for every $t \in [0,\, T]$, the nonlinear Kolmogorov equation* (33) *has a unique solution $v(t,x,a)$ in the sense of Definition 1.*

Moreover, for every fixed $t \in [0,\, T]$, for every $(x,a) \in S$ and $s \in [0,\, T-t]$ we have

$$Y_{s,t}^{x,a} = v(t+s, X_{s-}, \theta_{s-}), \qquad (40)$$

$$Z_{s,t}^{x,a}(y) = v(t+s, y, 0) - v(t+s, X_{s-}, \theta_{s-}), \qquad (41)$$

so that in particular $v(t,x,a) = Y_{0,t}^{x,a}$.

Remark 10 The equalities (40) and (41) are understood as follows.

- $\mathbb{P}^{x,a}$-a.s., equality (40) holds for all $s \in [0, T-t]$. The trajectories of $(X_s)_{s\in[0,T-t]}$ are piecewise constant and cÃ dlÃ g, while the trajectories of $(\theta_s)_{s\in[0,T-t]}$ are piecewise linear in s (with unitary slope) and cÃ dlÃ g; moreover the processes $(X_s)_{s\in[0,T-t]}$ and $(\theta_s)_{s\in[0,T-t]}$ have the same jump times $(T_n)_{n\geqslant 1}$. Then the equality (40) is equivalent to the condition

$$\mathbb{E}^{x,a}\left[\int_0^{T-t} \left|Y_{s,t}^{x,a} - v(t+s, X_s, \theta_s)\right|^2 ds\right] = 0.$$

- The equality (41) holds for all (ω, s, y) with respect to the measure $\lambda(X_{s-}(\omega), \theta_{s-}(\omega))\, \bar{q}(X_{s-}(\omega), \theta_{s-}(\omega), dy)\, \mathbb{P}^{x,a}(d\omega)ds$, i.e.,

$$\mathbb{E}^{x,a}\Big[\int_0^{T-t}\int_K \left|Z_{s,t}^{x,a}(y) - v(t+s, y, 0) + v(t+s, X_s, \theta_s)\right|^2 \cdot \\ \cdot\, \lambda(X_s, \theta_s)\, \bar{q}(X_s, \theta_s, dy)\, ds\Big] = 0.$$

Proof Uniqueness. Let v be a solution of the nonlinear Kolmogorov equation (33). It follows from equality (33) itself that for every $x \in K$ and every $\tau \in [-T, +\infty)$, $t \mapsto v(t, x, t+\tau)$ is absolutely continuous on $[0, T]$. Indeed, applying in (33) the change of variable $\tau := a - t$, we obtain $\forall t \in [0, T]$, $\forall \tau \in [-T, +\infty)$,

$$v(t, x, t+\tau) = g(x, T+\tau) + \int_t^T \mathcal{L}v(s, x, s+\tau)\, ds \\ + \int_t^T f(s, x, s+\tau, v(s, x, s+\tau), v(s, \cdot, 0) - v(s, x, s+\tau))\, ds.$$

Then, since by assumption the process $v(t+s, y, 0) - v(t+s, X_{s-}, \theta_{s-})$ belongs to $\mathcal{L}^2(p)$, we are in a position to apply the Itô formula (38) to the process $v(t+s, X_{s-}, \theta_{s-})$, $s \in [0, T-t]$. We get: $\mathbb{P}^{x,a}$-a.s.,

$$v(t+s, X_{s-}, \theta_{s-}) = v(t, x, a) + \int_0^s Dv(t+r, X_r, \theta_r)\, dr \\ + \int_0^s \mathcal{L}v(t+r, X_r, \theta_r)\, dr \\ + \int_0^s \int_K (v(t+r, y, 0) - v(t+r, X_r, \theta_r))\, q(dr, dy),$$

for every $s \in [0, T-t]$. We know that v satisfies (35); moreover the process X has piecewise constant trajectories, the process a has linear trajectories in s, and they have the same time jumps. Then, $\mathbb{P}^{x,a}$-a.s.,

$$Dv(t+s, X_{s-}, \theta_{s-}) + \mathcal{L}v(t+s, X_{s-}, \theta_{s-}) \\ + f(t+s, X_{s-}, \theta_{s-}, v(t+s, X_{s-}, \theta_{s-}), v(t+s, \cdot, 0) - v(t+s, X_{s-}, \theta_{s-})) = 0,$$

for almost $s \in [0, T-t]$. In particular, $\mathbb{P}^{x,a}$-a.s.,

$$\begin{aligned}
&v(t+s, X_{s-}, \theta_{s-}) = v(t,x,a) \\
&+ \int_0^s \int_K (v(t+r,y,0) - v(t+r, X_{r-}, \theta_{r-}))\, q(dr,dy) \\
&- \int_0^s f(t+r, X_r, \theta_r, v(t+s, X_s, \theta_s), v(t+r,y,0) - v(t+r, X_r, \theta_r))\, dr,
\end{aligned}$$

for $s \in [0, T-t]$. Since $v(T,x,a) = g(x,a)$ for all $(x,a) \in S$, by simple computations we can prove that, $\forall s \in [0, T-t]$,

$$\begin{aligned}
&v(t+s, X_{s-}, \theta_{s-}) + \int_s^{T-t} \int_K (v(t+r,y,0) - v(t+r, X_{r-}, \theta_{r-}))\, q(dr,dy) \\
&= g(X_{T-t}, \theta_{T-t}) \\
&+ \int_s^{T-t} f(t+r, X_r, \theta_r, v(t+r, X_r, \theta_r), v(t+r,y,0) - v(t+r, X_r, \theta_r))\, dr.
\end{aligned}$$

Since the pairs $(Y^{x,a}_{s,t}, Z^{x,a}_{s,t}(\cdot))_{s\in[0,\,T-t]}$ and $(v(t+s, X_{s-}, \theta_{s-}), v(t+s,y,0) - v(t+s, X_{s-}, \theta_{s-}))_{s\in[0,\,T-t]}$ are both solutions to the same BSDE under $\mathbb{P}^{x,a}$, they coincide as members of the space $\mathbb{M}^{x,a}$. It follows that equalities (40) and (41) hold. In particular, $v(t,x,a) = Y^{x,a}_{0,t}$, and this yields the uniqueness of the solution.

Existence. We proceed by an approximation argument, following the same lines of the proof of Theorem 4.4 in [14]. We recall that, by Theorem 3, for every fixed $t \in [0, T]$, the BSDE (39) has a unique solution $(Y^{x,a}_{s,t}, Z^{x,a}_{s,t}(\cdot))_{s\in[0,\,T-t]}$ for every $(x,a) \in S$; moreover, $Y^{x,a}_{0,t}$ is deterministic, i.e., there exists a real number, denoted by $v(t,x,a)$, such that $\mathbb{P}^{x,a}(Y^{x,a}_{0,t} = v(t,x,a)) = 1$. At this point, we set $f^n = (f \wedge n) \vee (-n)$ and $g^n = (g \wedge n) \vee (-n)$ as the truncations of f and g at level n. By Lemma 2, for $t \in [0, T]$, $(x,a) \in S$, equation

$$\begin{aligned}
&v^n(t,x,a) = g^n(x, a+T-t) + \int_t^T \mathcal{L}v^n(s,x,a+s-t)\, ds \qquad (42)\\
&+ \int_t^T f^n(s,x,a+s-t, v^n(s,x,a+s-t), v^n(s,\cdot,0) - v^n(s,x,a+s-t))\, ds
\end{aligned}$$

admits a unique bounded measurable solution v^n. In particular, the first part of the proof yield the following identifications:

$$\begin{aligned}
v^n(t,x,a) &= Y^{x,a,n}_{0,t}, \\
v^n(t+s, X_{s-}, \theta_{s-}) &= Y^{x,a,n}_{s,t}, \\
v^n(t+s,y,0) - v^n(t+s, X_{s-}, \theta_{s-}) &= Z^{x,a,n}_{s,t}(y),
\end{aligned}$$

in the sense of Remark 10, where $(Y^{x,a,n}_{s,t}, Z^{x,a,n}_{s,t}(\cdot))_{s\in[0,\,T-t]}$ is the unique solution to the BSDE

$$Y^{x,a,n}_{s,t} + \int_s^{T-t} \int_K Z^{x,a,n}_{r,t}(y)\, q(dr\, dy)$$

$$= g^n(X_{T-t}, \theta_{T-t}) + \int_s^{T-t} f^n\left(t + r, X_r, \theta_r, Y^{x,a,n}_{r,t}, Z^{x,a,n}_{r,t}(\cdot)\right) dr,$$

for all $s \in [0,\, T-t]$. Recalling (39) and applying Theorem 3, we deduce that, for some constant c,

$$\begin{aligned}
&\sup_{s\in[0,\,T-t]} \mathbb{E}^{x,a}\left[|Y^{x,a}_{s,t} - Y^{x,a,n}_{s,t}|^2\right] + \mathbb{E}^{x,a}\left[\int_0^{T-t} |Y^{x,a}_{s,t} - Y^{x,a,n}_{s,t}|^2 ds\right] \\
&+ \mathbb{E}^{x,a}\left[\int_0^{T-t}\int_K |Z^{x,a}_{s,t}(y) - Z^{x,a,n}_{s,t}(y)|^2 \lambda(X_s,\theta_s)\, \bar{q}(X_s,\theta_s,dy)\, ds\right] \\
&\leqslant c\mathbb{E}^{x,a}\left[|g(X_{T-t},\theta_{T-t}) - g^n(X_{T-t},\theta_{T-t})|^2\right] \\
&+ c\mathbb{E}^{x,a}\Big[\int_0^{T-t} |f(t+s, X_s, \theta_s, Y^{x,a}_{s,t}, Z^{x,a}_{s,t}(\cdot)) \\
&\qquad\qquad - f^n(t+s, X_s, \theta_s, Y^{x,a}_{s,t}, Z^{x,a}_{s,t}(\cdot))|^2 ds\Big] \overset{n\to\infty}{\longrightarrow} 0,
\end{aligned} \tag{43}$$

where the two final terms tend to zero by monotone convergence. In particular (43) yields

$$\begin{aligned}
|v(t,x,a) - v^n(t,x,a)|^2 &= |Y^{x,a}_{0,t} - Y^{x,a,n}_{0,t}|^2 \\
&\leqslant \sup_{s\in[0,\,T-t]} \mathbb{E}^{x,a}\left[|Y^{x,a}_{s,t} - Y^{x,a,n}_{s,t}|^2\right] \overset{n\to\infty}{\longrightarrow} 0,
\end{aligned}$$

and therefore v is a measurable function. At this point, applying the Fatou Lemma we get

$$\begin{aligned}
&\mathbb{E}^{x,a}\left[\int_0^{T-t} \left|Y^{x,a}_{s,t} - v(t+s, X_s, \theta_s)\right|^2 ds\right] \\
&+ \mathbb{E}^{x,a}\Bigg[\int_0^{T-t}\int_K \left|Z^{x,a}_{s,t}(y) - v(t+s,y,0) + v(t+s,X_s,\theta_s)\right|^2 \cdot \\
&\qquad\qquad\qquad \cdot \lambda(X_s,\theta_s)\, \bar{q}(X_s,\theta_s,dy)\, ds\Bigg] \\
&\leqslant \liminf_{n\to\infty} \mathbb{E}^{x,a}\left[\int_0^{T-t} \left|Y^{x,a}_{s,t} - v^n(t+s, X_s, \theta_s)\right|^2 ds\right] \\
&+ \liminf_{n\to\infty} \mathbb{E}^{x,a}\Bigg[\int_0^{T-t}\int_K \left|Z^{x,a}_{s,t}(y) - v^n(t+s,y,0) + v^n(t+s,X_s,\theta_s)\right|^2 \cdot \\
&\qquad\qquad\qquad \cdot \lambda(X_s,\theta_s)\, \bar{q}(X_s,\theta_s,dy)\, ds\Bigg] \\
&= \liminf_{n\to\infty} \mathbb{E}^{x,a}\left[\int_0^{T-t} \left|Y^{x,a}_{s,t} - Y^{x,a,n}_{s,t}\right|^2 ds\right]
\end{aligned}$$

$$+ \liminf_{n\to\infty} \mathbb{E}^{x,a} \left[\int_0^{T-t} \int_K \left| Z_{s,t}^{x,a}(y) - Z_{s,t}^{x,a,n}(y) \right|^2 \lambda(X_s,\theta_s)\, \bar{q}(X_s,\theta_s,dy)\, ds \right]$$
$$= 0$$

by (43). The above calculations show that (40) and (41) hold. Moreover, they imply that

$$\mathbb{E}^{x,a} \left[\int_0^{T-t} |v(t+s,X_s,\theta_s)|^2\, ds \right]$$
$$+ \mathbb{E}^{x,a} \left[\int_0^{T-t} \int_K |v(t+s,y,0) - v(t+s,X_s,\theta_s)|^2 \lambda(X_s,\theta_s)\, \bar{q}(X_s,\theta_s,dy)\, ds \right]$$
$$= \mathbb{E}^{x,a} \left[\int_0^{T-t} \left| Y_{s,t}^{x,a} \right|^2 ds \right]$$
$$+ \mathbb{E}^{x,a} \left[\int_0^{T-t} \int_K \left| Z_{s,t}^{x,a}(y) \right|^2 \lambda(X_{s-},\theta_{s-})\, \bar{q}(X_s,\theta_s,dy)\, ds \right]$$
$$< \infty,$$

that accords to requirement of Definition 1.

It remains to show that v satisfies (33). This would follow from a passage to the limit in (42), provided we show that

$$\int_t^T \mathcal{L}v^n(s,x,a+s-t)ds \to \int_t^T \mathcal{L}v(s,x,a+s-t)ds, \tag{44}$$

and

$$\int_t^T f^n(s,x,a+s-t, v^n(s,x,a+s-t), v^n(s,\cdot,0) - v^n(s,x,a+s-t))\, ds$$
$$\to \int_t^T f(s,x,a+s-t, v(s,x,a+s-t), v(s,\cdot,0) - v(s,x,a+s-t))\, ds. \tag{45}$$

To prove (44), we observe that

$$\mathbb{E}^{x,a} \left| \int_0^{T-t} \mathcal{L}v(t+s,X_{s-},\theta_{s-})\, ds - \int_0^{T-t} \mathcal{L}v^n(t+s,X_{s-},\theta_{s-})\, ds \right|$$
$$= \mathbb{E}^{x,a} \left| \int_0^{T-t} \int_K (Z_{s,t}^{x,a} - Z_{s,t}^{x,a,n})\, \lambda(X_s,\theta_s)\, \bar{q}(X_s,\theta_s,dy)\, ds \right|$$
$$\leqslant (T-t)^{1/2} \sup_{x,a} [\lambda(x,a)\, \bar{q}(x,a,K)]^{1/2} \cdot$$
$$\cdot \left(\mathbb{E}^{x,a} \left[\int_0^{T-t} \int_K \left| Z_{s,t}^{x,a} - Z_{s,t}^{x,a,n} \right| \lambda(X_s,\theta_s)\, \bar{q}(X_s,\theta_s,dy)\, ds \right] \right)^{1/2} \stackrel{n\to\infty}{\longrightarrow} 0,$$

by (43). Then, for a subsequence (still denoted v^n) we get

$$\int_0^{T-t} \mathcal{L}v^n(t+s, X_s, \theta_s)\, ds \to \int_0^{T-t} \mathcal{L}v(t+s, X_s, \theta_s)\, ds, \quad \mathbb{P}^{x,a}\text{-a.s.}$$

Recalling the law (7) of the first jump T_1, we see that the set $\{\omega \in \Omega : T_1(\omega) > T\}$ has positive $\mathbb{P}^{x,a}$ probability, and on this set we have $X_{s-}(\omega) = x$, $\theta_{s-}(\omega) = a + s$. Choosing such an ω we have

$$\int_0^{T-t} \mathcal{L}v^n(t+s, x, a+s)ds \overset{n\to\infty}{\longrightarrow} \int_0^{T-t} \mathcal{L}v(t+s, x, a+s)ds,$$

i.e., by a translation of t in the temporal line,

$$\int_t^{T} \mathcal{L}v^n(s, x, a+s-t)ds \overset{n\to\infty}{\longrightarrow} \int_t^{T} \mathcal{L}v(s, x, a+s-t)ds.$$

To show (45), we compute

$$\begin{aligned}
&\mathbb{E}^{x,a}\left[\left|\int_0^{T-t} f(t+s, X_s, \theta_s, Y_{s,t}^{x,a}, Z_{s,t}^{x,a}) - f^n(t+s, X_s, \theta_s, Y_{s,t}^{x,a,n}, Z_{s,t}^{x,a,n}))\, ds\right|\right] \\
&\leqslant \mathbb{E}^{x,a}\left[\int_0^{T-t} \left|f(t+s, X_s, \theta_s, Y_{s,t}^{x,a}, Z_{s,t}^{x,a}) - f^n(t+s, X_s, \theta_s, Y_{s,t}^{x,a}, Z_{s,t}^{x,a})\right|\, ds\right] \\
&+ \mathbb{E}^{x,a}\left[\int_0^{T-t} \left|f^n(t+s, X_s, \theta_s, Y_{s,t}^{x,a}, Z_{s,t}^{x,a}) - f^n(t+s, X_s, \theta_s, Y_{s,t}^{x,a,n}, Z_{s,t}^{x,a,n})\right|\, ds\right].
\end{aligned}$$

The first integral term in the right-hand side tends to zero by monotone convergence. At this point, we notice that f^n is a truncation of f, and therefore it satisfies the Lipschitz condition (13) with the same constants L, L', independent of n. This yields the following estimate for the second integral:

$$\begin{aligned}
&L'\, \mathbb{E}^{x,a}\left[\int_0^{T-t} \left|Y_{s,t}^{x,a} - Y_{s,t}^{x,a,n}\right| ds\right] \\
&+ L\, \mathbb{E}^{x,a}\left[\int_0^{T-t} \Big(\int_K \left|Z_{s,t}^{x,a}(y) - Z_{s,t}^{x,a,n}(y)\right|^2 \lambda(X_s, \theta_s)\, \bar{q}(X_s, \theta_s, dy)\Big)^{1/2} ds\right] \\
&\leqslant L' \sqrt{T-t} \left(\mathbb{E}^{x,a}\left[\int_0^{T-t} \left|Y_{s,t}^{x,a} - Y_{s,t}^{x,a,n}\right|^2 ds\right]\right)^{1/2} + L\sqrt{T-t} \cdot \\
&\qquad \cdot \left(\mathbb{E}^{x,a}\left[\int_0^{T-t} \int_K \left|Z_{s,t}^{x,a}(y) - Z_{s,t}^{x,a,n}(y)\right|^2 \lambda(X_s, \theta_s)\, \bar{q}(X_s, \theta_s, dy)\, ds\right]\right)^{1/2},
\end{aligned}$$

which tends to zero, again by (43). Considering a subsequence (still denoted v^n) we get, $\mathbb{P}^{x,a}$-a.s.,

$$\int_0^{T-t} f^n(t+s, X_s, \theta_s, v^n(t+s, X_s, \theta_s), v^n(t+s, y, 0) - v^n(t+s, X_s, \theta_s))\, ds$$

$$\overset{n\to\infty}{\longrightarrow} \int_0^{T-t} f(t+s, X_s, \theta_s, v(t+s, X_s, \theta_s), v(t+s, y, 0) - v(t+s, X_s, \theta_s))\, ds.$$

Choosing also in this case an ω in the set $\{\omega \in \Omega :\, T_1(\omega) > T\}$, we find

$$\int_0^{T-t} f^n(t+s, x, a+s, v^n(t+s, x, a+s), v^n(t+s, y, 0) - v^n(t+s, x, a+s))\, ds$$
$$\overset{n\to\infty}{\longrightarrow} \int_0^{T-t} f(t+s, x, a+s, v(t+s, x, a+s), v(t+s, y, 0) - v(t+s, x, a+s))\, ds,$$

and a change of temporal variable allows us to prove that (33) holds, and to conclude the proof.

We finally introduce the Hamilton-Jacobi-Bellman (HJB) equation associated to the control problem considered in Section 3: for every $t \in [0, T]$ and $(x, a) \in S$,

$$\begin{aligned} v(t, x, a) = g(x, a+T-t) &+ \int_t^T \mathcal{L}v(s, x, a+s-t)\, ds \\ &+ \int_t^T f(s, x, a+s-t, v(s, \cdot, 0) - v(s, x, a+s-t))\, ds, \end{aligned} \tag{46}$$

where $\mathcal{L}$ denotes the operator introduced in (32), f is the Hamiltonian function defined by (23) and g is the terminal cost. Since (46) is a nonlinear Kolmogorov equation of the form (33), we can apply Theorem 7 and conclude that the value function and an optimal control law can be represented by means of the HJB solution $v(t, x, a)$.

Corollary 1 *Let Hypotheses 4 and 5 hold. For every fixed $t \in [0, T]$, for every $(x, a) \in S$ and $s \in [0, T-t]$, there exists a unique solution v to the HJB equation (46), satisfying*

$$\begin{aligned} v(t+s, X_{s-}, \theta_{s-}) &= Y_{s,t}^{x,a}, \\ v(t+s, y, 0) - v(t+s, X_{s-}, \theta_{s-}) &= Z_{s,t}^{x,a}(y), \end{aligned}$$

where the above equalities are understood as explained in Remark 10.
In particular an optimal control is given by the formula

$$u_s^{*\,t,x,a} \in \Gamma(t+s, X_{s-}, \theta_{s-}, v(t+s, \cdot, 0) - v(t+s, X_{s-}, \theta_{s-})),$$

while the value function coincides with $v(t, x, a)$, i.e. $J(t, x, a, u^{*\,t,x,a}(\cdot)) = v(t, x, a) = Y_{0,t}^{x,a}$.

Remark 11 In [48] the author considered the so called *strong formulation* of the control problem. In particular, he chose the admissible controls in order to obtain controlled Markov processes with stationary transitions probabilities, see also Remark 5. This allowed him to write an HJB equation similar to (46) involving the operator $\mathcal{L}$ introduced in (32), equation (4.14) in [48]. Our HJB equation (46) is obtained instead by means of a *weak formulation* of the control problem. The corresponding controlled processes are Markov but in general non-homogeneous. Our approach, based on the introduction of the directional derivative operator D, permits us to dispense with the stationarity restriction. Moreover we are able to deal with optimal control problems with current cost and final cost not necessarily bounded but satisfying weaker L^2 integrability conditions.

References

1. Almudevar, A. A dynamic programming algorithm for the optimal control of piecewise deterministic Markov processes. SIAM Journal on Control and Optimization, 40(2) (2001), 525-539.
2. Bandini, E. Existence and uniqueness for backward stochastic differential equations driven by a random measure. Electronic Communications in Probability, 20(71) (2015), 1-13.
3. Barles G., Buckdahn R., Pardoux E. Backward stochastic differential equations and integral-partial differential equations. Stochastics Stochastics Rep. 60 (1997), no. 1-2, 57-83.
4. Bauerle, N., Rieder, U. MDP algorithms for portfolio optimization problems in pure jump markets, Finance Stoch., 13 (2009), 591-611.
5. Bauerle, N., Rieder, U. Optimal control of piecewise deterministic Markov processes with finite time horizon. In Modern Trends in Controlled Stochastic Processes: Theory and Applications, pages 123-143. Luniver Press, United Kingdom, 2010.
6. Bayraktar, E., Ludkovski, M. Optimal trade execution in illiquid financial markets, Math. Finance, 21(4) (2011), 681-701.
7. Becherer D. Bounded solutions to backward SDEs with jumps for utility optimization and indifference hedging. Ann. of App. Prob. 16 (2006), 2027-2054.
8. Beneš V. E. Existence of Optimal Stochastic Control Laws. SIAM Journal on Control, 9(3) (1971), 446-472.
9. Boel R., Varaiya P., Wong E. Martingales on jump processes; Part I: Representation results; Part II: Applications; SIAM J. Control 13, 999-1061.
10. Boel R., Varaiya P. Optimal control of jump processes; SIAM J. Control 15(1) (1977), 92-119.
11. Brémaud P. Point processes and queues, Martingale dynamics. Springer Series in Statistics. Springer (1981).
12. Chitopekar S. S. Continuous time Markovian sequential control processes. SIAM J. Control 7 (1969), 367-389.
13. Confortola F., Fuhrman M. Backward stochastic differential equations and optimal control of marked point processes. SIAM J. Control Optimization 51(5) (2013), 3592-3623.
14. Confortola F., Fuhrman M. Backward stochastic differential equations associated to jump Markov processes and their applications. Stochastic Processes and their Applications 124 (2014), 289-316.
15. Confortola F., Fuhrman M., Jacod J. Backward stochastic differential equations driven by a marked point process: an elementary approach, with an application to optimal control. The Annals of Applied Probability, 26(3) (2016).
16. Costa O. L., Dufour F. Continuous Average Control of Piecewise Deterministic Markov Processes. Springer Briefs in Mathematics, Springer, 2013.
17. Crépey S., Matoussi A. Reflected and doubly reflected BSDEs with jumps, Ann. of Appl. Prob. 18 (2008), 2041-2069.
18. Davis, M.H.A. Control of piecewise-deterministic processes via discrete-time dynamic programming. In Stochastic differential systems (Bad Honnef, 1985), volume 78 of Lecture Notes in Control and Inform. Sci., Springer, Berlin (1986), 140-150.
19. Davis M.H.A. Markov models and optimization. Monographs on Statistics and Applied Probability 49 (1993), Chapman & Hall.

20. Davis M.H.A., Farid M. Piecewise deterministic processes and viscosity solutions. McEneaney, W. M. et al. (ed) Stochastic Analysis, Control Optimization and Applications. A Volume in Honor of W. H. Fleming on Occasion of His 70th Birthday, Birkhäuser (1999), 249-268.
21. Davis M.H.A., Varaya P. Dynamic programming conditions for patially observable stochastic systems. SIAM J. Control 11 (1973), 226-261.
22. Dempster M. A. H. Optimal control of piecewise deterministic Markov processes. Applied stochastic analysis (London, 1989), 303-325, Stochastics Monogr. 5, Gordon and Breach (New York 1991).
23. Dempster M.A.H., Ye J.J. Necessary and sufficient optimality conditions for control of piecewise deterministic processes. Stochastics and Stochastics Reports, 40(3-4) (1992), 125-145.
24. Dempster, M.A.H., Ye J.J. Generalized Bellman-Hamilton-Jacob optimality conditions for a control problem with boundary conditions. Applied Mathematics and Optimization, 33(3) (1996), 211-225.
25. El Karoui N. Les aspects probabilistes du contrôle stocastique. [The probabilistic aspects of stochastic control] Ninth Saint Flour Probabilistic Summer School,-1979. (Saint Flour, 1979), 73-238, Lecture Notes in Math. 876 (1981), Springer.
26. Elliott R.J. Stochastic Calculus and its Applications. Springer, (1982).
27. Forwick, L., Schäl, M., Schmitz, M. Piecewise deterministic Markov control processes with feedback controls and unbounded costs. Acta Applicandae Mathematicae, 82(3) (2004), 239-267.
28. Gihman I. I., Skorohod A. V. The Theory of Stochastic Processes II. Springer-Verlag Berlin Heidelberg New-York Tokyo (1983).
29. Guilbaud, F., Pham, H. Optimal High Frequency Trading with Limit and Market Orders. Quantitative Finance 13(1) (2011).
30. Guo X., Hernandez-Lerma O. Continuous-time Markov decision processes: theory and applications. Stochastic Modelling and Applied Probability 62 (2009). Springer-Verlag, Berlin.
31. Howard R. A. Dynamic Probabilistic Systems. John Wiley, New York (1971).
32. Jacod J. Multivariate point processes: predictable projection, Radon-Nikodym derivatives, representation of martingales. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 31 (1974/75), 235-253.
33. Jewell W. S. Markov-renewal programming I and II. Operational Res. 11 (1963), 938-971.
34. Kakumanu P. Continuously discounted markov decision model with countable state and action space. Ann. Math. Stat. 42 (1971), 919-926.
35. Kazi-Tani N., Possamaï D., Zhou C. Second-order BSDEs with jumps: formulation and uniqueness, The Annals of Applied Probability, 25(5) (2015), 2867-2908.
36. Kazi-Tani N., Possamaï D., Zhou C. Second order BSDEs with jumps: existence and probabilistic representation for fully-nonlinear PIDEs, Electronic Journal of Probability, 20(65) (2015), 1-31.
37. Kharroubi, I., Langrené, N., Pham, H. A numerical algorithm for fully nonlinear HJB equations: an approach by control randomization. Monte Carlo Methods Appl. 20(2) (2014), 145-165.
38. Kharroubi, I., Langrené, N., Pham, H. Discrete time approximation of fully nonlinear HJB equations via BSDEs with nonpositive jumps. Ann. Appl. Probab. 25 (2015), no. 4, 2301-2338.
39. Kharroubi I., Ma J., Pham H., Zhang J. Backward SDEs with constrained jumps and quasi-variational inequalities. Ann. Probab. 38(2) (2010), 794-840.
40. Lenhart S.-M., Yamada N. Perron's method for viscosity solutions associated with piecewise-deterministic processes. Funkcialaj Ekvacioj 34 (1991), 173-186.
41. Lim T., Quenez M.-C. Exponential utility maximization and indifference price in an incomplete market with defaults. Electronic Journal of Probability 16 (2011), 1434-1464.
42. Osaki S., Mine H. Linear programming algorithms for semi-Markovian decision processes. J. Math. Anal. 22 (1968), 356-381.
43. Pliska, S. Controlled jump processes. Stoch. Proc. Appl. 3 (1975), 259-282.
44. Revuz D., Yor M. Continuous Martingales and Brownian Motion. Grundlehren der mathematischen Wissenschaften. Springer, third edition (1999).

45. Rishel, R. A minimum principle for controlled jump processes. Lect. Notes Econ. Math. Syst. Control Theory, Numerical Methods and Computer Systems Modeling, vol. 107, pp. 493-508. Berlin, Heidelberg, New York: Springer 1975.
46. Ross S. M. Applied Probability Models with Optimization Applications. Holden-Day, San Francisco (1970).
47. Royer M. Backward stochastic differential equations with jumps and related nonlinear expectations. Stochastic Processes and their Applications, 116(10) (2006), 1358-1376.
48. Stone L. D. Necessary and sufficient conditions for optimal control of semi-Markov jump processes. SIAM J. Control Optim. 11(2) (1973), 187-201.
49. Striebel, C. Martingale conditions for the optimal control of continuous-time stochastic systems, 5th Symposium on nonlinear estimation and applications, San Diego (1974).
50. Tang S., Li X. Necessary Conditions for Optimal Control of Systems with Random Jumps. SIAM J. Control Optim. 32 (1994), 1447-1475.
51. Vermes D. Optimal control of piecewise deterministic Markov process. Stochastics 14(3) (1985), 165-207.
52. Xia J. Backward stochastic differential equations with random measures. Acta Mathematicae Applicatae Sinica 16(3) (2000), 225-234.
53. Yushkevich, A. A. On reducing a jump controllable Markov model to a model with discrete time. Theory of Probability and Its Applications, 25(1) (1980), 58-69.
54. Yushkevich, A. A. Bellman inequalities in Markov decision deterministic drift processes. Stochastics, 23(1) (1987), 25-77.