The Annals of Applied Probability
2016, Vol. 26, No. 3, 1743–1773
DOI: 10.1214/15-AAP1132

BACKWARD STOCHASTIC DIFFERENTIAL EQUATION DRIVEN BY A MARKED POINT PROCESS: AN ELEMENTARY APPROACH WITH AN APPLICATION TO OPTIMAL CONTROL

BY FULVIA CONFORTOLA[*,1], MARCO FUHRMAN[*,1] AND JEAN JACOD[†]

Politecnico di Milano[*] and Université Pierre et Marie Curie[†]

We address a class of backward stochastic differential equations on a bounded interval, where the driving noise is a marked, or multivariate, point process. Assuming that the jump times are totally inaccessible and a technical condition holds (see Assumption (A) below), we prove existence and uniqueness results under Lipschitz conditions on the coefficients. Some counter-examples show that our assumptions are indeed needed. We use a novel approach that allows reduction to a (finite or infinite) system of deterministic differential equations, thus avoiding the use of martingale representation theorems and allowing potential use of standard numerical methods. Finally, we apply the main results to solve an optimal control problem for a marked point process, formulated in a classical way.

1. Introduction. Since the paper by Pardoux and Peng [17], the topic of backward stochastic differential equations (BSDE in short) has been in constant development, due to its utility in finance (see, e.g., El Karoui, Peng and Quenez [12]), in control theory, and in the theory of nonlinear PDEs.

The first papers, and most of the subsequent ones, assume that the driving term is a Brownian motion, but the case of a discontinuous driving process has also been considered rather early; see, for example, Buckdahn and Pardoux [4], Tang and Li [19] and more recently Barles, Buckdahn and Pardoux [2], Xia [20], Becherer [3], Crépey and Matoussi [10], or Carbone, Ferrario and Santacroce [5] among many others.

The case of a driving term which is purely discontinuous has attracted less attention; see, however, Shen and Elliott [18] for the particularly simple "one-jump" case, or Cohen and Elliott [6, 7] and Cohen and Szpruch [8] for BSDEs associated to Markov chains. The pure jump case has certainly less potential applications than the continuous or continuous-plus-jumps case, but on the other hand it exhibits a much simpler structure, which provides original insight on BSDEs.

Received July 2014; revised July 2015.

[1]Supported in part by MIUR-PRIN 2010–2011 "Evolution differential problems: deterministic and stochastic approaches and their interactions".

MSC2010 subject classifications. Primary 60H10; secondary 93E20.

Key words and phrases. Backward stochastic differential equations, marked point processes, stochastic optimal control.

To illustrate the latter point, in this paper we consider BSDEs driven by a marked (or, multivariate) point process. The time horizon is a finite (nonrandom) time T. The point process is nonexplosive, that is, there are almost surely finitely many points within the interval $[0, T]$, and it is also *quasi-left continuous*, that is, the jump times are totally inaccessible: the main examples of this situation are the Poisson process and the compound Poisson process. We also make the (rather strong) assumption that the generator is uniformly Lipschitz.

In contrast with most of the literature, in which the martingale representation theorem and the application of a suitable fixed-point theorem play a central role, in the setting of point processes it is possible to solve the equation recursively, by replacing the BSDE by an ordinary differential equation in between jumps, and match the pre- and post-jump values at each jump time (such a method has already been used for a BSDE driven by a Brownian motion plus a Poisson process; see, e.g., Kharroubi and Lim [16], but then between any two consecutive jumps one has to solve a genuine BSDE).

Reducing the BSDE to a sequence of ODEs allows us for a very simple solution, although we still need some elementary a priori estimates, though, for establishing the existence when the number of jumps is unbounded. Apart from the intrinsic interest of a simple method, this might also give rise to simple numerical ways for solving the equation. Another noticeable point is that it provides an $\mathbf{L}^1$ theory, which is more appropriate for point processes than the usual $\mathbf{L}^2$ theory.

There are two main results about the BSDE: one is when the number of jumps is bounded, and then we obtain uniqueness within the class of all possible solutions. The other is, in the general case, an existence and uniqueness result within a suitable weighted $\mathbf{L}^1$ space. We also state a third important result, showing how an optimal control problem on a marked process reduces to solving a BSDE. Existence and uniqueness results for the BSDE are stated in the case of a scalar equation, but the extension to the vector-valued case is immediate.

The paper is organized as follows: in Section 2, we present the setting and the two main results (as will be seen, the setting is somewhat complicated to explain, because in the multivariate case there are several distinct but natural versions for the BSDE). Section 3 is devoted to a few simple a priori estimates. In Section 4, we explain how the BSDE can be reduced to a sequence of (nonrandom) ODEs, and also exhibit a few counter-examples when the basic assumptions on the point process are violated. The proof of the main results is in Section 5, and in Section 6 the control problem is considered.

2. Main results.

2.1. *The setting.* We have a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a fixed time horizon $T \in (0, \infty)$, so all processes defined on this space are indexed by $[0, T]$, and all random times take their values in $[0, T] \cup \{\infty\}$.

This space is endowed with a nonexplosive multivariate point process (also called marked point process) on $[0,T] \times E$, where $(E, \mathcal{E})$ is a Lusin space: this is a sequence (S_n, X_n) of points, with distinct times of occurrence S_n and with marks X_n, so it can be viewed as a random measure of the form

$$\mu(dt, dx) = \sum_{n \geq 1 : S_n \leq T} \varepsilon_{(S_n, X_n)}(dt, dx), \tag{2.1}$$

where $\varepsilon_{(t,x)}$ denotes the Dirac measure. Here, the S_n's are $(0,T] \cup \{\infty\}$-valued and the X_n's are E-valued, and $S_1 > 0$, and $S_n < S_{n+1}$ if $S_n \leq T$, and $S_n \leq S_{n+1}$ everywhere, and $\Omega = \bigcup\{S_n > T\}$. Note that the "mark" X_n is relevant on the set $\{S_n \leq T\}$ only, but it is convenient to have it defined on the whole set Ω, and without restriction we may assume that $X_n = \Delta$ when $S_n = \infty$, where Δ is a distinguished point in E.

We denote by $(\mathcal{F}_t)_{t \geq 0}$ the filtration generated by the point process, which is the smallest filtration for which each S_n is a stopping time and X_n is $\mathcal{F}_{S_n}$-measurable. As we will see, the special structure of this filtration plays a fundamental role in all what follows. We let $\mathcal{P}$ be the predictable σ-field on $\Omega \times [0,T]$, and for any auxiliary measurable space $(G, \mathcal{G})$ a function on the product $\Omega \times [0,T] \times G$ which is measurable with respect to $\mathcal{P} \otimes \mathcal{G}$ is called *predictable*.

We denote by ν the predictable compensator of the measure μ, relative to the filtration $(\mathcal{F}_t)$. The measure ν admits the disintegration

$$\nu(\omega, dt, dx) = dA_t(\omega)\phi_{\omega,t}(dx), \tag{2.2}$$

where ϕ is a transition probability from $(\Omega \times [0,T], \mathcal{P})$ into $(E, \mathcal{E})$, and A is an increasing càdlàg predictable process starting at $A_0 = 0$, which is also the predictable compensator of the univariate point process

$$N_t = \mu\big([0,t] \times E\big) = \sum_{n \geq 1} 1_{\{S_n \leq t\}}. \tag{2.3}$$

Of course, the multivariate point process μ reduces to the univariate N when E is a singleton.

Unless otherwise specified, the following assumption, where we set $S_0 = 0$, will hold throughout.

ASSUMPTION (A). (A_1) The process A is continuous (equivalently: the jump times S_n are totally inaccessible).

(A_2) $\mathbb{P}(S_{n+1} > T \mid \mathcal{F}_{S_n}) > 0$ for all $n \geq 0$.

The first condition amounts to the quasi-left continuity of N. We will briefly examine what happens when (A_1) and (A_2) fail in Section 4.

2.2. *The BSDE in the univariate case.* Now, we turn to the BSDE. In addition to the driving point process, the ingredients are:

- a *terminal condition* ξ, which is always an $\mathcal{F}_T$-measurable random variable;
- a *generator* f, which is real-valued function depending on ω, on time, possibly on the mark x of the point process, and also in a suitable way on the solution of the BSDE. In all cases below, the dependence of the generator upon the solution will be assumed Lipschitz, typically involving two nonnegative constants L, L', as specified below.

We begin with the univariate case, which is simpler to formulate. In this case, the BSDE takes the form

$$Y_t + \int_{(t,T]} Z_s \, dN_s = \xi + \int_{(t,T]} f(\cdot, s, Y_{s-}, Z_s) \, dA_s, \tag{2.4}$$

where f is a predictable function on $\Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}$, satisfying

$$\begin{gathered} |f(\omega, t, y', z') - f(\omega, t, y, z)| \le L'|y' - y| + L|z' - z|, \\ \int_0^T |f(t, 0, 0)| \, dA_t < \infty \qquad \text{a.s.} \end{gathered} \tag{2.5}$$

A *solution* is a pair (Y, Z) consisting in an adapted càdlàg process Y and a predictable process Z satisfying $\int_0^T |Z_t| \, dA_t < \infty$ a.s., such that (2.4) holds for all $t \in [0, T]$, outside a $\mathbb{P}$-null set [this implicitly supposes that $\int_0^T |f(\cdot, s, Y_s, Z_s)| \, dA_s < \infty$ a.s.].

REMARK 1. Quite often the BSDE is written, in a slightly different form, as

$$Y_t + \int_{(t,T]} Z_s(dN_s - dA_s) = \xi + \int_{(t,T]} f(\cdot, s, x, Y_{s-}, Z_s) \, dA_s. \tag{2.6}$$

Upon a trivial modification of f, this is clearly the same as (2.4), and it explains the integrability restriction on Z. The reason underlying the formulation (2.6) is that it singles out the "martingale increment" $\int_{(t,T]} Z_s(dN_s - dA_s)$.

2.3. *The BSDE in the multivariate case.* In the multivariate case, the predictable process Z of (2.4) should be replaced by a predictable function $Z(\omega, t, x)$ on $\Omega \times [0, T] \times E$, and this function may enter the generator in different guises. We start with the most general formulation, and will single out two special, easier to formulate, cases afterward.

We need some additional notation: we let $\mathcal{B}(E)$ be the set of all Borel functions on E; if Z is a measurable function on $\Omega \times [0, T] \times E$, we write $Z_{\omega,t}(x) = Z(\omega, t, x)$, so each $Z_{\omega,t}$, often abbreviated as Z_t or $Z_t(\cdot)$, is an element of $\mathcal{B}(E)$.

With this notation, the BSDE takes the form

$$
\begin{aligned}
&Y_t + \int_{(t,T]} \int_E Z(s,x)\mu(ds,dx) \\
&\quad = \xi + \int_{(t,T]} \int_E f\big(\cdot, s, x, Y_{s-}, Z_s(\cdot)\big)\nu(ds,dx),
\end{aligned}
\tag{2.7}
$$

where f is a real-valued function on $\Omega \times [0,T] \times E \times \mathbb{R} \times \mathcal{B}(E)$, such that $f(\omega,t,x,y,Z_{\omega,t}(\cdot))$ is predictable for any predictable function Z on $\Omega \times [0, T] \times E$, and

$$
\begin{aligned}
&\big|f(\omega,t,x,y',\zeta) - f(\omega,t,x,y,\zeta)\big| \le L'|y'-y|, \\
&\int_E \big|f(\omega,t,x,y,\zeta) - f(\omega,t,x,y,\zeta')\big|\phi_{\omega,t}(dx) \\
&\qquad \le L \int_E |\zeta'(x) - \zeta(x)|\phi_{\omega,t}(dx), \\
&\int_0^T \int_E \big|f(t,x,0,0)\big|\nu(dt,dx) < \infty \qquad \text{a.s.}
\end{aligned}
\tag{2.8}
$$

[in the expression $f(t,x,0,0)$, the last "0" stands for the function in $\mathcal{B}(E)$ which vanishes identically].

A *solution* is a pair (Y,Z) consisting in an adapted càdlàg process Y and a predictable function Z on $\Omega \times [0,T] \times E$ satisfying $\int_0^T \int_E |Z(t,x)|\nu(ds,dx) < \infty$ a.s., such that (2.7) holds for all $t \in [0,T]$, outside a $\mathbb{P}$-null set.

The measurability condition imposed on the generator is somewhat awkward, and probably difficult to check in general. However, it is satisfied in the following two types of equations.

Type I equation: This is the simplest one to state, and it takes the form

$$
\begin{aligned}
&Y_t + \int_{(t,T]} \int_E Z(s,x)\mu(ds,dx) \\
&\quad = \xi + \int_{(t,T]} \int_E f_I\big(\cdot, s, x, Y_{s-}, Z(s,x)\big)\nu(ds,dx),
\end{aligned}
\tag{2.9}
$$

where f_I is a predictable function on $\Omega \times [0,T] \times E \times \mathbb{R} \times \mathbb{R}$, satisfying

$$
\begin{aligned}
&\big|f_I(\omega,t,x,y',z') - f_I(\omega,t,x,y,z)\big| \le L'|y'-y| + L|z'-z|, \\
&\int_0^T \int_E \big|f_I(t,x,0,0)\big|\nu(dt,dx) < \infty \qquad \text{a.s.}
\end{aligned}
\tag{2.10}
$$

That (2.9) is a special case of (2.7) is obvious; we simply have to take for f the function on $\Omega \times [0,T] \times E \times \mathbb{R} \times \mathcal{B}(E)$ defined by

$$
f(\omega,s,x,y,\zeta) = f_I\big(\omega,s,x,y,\zeta(x)\big),
\tag{2.11}
$$

and (2.10) for f_I yields (2.8) for f.

Type II equations: The BSDE (2.9) cannot in general be used as a tool for solving control problems driven by a multivariate point process, whereas this is one of the main motivations for introducing them. We rather need the following formulation:

$$Y_t + \int_{(t,T]} \int_E Z(s,x)\mu(ds,dx) = \xi + \int_{(t,T]} f_{\mathrm{II}}(\cdot, s, Y_{s-}, \eta_s Z_s)\, dA_s, \tag{2.12}$$

where, recalling that $\phi_{\omega,t}$ are the measures occurring in (2.2) and $Z_{\omega,t}(x) = Z(\omega,t,x)$,

$$\begin{aligned} &\eta_{\omega,t} \text{ is a real-valued map on } \mathcal{B}(E), \\ &\qquad \text{with } \left|\eta_{\omega,t}\zeta - \eta_{\omega,t}\zeta'\right| \le \int_E \left|\zeta'(v) - \zeta(v)\right| \phi_{\omega,t}(dv), \\ &Z \text{ predictable on } \Omega \times [0,T] \times E \quad \Rightarrow \\ &\qquad \text{the process } (\omega,t) \mapsto \eta_{\omega,t} Z_{\omega,t} \text{ is predictable,} \\ &f_{\mathrm{II}} \text{ is a function satisfying (2.5).} \end{aligned} \tag{2.13}$$

Again, (2.12) reduces to (2.7) upon taking

$$f(\omega, s, x, y, \zeta) = f_{\mathrm{II}}(\omega, s, y, \eta_{\omega,s}\zeta), \tag{2.14}$$

and (2.5) for f_{II} plus (2.13) for $\eta_{\omega,t}$ yield (2.8) for f. As we will see in Section 6, this type of equation is well suited to control problem.

In the univariate case, all three formulations (2.7), (2.9) and (2.12) coincide with (2.4).

Finally, we describe another notion of a solution, starting with the following remark: we can of course rewrite (2.7) as follows:

$$\begin{aligned} &Y_t + \sum_{n \ge 1} Z(S_n, X_n) 1_{\{t < S_n \le T\}} \\ &\qquad = \xi + \int_{(t,T]} \int_E f\big(s, x, Y_{s-}, Z_s(\cdot)\big) \nu(ds,dx). \end{aligned} \tag{2.15}$$

Since A is continuous, (2.15) yields, outside a $\mathbb{P}$-null set,

$$\begin{aligned} &\Delta Y_{S_n} = Z(S_n, X_n) \qquad \text{if } S_n \le T \text{ and } n \ge 1, \\ &Y \text{ is continuous outside } \{S_1, \dots, S_n, \dots\}. \end{aligned} \tag{2.16}$$

In other words, Y completely determines the predictable function Z outside a null set with respect to the measure $\mathbb{P}(d\omega)\mu(\omega, dt, dx)$, hence also outside a $\mathbb{P}(d\omega)\nu(\omega, dt, dx)$-null set. Equivalently, if (Y, Z) is a solution and Z' is another predictable function, then (Y, Z') being another solution is the same as having $Z' = Z$ outside a $\mathbb{P}(d\omega)\mu(\omega, dt, dx)$-null set, and the same as having $Z' = Z$ outside a $\mathbb{P}(d\omega)\nu(\omega, dt, dx)$-null set.

Therefore, another way of looking at equation (2.7) is as follows: a *solution* is an adapted càdlàg process Y for which there exists a predictable function Z satisfying

$$\int_0^T \int_E |Z(s,x)| \nu(ds,dx) < \infty \qquad \text{a.s.},$$

such that the pair (Y, Z) satisfies (2.7) for all $t \in [0, T]$, outside a $\mathbb{P}$-null set. Then *uniqueness* of the solution means that, for any two solutions Y and Y' we have $Y_t = Y'_t$ for all $t \in [0, T]$, outside a $\mathbb{P}$-null set.

2.4. *Statement of the main results*. We have two main results. The first one is when the point process has at most M points, for a nonrandom integer M, that is,

$$\mathbb{P}(S_{M+1} = \infty) = 1. \tag{2.17}$$

THEOREM 2. *Assume* (A) *and* (2.17). *The solution Y of* (2.7), *if it exists*, *is unique up to null sets*. *Moreover*, *if the variable A_T is bounded*, *and if*

$$\mathbb{E}(|\xi|) < \infty, \qquad \mathbb{E}\left(\int_0^T \int_E |f(s,x,0,0)| \nu(ds,dx)\right) < \infty, \tag{2.18}$$

the solution exists and satisfies $\mathbb{E}(\int_0^T |Y_t|\,dA_t) < \infty$ *and* $\mathbb{E}(\int_0^T \int_E |Z(t,x)| \nu(dt, dx)) < \infty$.

The existence result above is "almost" a special case of the next theorem. In contrast, the uniqueness within the class of *all* possible solutions is specific to the situation (2.17). When this fails, uniqueness holds only within smaller subclasses, which we now describe. For any $\alpha > 0$ and $\beta \geq 0$, we set

$$\begin{aligned} \mathcal{L}^1_{\alpha,\beta} &= \text{the set of all pairs } (Y,Z) \text{ with } Y \text{ càdlàg adapted and} \\ &\quad\; Z \text{ predictable, satisfying,} \\ \|(Y,Z)\|_{\alpha,\beta} &:= \mathbb{E}\left(\int_0^T \int_E (|Y_t| + |Z(t,x)|) e^{\beta A_t} \alpha^{N_t} \nu(dt,dx)\right) < \infty. \end{aligned} \tag{2.19}$$

The space $\mathcal{L}^1_{\alpha,\beta}$ decreases when α and/or β increases.

THEOREM 3. *Assume* (A).

(a) *If*

$$\begin{gathered} \mathbb{E}(e^{\beta A_T} \alpha^{N_T} |\xi|) < \infty, \\ \mathbb{E}\left(\int_0^T \int_E \alpha^{N_s} e^{\beta A_s} |f(s,x,0,0)| \nu(ds,dx)\right) < \infty, \end{gathered} \tag{2.20}$$

for some $\alpha > L$ and $\beta > 1 + \alpha + L'$, where L, L' are the constants occurring in (2.8), *then* (2.7) *admits one and only one* (*up to null sets*) *solution* (Y, Z) *belonging to* $\mathcal{L}^1_{\alpha,\beta}$.

(b) *When moreover the variable A_T is bounded, the conditions*

$$\mathbb{E}(|\xi|^{1+\varepsilon}) < \infty, \qquad \mathbb{E}\biggl(\biggl(\int_0^T \int_E |f(s,x,0,0)|\nu(ds,dx)\biggr)^{1+\varepsilon}\biggr) < \infty \tag{2.21}$$

for some $\varepsilon > 0$ imply (2.20) for all $\beta \ge 0$ and $\alpha > 0$, hence (2.7) admits one and only one (up to null sets) solution (Y, Z) belonging to $\bigcup_{\alpha>L,\beta>1+\alpha+L'} \mathcal{L}^1_{\alpha,\beta}$, and this solution also belongs to $\bigcap_{\alpha>0,\beta\ge 0} \mathcal{L}^1_{\alpha,\beta}$.

The claim (b) is interesting, because it covers the most usual situation where μ is a Poisson random measure (so that $A_t = \lambda t$ for some constant $\lambda > 0$). Note that, even in this case, we do not know whether (2.7) admits other solutions, which are not in $\bigcup_{\alpha>L,\beta>1+\alpha+L'} \mathcal{L}^1_{\alpha,\beta}$.

We note that if we apply Theorem 3 with the assumptions of Theorem 2, namely $A_T \le K$ and $N_T \le M$, condition (2.20) is equivalent to (2.18) since the exponential factors are bounded. In this sense, Theorem 2 is a special case of Theorem 3, except that in the latter theorem uniqueness is guaranteed only within the smaller class $\mathcal{L}^1_{\alpha,\beta}$. The occurrence of exponential weights in the definition of the norm in this space is due to the fact that we are dealing with BSDEs driven by a general random compensator $\nu(\omega, dt, dx) = dA_t(\omega)\phi_{\omega,t}(dx)$, where A is an increasing but not necessarily bounded predictable processes. The same happens in the $\mathbf{L}^2$ theory for BSDEs associated to marked point processes (see [9, 20]) and for BSDEs driven by a general càdlàg martingale (see [11]). On the other hand, in case of compensators absolutely continuous with respect to a deterministic measure, [3, 10, 19], a standard $\mathbf{L}^2$ theory holds (the norm reduces to a simpler form, not involving exponentials of stochastic processes).

3. A priori estimates. In this section, we provide some a priori estimates for the solutions of equation (2.7). Without special mention, Assumption (A_1) is assumed throughout.

LEMMA 4. *Let $\alpha > 0$ and $\beta \in \mathbb{R}$. If (Y, Z) is a solution of (2.7) we have almost surely*

$$\begin{aligned} &|Y_t|e^{\beta A_t}\alpha^{N_t} + \int_t^T \int_E (\alpha|Y_{s-} + Z(s,x)| - |Y_{s-}|)e^{\beta A_s}\alpha^{N_{s-}}\mu(ds,dx) \\ &\quad + \beta \int_t^T |Y_s|e^{\beta A_s}\alpha^{N_s}\,dA_s \\ &\qquad = |\xi|^p e^{\beta A_T}\alpha^{N_T} + \int_t^T \int_E \operatorname{sign}(Y_s) f(s,x,Y_s,Z_s(\cdot))e^{\beta A_s}\alpha^{N_s}\nu(ds,dx). \end{aligned} \tag{3.1}$$

PROOF. Letting U_t and V_t be the left-hand and right-hand sides of (3.1), and since these processes are càdlàg, and continuous outside the S_n's, and $U_T = V_T$, it suffices to check that outside a null set we have $\Delta U_{S_n} = \Delta V_{S_n}$ and also $U_t - U_s = V_t - V_s$ if $S_n \le t < s < S_{n+1} \wedge T$, for all $n \ge 0$. The first property is obvious because $\Delta Y_{S_n} = Z(S_n, X_n)$ a.s. and A is continuous. The second property follows from $Y_t - Y_s = \int_t^s \int_E f(v, x, Y_v, Z_v(\cdot))\nu(dv, dx)$, implying $|Y_t| - |Y_s| = \int_t^s \int_E \operatorname{sign}(Y_v) f(v, x, Y_v, Z_v(\cdot))\nu(dv, dx)$ and $\alpha^{N_v} = \alpha^{N_t}$ for all $v \in [t, s]$, plus a standard change of variables formula. $\square$

For any $\alpha > 0$ and $\beta \ge 0$, and with any measurable process Y and measurable function Z on $\Omega \times [0, T] \times E$ we set for $0 \le t < s \le T$

$$\mathcal{W}^{\alpha,\beta}_{(t,s]}(Y, Z) = \int_t^s \int_E (|Y_v| + |Z(v, x)|) e^{\beta A_v} \alpha^{N_v} \nu(dv, dx), \tag{3.2}$$

so with the notation (2.19) we have $\|(Y, Z)\|_{\alpha,\beta} = \mathbb{E}(\mathcal{W}^{\alpha,\beta}_{(0,T]}(Y, Z))$. Below, L and L' are as in (2.8).

LEMMA 5. *Let $\alpha > L$ and $\beta > 1 + \alpha + L'$. There is a constant C only depending on (α, β, L, L'), such that any pair (Y, Z) in $\mathcal{L}^1_{\alpha,\beta}$ which solves (2.7) satisfies, for any stopping time S with $S \le T$ and outside a null set,*

$$\begin{aligned} &|Y_S| e^{\beta A_S} \alpha^{N_S} \\ &\quad \le \mathbb{E}\left(|\xi| e^{\beta A_T} \alpha^{N_T} + \int_S^T \int_E |f(s, x, 0, 0)| e^{\beta A_s} \alpha^{N_s} \nu(ds, dx) \Big| \mathcal{F}_S \right), \end{aligned} \tag{3.3}$$

$$\begin{aligned} &\mathbb{E}\big(\mathcal{W}^{\alpha,\beta}_{(S,T]}(Y, Z) | \mathcal{F}_S\big) \\ &\quad \le C\mathbb{E}\left(|\xi| e^{\beta A_T} \alpha^{N_T} + \int_S^T \int_E |f(s, x, 0, 0)| e^{\beta A_s} \alpha^{N_s} \nu(ds, dx) \Big| \mathcal{F}_S \right). \end{aligned} \tag{3.4}$$

PROOF. We have $\alpha|Y_{s-} + Z(s, x)| - |Y_{s-}| \ge \alpha|Z(s, x)| - (1 + \alpha)|Y_{s-}|$, hence (3.1), and the Lipschitz condition (2.8) plus the fact that $\phi_{t,\omega}(E) = 1$ yield almost surely

$$\begin{aligned} &|Y_S| e^{\beta A_S} \alpha^{N_S} + \alpha \int_S^T \int_E |Z(s, x)| e^{\beta A_s} \alpha^{N_{s-}} \mu(ds, dx) + \beta \int_S^T |Y_s| e^{\beta A_s} \alpha^{N_s} \, dA_s \\ &\le |\xi| e^{\beta A_T} \alpha^{N_T} + (1 + \alpha) \int_S^T |Y_{s-}| e^{\beta A_s} \alpha^{N_{s-}} \, dN_s \\ &\quad + \int_S^T \int_E (|f(s, x, 0, 0)| + L'|Y_s| + L|Z(s, x)|) e^{\beta A_s} \alpha^{N_s} \nu(ds, dx). \end{aligned} \tag{3.5}$$

Since $\mathbb{E}(\int_S^T \int_E \psi(s, x)\mu(ds, dx) | \mathcal{F}_S) = \mathbb{E}(\int_S^T \int_E \psi(s, x)\nu(ds, dx) | \mathcal{F}_S)$ for any nonnegative predictable function ψ, taking the $\mathcal{F}_S$-conditional expectation in (3.5) yields

$$
\begin{aligned}
&|Y_S|e^{\beta A_S}\alpha^{N_S} + \mathbb{E}\biggl(\int_S^T\int_E \bigl(\alpha|Z(s,x)| + \beta|Y_s|\bigr)e^{\beta A_s}\alpha^{N_s}\nu(ds,dx)\Big|\mathcal{F}_S\biggr) \\
&\quad\le \mathbb{E}\bigl(|\xi|e^{\beta A_T}\alpha^{N_T}\bigr) \\
&\qquad + \mathbb{E}\biggl(\int_S^T\int_E \bigl(|f(s,x,0,0)| + (1+\alpha+L')|Y_{s-}| + L|Z(s,x)|\bigr) \\
&\qquad \times e^{\beta A_s}\alpha^{N_s}\nu(ds,dx)\Big|\mathcal{F}_S\biggr).
\end{aligned}
$$

When $\mathbb{E}(\mathcal{W}^{\alpha,\beta}_{(0,T]}(Y,Z)) < \infty$, this implies almost surely

$$
\begin{aligned}
&|Y_S|e^{\beta A_S}\alpha^{N_S} + \mathbb{E}\biggl(\int_S^T\int_E \bigl((\beta-1-\alpha-L')|Y_s| + (\alpha-L)|Z(s,x)|\bigr) \\
&\quad \times e^{\beta A_s}\alpha^{N_s}\nu(ds,dx)\Big|\mathcal{F}_S\biggr) \\
&\qquad \le \mathbb{E}\biggl(|\xi|e^{\beta A_T}\alpha^{N_T} + \int_S^T\int_E |f(s,x,0,0)|e^{\beta A_s}\alpha^{N_s}\nu(ds,dx)\Big|\mathcal{F}_S\biggr),
\end{aligned}
$$

giving us both (3.3) and (3.4). $\square$

LEMMA 6. *Let $\alpha > L$ and $\beta > 1+\alpha+L'$. If (Y,Z) is a solution of* (2.7) *and (Y',Z') is a solution of the same equation with the same generator f and another terminal condition ξ', both pairs (Y,Z) and (Y',Z') being in $\mathcal{L}^1_{\alpha,\beta}$, we have for any stopping time S with $S \le T$ and outside a null set*

$$
|Y'_S - Y_S|e^{\beta A_S}\alpha^{N_S} \le \mathbb{E}\bigl(|\xi'-\xi|e^{\beta A_T}\alpha^{N_T}|\mathcal{F}_S\bigr), \tag{3.6}
$$

$$
\mathbb{E}\bigl(\mathcal{W}^{\alpha,\beta}_{(0,T]}(Y'-Y, Z'-Z)\bigr) \le C\mathbb{E}\bigl(|\xi'-\xi|e^{\beta A_T}\alpha^{N_T}\bigr). \tag{3.7}
$$

In particular, (2.7) *admits*, *up to null sets*, *at most one solution (Y,Z) belonging to $\mathcal{L}^1_{\alpha,\beta}$.*

PROOF. Set [with ζ arbitrary in $\mathcal{B}(E)$, and recalling the notation $Z_{\omega,t}(x) = Z(\omega,t,x)$]

$$
\begin{aligned}
&\overline{Y} = Y' - Y, \qquad \overline{Z} = Z' - Z, \qquad \overline{\xi} = \xi' - \xi, \\
&\overline{f}(\omega,s,x,y,\zeta) \\
&\qquad = f\bigl(\omega,s,x,Y_{s-}(\omega)+y, Z_{\omega,s}(\cdot)+\zeta\bigr) - f\bigl(\omega,s,x,Y_{s-}(\omega),Z_{\omega,s}(\cdot)\bigr).
\end{aligned}
$$

Then $\overline{f}$ is satisfies (2.8) with the same constants L, L', and also $\overline{f}(s,x,0,0) = 0$, and clearly $(\overline{Y},\overline{Z})$ belongs to $\mathcal{L}^1_{\alpha,\beta}$ and satisfies (2.7) with the generator $\overline{f}$ and the terminal condition $\overline{\xi}$. Hence, (3.6) and (3.7) are exactly (3.3) and (3.4) written for $(\overline{Y},\overline{Z})$.

Finally, the last claim follows by taking $\xi' = \xi$. $\square$

4. The structure of the solutions. In this section, we show how it is possible to reduce the problem of solving equation (2.7) to solving a sequence of ordinary differential equations. This reduction needs a number of rather awkward notation, but it certainly has interest in its own sake. Except in the last subsection, devoted to some counter-examples, we assume (A). We stress that both A_1 and A_2 are crucial here, in particular to characterize the $\mathcal{F}_{S_n}$-conditional law of (S_{n+1}, X_{n+1}) and the compensator ν of μ.

4.1. *Some basic facts.* Recall that (S_n, X_n) takes its values in the set $\mathcal{S} = ([0, T] \times E) \cup \{(\infty, \Delta)\}$. For any integer $n \geq 0$, we let H_n be the subset of $\mathcal{S}^{n+1}$ consisting in all $D = ((t_0, x_0), \dots, (t_n, x_n))$ satisfying

$$\begin{aligned} t_0 = 0, \qquad x_0 = \Delta, \qquad t_{j+1} \geq t_j, \qquad t_j \leq T \\ \Rightarrow \quad t_{j+1} > t_j, \qquad t_j > T \\ \Rightarrow \quad (t_j, x_j) = (\infty, \Delta). \end{aligned}$$

We set $D^{\max} = t_n$ and endow H_n with its Borel σ-field $\mathcal{H}_n$. We set $S_0 = 0$ and $X_0 = \Delta$, so

$$D_n = \big((S_0, X_0), \dots, (S_n, X_n)\big) \tag{4.1}$$

is a random element with values in H_n, whose law is denoted as Λ_n [a probability measure on $(H_n, \mathcal{H}_n)$].

The filtration $(\mathcal{F}_t)$ generated by the point process μ has a very special structure, which reflects on adapted or predictable processes, and below we explain some of these properties; see [13] for more details. They might look complicated at first glance, but they indeed allow us to replace random elements by deterministic functions of all the D_n's.

(a) *The variable* ξ: Since ξ is $\mathcal{F}_T$-measurable, for each $n \geq 0$ there is an $\mathcal{H}_n$-measurable map $D \mapsto u^n_D$ on H_n with

$$\begin{aligned} & D^{\max} = \infty \quad \Rightarrow \quad u^n_D = 0, \\ & S_n(\omega) \leq T < S_{n+1}(\omega) \quad \Rightarrow \quad \xi(\omega) = u^n_{D_n(\omega)}. \end{aligned} \tag{4.2}$$

(b) *Adapted càdlàg processes*: A càdlàg process Y, which further is continuous outside the times S_n, is adapted if and only if for each $n \geq 0$ there is a Borel function $y^n = y^n_D(t)$ on $H_n \times [0, T]$ such that

$$\begin{aligned} & D^{\max} = \infty \quad \Rightarrow \quad y^n_D(t) = 0, \\ & t \mapsto y^n_D(t) \text{ is continuous on } [0, T] \text{ and constant on } \big[0, T \wedge D^{\max}\big], \\ & S_n(\omega) \leq t < S_{n+1}(\omega), \qquad t \leq T \quad \Rightarrow \quad Y_t(\omega) = y^n_{D_n(\omega)}(t), \end{aligned} \tag{4.3}$$

and we express this as $Y \equiv (y^n)$.

(c) *Predictable functions*: A function Z on $\Omega \times [0,T] \times E$ is predictable if and only if for each $n \geq 0$ there is a Borel function $z^n = z^n_D(t,x)$ on $H_n \times [0,T] \times E$ such that

$$
\begin{aligned}
&D^{\max} = \infty \quad \Rightarrow \quad z^n_D(t,x) = 0, \\
&S_n(\omega) < t \leq S_{n+1}(\omega) \wedge T \quad \Rightarrow \quad Z(\omega,t,x) = z^n_{D_n(\omega)}(t,x).
\end{aligned}
\tag{4.4}
$$

We express this as $Z \equiv (z^n)$, and also write $z^n_{D,t}$ for the function $z^n_{D,t}(x) = z^n_D(t,x)$ on E.

(d) *The $\mathcal{F}_{S_n}$-conditional law of (S_{n+1}, X_{n+1})*: This conditional law takes the form $G^n_{D_n}$, where $G^n_D(dt,dx)$ is a transition probability from H_n into $[0,\infty] \times E$, and upon using (A) we may further assume the following structure on G^n_D, where $\phi^n_{D,t}(dx)$ is a transition probability from $H_n \times [0,\infty]$ into E:

$$
\begin{aligned}
&G^n_D(dt,dx) = G'^n_D(dt)\phi^n_{D,t}(dx) \qquad \text{where } G'^n_D(dt) = G^n_D(dt,E), \\
&G'^n_D\big((T,\infty)\big) = 0, \qquad t > T \quad \Rightarrow \quad \phi^n_{D,t}(dx) = \varepsilon_\Delta(dx), \\
&t \mapsto g^n_D(t) := G'^n_D\big((t,\infty]\big) \text{ is continuous } \big(\text{by } (\mathrm{A}_1)\big), \\
&g^n_D(T) > 0 \qquad \big(\text{by } (\mathrm{A}_2)\big), \\
&D^{\max} < \infty \quad \Rightarrow \quad g^n_D\big(D^{\max}\big) = 1.
\end{aligned}
\tag{4.5}
$$

The last property $D^{\max} < \infty \Rightarrow g^n_D(D^{\max}) = 1$, which plays an important role later, simply expresses the fact that $S_{n+1} > S_n$ if $S_n < \infty$.

(e) *The compensator ν of μ*: The following gives us versions of ν and A and $\phi_{\omega,t}$ in (2.2):

$$
\begin{aligned}
&\nu(\omega;dt,dx) = \sum_{n=0}^\infty \nu^n_{D_n(\omega)}(dt,dx) 1_{\{S_n < t \leq S_{n+1} \wedge T\}}, \\
&\nu^n_D(dt,dx) = \frac{1}{g^n_D(t)} G^n_D(dt,dx), \\
&S_n(\omega) < t \leq S_{n+1}(\omega) \quad \Rightarrow \quad \phi_{\omega,t} = \phi^n_{D_n(\omega),t}, \\
&A_t(\omega) = \sum_{n=0}^\infty a^n_{D_n(\omega)}\big(t \wedge S_{n+1}(\omega)\big), \qquad a^n_D(t) = -\log g^n_D(t),
\end{aligned}
\tag{4.6}
$$

hence $a^n_D(t) = 0$ for $t \leq D^{\max}$, and $a^n_D(T) < \infty$.

(f) *The generator*: Recall that we are interested in equation (2.7), so by (2.8) the generator f has a nice predictability property only after plugging in a predictable function Z. This implies that, for any $n \geq 0$, and if $z^n = z^n_D(t,x)$ is as in (c) above, one has a Borel function $f\{z^n\}^n = f\{z^n\}^n_D(t,x,y,w)$ on $H_n \times [0,T] \times E \times \mathbb{R} \times \mathbb{R}$, such that (with $t \le T$ below)

$$
\begin{aligned}
&D^{\max} = \infty \quad \Rightarrow \quad f\{z^n\}^n_D(t,x,y) = 0,\\
&S_n(\omega) < t \le S_{n+1}(\omega), \qquad \zeta(x) = w + z^n_{D_n(\omega)}(t,x) \\
&\qquad \Rightarrow \quad f(\omega,t,x,y,\zeta) = f\{z^n\}^n_{D_n(\omega)}(t,x,y,w).
\end{aligned} \tag{4.7}
$$

Moreover, the last two conditions in (2.8) imply that one can take a version which satisfies identically (where z^n and z'^n are two terms as in (c), and $f\{0\}^n_D$ below is $f\{z^n\}^n_D$ for $z^n_D(t,x) \equiv 0$)

$$
\begin{aligned}
&\big|f\{z^n\}^n_D(t,x,y',w') - f\{z'^n\}^n_D(t,x,y,w)\big| \\
&\quad \le L'|y'-y| + L|w'-w| + L\int_E |z'^n_D(t,v) - z^n_D(t,v)|\phi^n_{D,t}(dv)
\end{aligned} \tag{4.8}
$$

$$
\int_0^T |f\{0\}^n_D(t,x,0,0)|\nu^n_D(dt,dx) < \infty.
$$

4.2. *Reduction to ordinary differential equations.* By virtue of (2.16), if $Y \equiv (y^n)$ is a solution of (2.7), we can, and always will, take for the associated process $Z \equiv (z^n)$ the one defined for $t \in [0,T]$ by

$$
z^n_D(t,x) = y^{n+1}_{D\cup\{(t,x)\}}(t) 1_{\{t>D^{\max}\}} - y^n_D(t), \tag{4.9}
$$

because $Y_{S_{n+1}} = y^{n+1}_{D_n\cup\{(S_{n+1},X_{n+1})\}}(S_{n+1})$ and $Y_{S_{n+1}-} = y^n_{D_n}(S_{n+1})$, when $S_{n+1} \le T$. We will in fact write the above in another form, suitable for plugging into the generator f, as represented by (4.7). Namely, we set

$$
\begin{aligned}
&\widehat{y}^{n+1} = \big(\widehat{y}^{n+1}_D(t,x) : (D,t,x) \in H_n \times [0,T] \times E\big): \\
&\qquad \widehat{y}^{n+1}_D(t,x) = y^{n+1}_{D\cup\{(t,x)\}}(t) 1_{\{t>D^{\max}\}}.
\end{aligned} \tag{4.10}
$$

Then we take $Z \equiv (z^n)$ as follows:

$$
z^n_D(t,x) = \widehat{y}^{n+1}_D(t,x) - y^n_D(t), \tag{4.11}
$$

and it follows that

$$
\begin{aligned}
&S_n(\omega) < t \le S_{n+1}(\omega) \\
&\quad \Rightarrow \quad f\big(\omega,t,x,Y_{t-},Z_t(\cdot)\big) \\
&\qquad\qquad = f\{\widehat{y}^{n+1}\}^n_{D_n(\omega)}\big(t,x,y^n_{D_n(\omega)}(t), -y^n_{D_n(\omega)}(t)\big).
\end{aligned} \tag{4.12}
$$

The following lemma is a key point for our analysis.

LEMMA 7. *A càdlàg adapted process $Y \equiv (y^n)$ solves* (2.7) *if and only if for $\mathbb{P}$-almost all ω and all $n \ge 0$ we have*

$$
\begin{aligned}
&y^n_{D_n(\omega)}(t)\\
&\quad = u^n_{D_n(\omega)}\\
&\qquad + \int_t^T \int_E f\{\widehat{y}^{n+1}\}^n_{D_n(\omega)}\bigl(s, x, y^n_{D_n(\omega)}(s), -y^n_{D_n(\omega)}(s)\bigr)\nu^n_{D_n(\omega)}(ds, dx),\\
&\hspace{20em} t \in [0, T].
\end{aligned}
\tag{4.13}
$$

If further (2.17) *holds, then $Y \equiv (y^n)$ is a solution if and only if for $\mathbb{P}$-almost all ω we have* (4.13) *for all $n = 0, \dots, M-1$ and*

$$
t \in [0, T] \quad \Rightarrow \quad y^M_{D_M(\omega)}(t) = u^M_{D_M(\omega)} = \xi(\omega). \tag{4.14}
$$

PROOF. Considering the restriction of the BSDE to each interval $[S_n, S_{n+1}) \cap [0, T]$ and recalling (2.16), we see that Y is a solution if and only if, outside some null set $\mathcal{N}$, we have for $n \ge 0$

$$
\begin{aligned}
S_n \le t < S_{n+1} \le T \quad &\Rightarrow \quad Y_t = Y_{S_{n+1}-} + \int_t^{S_{n+1}} \int_E f\bigl(s, x, Y_s, Z_s(\cdot)\bigr)\nu(ds, dx),\\
S_n \le t \le T < S_{n+1} \quad &\Rightarrow \quad Y_t = \xi + \int_t^T \int_E f\bigl(s, x, Y_s, Z_s(\cdot)\bigr)\nu(ds, dx).
\end{aligned}
$$

Using the form $Y \equiv (y.)$, and $Z \equiv (z^n)$ as defined by (4.11), this is equivalent to having for $\omega \notin \mathcal{N}$

$$
\begin{aligned}
&S_n(\omega) \le t < S_{n+1}(\omega) \le T\\
&\qquad \Rightarrow \quad y^n_{D_n(\omega)}(t) = y^n_{D_n(\omega)}\bigl(S_{n+1}(\omega)\bigr)\\
&\qquad\qquad + \int_t^{S_{n+1}(\omega)} \int_E f\{\widehat{y}^{n+1}\}^n_{D_n(\omega)}\bigl(s, x, y^n_{D_n(\omega)}(s),\\
&\qquad\qquad - y^n_{D_n(\omega)}(s)\bigr)\nu^n_{D_n(\omega)}(ds, dx),
\end{aligned}
\tag{4.15}
$$

$$
\begin{aligned}
&S_n(\omega) \le t \le T < S_{n+1}(\omega)\\
&\qquad \Rightarrow \quad y^n_{D_n(\omega)}(t) = u^n_{D_n(\omega)}\\
&\qquad\qquad + \int_t^T \int_E f\{\widehat{y}^{n+1}\}^n_{D_n(\omega)}\bigl(s, x, y^n_{D_n(\omega)}(s),\\
&\qquad\qquad - y^n_{D_n(\omega)}(s)\bigr)\nu^n_{D_n(\omega)}(ds, dx).
\end{aligned}
\tag{4.16}
$$

Thus, if Y is a solution and $\omega \notin \mathcal{N}$, the function $y^n_{D_n(\omega)}$ satisfies the differential equation in (4.16) on the interval $[S_n(\omega) \wedge T, T]$, hence also on the interval $[0, T]$ because $\nu^n_{D_n(\omega)}([0, S_n(\omega)] \times E) = 0$ and $y^n_{D_n(\omega)}(t) = y^n_{D_n(\omega)}(S_n(\omega))$ if $t \le S_n(\omega)$ and also $u^n_{D_n(\omega)} = 0$ and $y^n_{D_n(\omega)}(t) = 0$ if $S_n(\omega) > T$: we thus have (4.13).

Conversely, assume that outside a null set $\mathcal{N}$ we have (4.13) for all n. Then obviously (4.16) holds, and (4.15) as well by taking the difference $y^n_{D_n(\omega)}(t) - y^n_{D_n(\omega)}(S_{n+1}(\omega))$. Therefore, Y solves the BSDE. This proves the first claim.

Assume further $\mathbb{P}(S_{M+1} = \infty) = 1$. Outside a null set, we have $S_n = \infty$ for all $n > M$, so (4.13) is trivially satisfied (with both members equal to 0) if $n > M$, and it reduces to (4.14) when $n = M$ because then $\nu^M_{D_M(\omega)}([0, T] \times E) = 0$, hence the second claim. $\square$

Equation (4.13) leads us to consider the following equation with unknown function y, for any given n,

$$y(t) = u^n_D + \int_t^T \int_E f\{\widehat{y}\}^n_D(s, x, y(s), -y(s))\nu^n_D(ds, dx), \qquad t \in [0, T], \tag{4.17}$$

where $D \in H_n$ is given, as well as the Borel function $\widehat{y}$ on $[0, T] \times E$ with further $\widehat{y}(t, x) = 0$ if $t \le D^{\max}$. When $D^{\max} = \infty$, and in view of our prevailing convention $u^D = 0$, plus $\nu^n_D([0, T] \times E) = 0$ in this case, this reduces to $y(t) = 0$. Otherwise, this equation is a backward ordinary integro-differential equation, and we have the following.

LEMMA 8. *Equation* (4.17) *has at most one solution*, *and it has one as soon as*

$$\int_0^T \int_E |\widehat{y}(s, x)| \nu^n_D(ds, dx) < \infty. \tag{4.18}$$

In this case, *the unique solution y satisfies*, *for all $\rho \ge L + L'$*,

$$\begin{aligned} |y(t)|e^{\rho a^n_D(t)} \le{}& |u^n_D|e^{\rho a^n_D(T)} \\ &+ \int_t^T \int_E \big(|f\{0\}^n_D(s, x, 0, 0)| + L|\widehat{y}(s, x)|\big)e^{\rho a^n_D(s)}\nu^n_D(ds, dx) \end{aligned} \tag{4.19}$$

and also, *if $\rho > L + L'$ and with a constant $\overline{C}$ depending only on (ρ, L, L')*,

$$\begin{aligned} &\int_t^T |y(s)|e^{\rho a^n_D(s)}\, da^n_D(s) \\ &\quad \le \overline{C}\bigg(|u^n_D|e^{\rho a^n_D(T)} + \int_t^T \int_E \big(|f\{0\}^n_D(s, x, 0, 0)| \\ &\qquad + L|\widehat{y}(s, x)|\big)e^{\rho a^n_D(s)}\nu^n_D(ds, dx)\bigg). \end{aligned} \tag{4.20}$$

PROOF. We have $f\{\widehat{y}\}^n_D(s, x, y(s), -y(s)) = g(s, x, y(s))$, where g is a Borel function on $[0, T] \times E \times \mathbb{R}$, which by (4.8) satisfies

$$|g(s, x, y') - g(s, x, y)| \le (L + L')|y' - y|,$$

$$\int_0^T \int_E |g(s,x,0)| \nu_D^n(ds,dx)$$
$$\leq \int_0^T |f\{0\}_D^n(t,x,0,0)| \nu_D^n(dt,dx) + L \int_0^T \int_E |\widehat{y}(s,x)| \nu_D^n(ds,dx).$$

The Lipschitz property of g implies the uniqueness, and the existence is classically implied by the finiteness of $\int_0^T \int_E |g(s,x,0)| \nu_D^n(ds,dx)$, which holds under (4.18) because of the last condition in (4.8).

Next, under (4.18), the proof of the estimates is the same as in Lemma 5. Namely, there is no jump here, so (3.5) is replaced by

$$|y(t)| e^{\rho a_D^n(t)} + \rho \int_t^T |y(s)| e^{\rho a_D^n(s)} \, da_D^n(s)$$
$$\leq |u_D^n| e^{\rho a_D^n(T)} + \int_t^T \int_E \big(|g(s,x,0)| + (L+L')|y(s)|\big) e^{\rho a_D^n(s)} \nu_D^n(ds,dx).$$

Note that here $\int_t^T |y(s)| e^{\rho a_D^n(s)} \, da_D^n(s) < \infty$ because $a_D^n(T) < \infty$. We readily get (4.19) if $\rho \geq L + L'$, and (4.20) if $\rho > L + L'$. □

We end this subsection with a technical lemma.

LEMMA 9. *For any $n \geq 0$ and any nonnegative Borel function g on $[0,T] \times E \times H_n \times H_{n+1}$, we have*

$$\begin{aligned} &\int_0^T \int_E g\big(s,x,D_n,D_n \cup \{(s,x)\}\big) \nu_{D_n}^n(ds,dx) \\ &\qquad = \mathbb{E}\big(g(S_{n+1},X_{n+1},D_n,D_{n+1}) e^{a_{D_n}^n(S_{n+1})} 1_{\{S_{n+1} \leq T\}} | \mathcal{F}_{S_n}\big). \end{aligned} \tag{4.21}$$

Moreover, the set $C' = \{D \in \mathcal{H}_n : \int_0^T \int_E 1_{\{D \cup \{(s,x)\} \in C\}} \nu_D^n(ds,dx) > 0\}$ is Λ_n-negligible, if $C \subset H_{n+1}$ is Λ_{n+1}-negligible.

PROOF. In view of (4.6), the left-hand side of (4.21) is

$$\int_0^T \int_E g\big(s,x,D_n,D_n \cup \{(s,x)\}\big) e^{a_{D_n}^n(s)} G_{D_n}^n(ds,dx),$$

so the first claim follows from the fact that $G_{D_n}^n$ is the $\mathcal{F}_{S_n}$-conditional law of $(S_{n+1} X_{n+1})$. For the last claim, it suffices to take the expectation of both sides of (4.19) with $g = 1_{[0,T] \times E \times H_n \times C}$: the right-hand side becomes $\mathbb{E}(e^{a_{D_n}^n(S_{n+1})} \times 1_C(D_{n+1}) 1_{\{S_{n+1} \leq T\}})$, which vanishes because $\Lambda_{n+1}(C) = 0$, whereas the left-hand side is positive if $\Lambda_n(C') > 0$. □

An example of an explicit solution: We will prove Theorem 2 later, but here we show how Lemma 7 allows us to give an explicit solution, in a special (but nontrivial) case of this theorem, with $M = 2$.

We consider a state space $E = \{x_1, x_2, x_3\}$ with three elements and suppose that $S_n = \infty$ for $n \geq 3$ and that $X_1 = x_1$ if $S_1 < \infty$, whereas conditionally on (S_1, S_2) and if $S_2 < \infty$ then X_2 takes the two values x_2 and x_3 with probability $\frac{1}{2}$. The law of the point process is thus completely characterized by the law $H^1(dt)$ of S_1, and by the conditional law $H^2(S_1, dt)$ of S_2 knowing S_1 (so $H^2(s, dt)$ is a transition probability from $[0, \infty]$ into itself, satisfying $H^2(\infty, \{\infty\}) = 1$ and $H^2(s, (s, \infty]) = 1$ if $s < \infty$). We also assume (A), which amounts to the facts that H^1 and H^2_s have no atom except ∞, plus $H^1((T, \infty]) > 0$ and $H^2(s, (T, \infty]) > 0$.

We consider the linear equation

$$
\begin{aligned}
&Y_t + \int_{(t,T]} \int_E Z(s,x)\mu(ds,dx) \\
&\qquad = 1_{\{S_2 \leq T, X_2 = x_2\}} + \int_{(t,T]} \int_E Z(s,x)\nu(ds,dx).
\end{aligned}
\tag{4.22}
$$

With the notation (4.1) and $\Delta = x_1$, say, we have $D_0 = (0, x_1)$ and $D_1 = ((0, x_1), (S_1, x_1))$ reduces to S_1. Thus, we may take

$$
\begin{aligned}
&u^0_D = 0, \qquad u^1_D = 0, \qquad u^2_{D_2} = 1_{\{S_2 \leq T, X_2 = x_2\}}, \\
&G'^0_D = H^1, \qquad \phi^0_{D,t} = \varepsilon_{x_1}, \qquad G'^1_{D_1} = H^2(S_1, \cdot), \qquad \phi^1_{D_1,t} = \tfrac{1}{2}(\varepsilon_{x_1} + \varepsilon_{x_2}), \\
&a^0_{D_0}(t) = a^0(t) = -\log H^1\big((t, \infty]\big), \\
&a^1_{D_1}(t) = a^1_{S_1}(t) = -\log H^2\big(S_1, (t, \infty]\big).
\end{aligned}
$$

Moreover, in (4.4) $y^0_D(t)$ is a function $y^0(t)$, and $y^1_{D_1}(t)$ takes the form $y^1_{S_1}(t) \times 1_{\{S_1 \leq T\}}$ for some function $(r, t) \mapsto y^1_r(t)$ on $[0, T]^2$, whereas by (4.14) we may take $y^2_{D_2}(t) = u^2_{D_2}$ for all t. The form of the generator implies that in (4.7) we have $f\{z^n\}^n_{D_n}(t, x, y, w) = w - z^n_{D_n}(t, x)$. Then, writing (4.13) for $n = 1$ and $n = 0$ gives us (below, r stands for S_1)

$$
\begin{aligned}
y^1_r(t) &= \frac{1}{2}\big(a^1_r(T) - a^1_r(t)\big) - \int_t^T y^1_r(s)\, da^1_r(s), \\
y^0(t) &= \int_t^T y^1_s(s)\, da^0(s) - \int_t^T y^0(s)\, da^0(s).
\end{aligned}
$$

This is a system of linear ODEs, whose explicit solution is [recall $a^1_s(s) = 0$]

$$
\begin{aligned}
y^1_r(t) &= \tfrac{1}{2}\big(1 - e^{a^1_r(t) - a^1_r(T)}\big), \\
y^0(t) &= \frac{1}{2} \int_t^T e^{a^0(t) - a^0(s)}\big(1 - e^{-a^1_s(T)}\big)\, da^0(s).
\end{aligned}
$$

Upon replacing $a^1_s(t)$ and $a^0(t)$ by $-\log \overline{H}^2_s(t)$ and $-\log \overline{H}^1(t)$, and using $y^2_{D_2}(t) = 1_{\{S_2 \leq T, X_2 = x_2\}}$, we obtain the following explicit form for the unique solution:

$$t \in [S_2, T] \quad \Rightarrow \quad Y_t = 1_{\{X_2 = x_2\}},$$

$$t \in [S_1 \wedge T, S_2 \wedge T] \quad \Rightarrow \quad Y_t = \frac{H^2(S_1, (t, T])}{2\overline{H}^2(S_1, (t, \infty])},$$

$$t \in [0, S_1 \wedge T] \quad \Rightarrow \quad Y_t = \frac{1}{2H^1((t, \infty])} \int_t^T H^2\bigl(s, (s, T]\bigr) H^1(ds).$$

REMARK 10. In this example, we have (2.17) with $M = 2$, so the uniqueness holds by Theorem 2. We also have (2.18), but the process A is not necessarily bounded: nevertheless we do have existence.

4.3. *Some counter-examples when* (A) *fails*. In all the paper, we assume (A), and it is enlightening to see what happens when this assumption fails. We are not going to do any deep study of this case, and will content ourselves with the simple situation where the point process is univariate and has a single point, that is, $E = \{\Delta\}$ is a singleton, and

$$N_t = 1_{\{S \le t\}},$$

where S is a variable with values in $(0, T] \cup \{\infty\}$. The filtration $(\mathcal{F}_t)$ is still the one generated by N, and G denotes the law of S, whereas $g(t) = G(t, \infty]$: those are the same as in (4.5), in our simplified setting.

The equation is (2.4), but since $A_t = A_{t \wedge S}$ and any predictable process is non-random, up to time S, it now reads as

$$Y_t + Z_S 1_{\{t < S \le T\}} = \xi + \int_{(t, S \wedge T]} f(s, Y_{s-}, Z_s)\, dA_s, \tag{4.23}$$

with f a Borel function on $[0, T] \times \mathbb{R} \times \mathbb{R}$, Lipschitz in its last two arguments, and such that $\int_0^T |f(s, 0, 0)|\, dA_s < \infty$.

Assumption (A) fails if (A_1) or (A_2) or both fail. Below, we examine what happens if either one of these two partial assumptions fails.

(1) *When G has an atom*. Here, we assume that (A_1) does not hold, that is, A is discontinuous, whereas $\mathbb{P}(S = \infty) > 0$, so (A_2) holds. We will see that in this case the existence of a solution to (4.23) is not guaranteed.

To see this, we consider the special case where S only takes the two values $r \in (0, T]$ and ∞, with respective positive probabilities p and $1 - p$. We have $N_t = 1_{\{r \le t\}} 1_{\{S = r\}}$ and $A_t = p 1_{\{t \ge r\}}$, so only the values of $f(t, y, z)$ at time $t = r$ are relevant, and we may assume that $f = f(y, z)$ only depends on y, z. Note also that ξ takes the form

$$\xi = a 1_{\{S = r\}} + b 1_{\{S = \infty\}} \qquad \text{where } a, b \in \mathbb{R}.$$

Moreover, only the value $Z_r(\omega)$ is involved, and it is nonrandom, and any solution Y is constant on $[0, r)$ and on $[r, T]$, that is, we have for $t \in [0, T]$

$$Z_r(\omega) = \gamma, \qquad Y_t = \delta 1_{\{t<r\}} + \rho 1_{\{t \geq r, S=r\}} + \eta 1_{\{t \geq r, S=\infty\}}$$
$$\text{where } \gamma, \delta, \rho, \eta \in \mathbb{R}.$$

Here, a, b are given, and $\gamma, \delta, \rho, \eta$ constitute the "solution" of (4.23), which reduces to the four equalities

$$\eta = b, \qquad \rho = a, \qquad \delta = b + pf(\delta, \gamma), \qquad \delta + \gamma = a + pf(\delta, \gamma),$$

which in turn give us

$$\gamma = a - b, \qquad \delta = b + pf(\delta, a - b).$$

The problem is that the last equation may not have a solution, and if it has one it is not necessarily unique. For example, we have:

$$\text{if } f(y, z) = \frac{1}{p}\big(y + g(z)\big),$$
$$\text{then } \begin{cases} \text{if } a + g(a - b) = 0 \text{ there are infinitely many solutions,} \\ \text{if } a + g(a - b) \neq 0 \text{ there is no solution.} \end{cases}$$

(2) *When G is supported by* $[0, T]$. Here, we suppose that G has no atom, but is supported by $[0, T]$. This corresponds to having (A_1), but not (A_2), and we have $A_t = a(t \wedge S)$, where $a(t) = -\log g(t)$ is increasing, finite for $t < v$ and infinite if $t \geq v$, where $v = \inf(t : g(t) = 0) \leq T$ is the right end point of the support of the measure G.

We will also consider a special generator, and more specifically the equation

$$Y_t + \int_{(t,T]} Z_s (dN_s - dA_s) = \xi, \tag{4.24}$$

which is (2.6) with $f \equiv 0$, and (2.4) with $f(t, y, z) = z$.

When ξ is integrable, the martingale representation theorem for point processes yields that $\xi = \mathbb{E}(\xi) + \int_0^T Z_s(dN_s - dA_s)$ for some predictable and dA_t-integrable process Z, hence $Y_t = \mathbb{E}(\xi | \mathcal{F}_t)$ is a solution. But this is *not* the only one. Indeed, recalling that here $\xi = h(S)$ is a (Borel) function of S, we have the following.

PROPOSITION 11. *Assume that $\mathbb{P}(S \leq T) = 1$ and that the law of S has no atom, and also that ξ is integrable. Then a process Y is a solution of* (4.24) *if and only if, outside a $\mathbb{P}$-null set, it takes the form*

$$Y_t = \xi 1_{\{t \geq S\}} + \left(w - \int_0^t e^{-A_s} h(s)\, dA_s\right) e^{A_t} 1_{\{t<S\}} \tag{4.25}$$

for an arbitrary real number w, and the associated process Z can be taken as $Z_t = h(t) - Y_{t-}$.

Note that $Y_0 = w$ in (4.25), so in particular it follows that (4.24) has a unique solution for any initial condition $Y_0 = w \in \mathbb{R}$. This is in deep contrast with Theorems 2 or 3, and it holds even for the trivial case $\xi \equiv 0$: in this trivial case, $Y_t = 0$ is of course a solution, but $Y_t = we^{A_t} 1_{\{t<S\}}$ for any $w \in \mathbb{R}$ is also a solution.

PROOF OF PROPOSITION 11. Any solution (Y, Z) satisfies $Y_t = \xi$ if $t \geq S$ and $Y_t = y(t)$ if $t < S$, where y is a continuous (nonrandom) function on $[0, v)$ (recall that $S < v$ a.s., and ess sup $S = v$). Since further (2.16) holds, one may always take the associated predictable process Z to be $Z_t = h(t) - Y_{t-}$. Then writing (4.24) for $t = 0$ and t arbitrary in $[0, v)$, we see that Y is a solution if and only if

$$t \in [0, v) \quad \Rightarrow \quad y(t) = y(0) + \int_0^t \bigl(y(s) - h(s)\bigr)\, da(s).$$

This is a linear ODE whose solutions are exactly the functions

$$y(t) = \left(w - \int_0^t e^{-a(s)} h(s)\, da(s)\right) e^{a(t)}$$

for $w \in \mathbb{R}$ arbitrary [since $\int_0^t |h(s)|\, da(s) \leq \frac{1}{g(t)} \mathbb{E}(|\xi|)$ is finite for all $t < v$]. This completes the proof. $\square$

REMARK 12. The previous result does not depend on the special form of the generator f, in the sense that for any f satisfying (2.5) and under the assumptions of Proposition 11, for any $w \in \mathbb{R}$ the BSDE admits a unique solution starting at $Y_0 = w$: of course an explicit form such as (4.25) is no longer available, but the proof of this result follows exactly the same argument as above.

REMARK 13. Jeanblanc and Réveillac [15] have studied some cases of BSDEs driven by a Wiener process, for which the generator "explodes" at the terminal time T. This bears some resemblance with the previous setting, in which $A_t = a(t \wedge S)$ and $a(t) \to \infty$ as $t \to T$. They show for example that, in the affine case, and under appropriate assumptions, there is no solution when $\mathbb{P}(\xi \neq 0) > 0$, and infinitely many solutions when $\xi \equiv 0$. Of course, the setting is quite different (a Wiener process instead of a point process), so the results are not really comparable, but they find cases like when (A_2) fails (no solutions) and like when (A_1) fails (infinitely many solutions).

5. Proof of the main results.

We start with an auxiliary lemma needed for proving the existence of a solution.

LEMMA 14. *Assume* (2.17) *and that* $A_T \leq K$ *for some constant* K. *Let* $m \in \{1, \dots, M\}$, *and suppose that we have* $y^n_{D_n}(t)$ *for* $n = m, m+1, \dots, M$, *such* *that* (4.13) *holds if* $m \le n < M$ *and* (4.14) *holds if* $n = M$, *outside a null set*. *Then for* n *between* m *and* $M-1$, *we have the* (*rather coarse*) *estimate*

$$\begin{aligned} v_n &:= \int_0^T \int_E \bigl(\bigl|f\{0\}^n_{D_n}(s,x,0,0)\bigr| + L\bigl|y^{n+1}_{D_n \cup \{(s,x)\}}(s)\bigr|\bigr)\nu^n_{D_n}(ds,dx) \\ &\le (1+L)^M e^{MK(2+L+L')} \\ &\quad \times \mathbb{E}\biggl(\int_{S_n \wedge T}^T \int_E \bigl|f(s,x,0,0)\bigr|\nu(ds,dx) + |\xi| 1_{\{S_n \le T\}} | \mathcal{F}_{S_n}\biggr). \end{aligned} \tag{5.1}$$

PROOF. (1) We first prove that $A_T \le K$ implies

$$n \ge 0, \qquad D \in H_n \quad \Rightarrow \quad a^n_D(T) \le K \tag{5.2}$$

for a suitable version of the a^n_D's, which amounts to proving $a^n_{D_n}(T) \le K$ a.s. To check this, we observe that for any $\gamma > 1$

$$\begin{aligned} e^{(\gamma-1)a^n_{D_n}(T)} &= \mathbb{E}\bigl(e^{\gamma a^n_{D_n}(T)} 1_{\{S_{n+1} > T\}} | \mathcal{F}_{S_n}\bigr) \\ &= \mathbb{E}\bigl(e^{\gamma a^n_{D_n}(T \wedge S_{n+1})} 1_{\{S_{n+1} > T\}} | \mathcal{F}_{S_n}\bigr) \le e^{K\gamma}, \end{aligned}$$

because $a^n_{D_n}(T \wedge S_{n+1}) \le A_T$ by (4.6). This implies $a^n_{D_n}(T) \le \frac{K\gamma}{\gamma-1}$ a.s. and, being true for all $\gamma > 1$, it yields (5.2).

(2) By Lemma 9, we have outside a null set

$$v_n = \mathbb{E}\bigl(e^{a^n_{D_n}(S_{n+1})}\bigl(\bigl|f\{0\}^n_{D_n}(S_{n+1}, X_{n+1}, 0, 0)\bigr| + L\bigl|y^{n+1}_{D_{n+1}}(S_{n+1})\bigr|\bigr) 1_{\{S_{n+1} \le T\}} | \mathcal{F}_{S_n}\bigr).$$

Equation (4.7) yields $f\{0\}^n_{D_n}(t,x,0,0) = f(t,x,0,0)$ if $S_n < t \le S_{n+1}$, whereas $u^n_{D_n} = 0$ if $S_n > T$, and $u^n_{D_n} = \xi$ if $S_n \le T < S_{n+1}$. In view of (4.14) and (5.2), we first deduce

$$v_{M-1} \le e^K \mathbb{E}\bigl(\bigl(\bigl|f(S_M, X_M, 0, 0)\bigr| + L|\xi|\bigr) 1_{\{S_M \le T\}} | \mathcal{F}_{S_{M-1}}\bigr).$$

It also gives us for $n \le M-2$, upon using (4.19) with $n+1$ and $\rho = L + L'$, and (5.2) again

$$\begin{aligned} v_n &\le e^K \mathbb{E}\bigl(\bigl(\bigl|f(S_{n+1}, X_{n+1}, 0, 0)\bigr| + Le^{(L+L')K}\bigl(\bigl|u^{n+1}_{D_{n+1}}\bigr| + v_{n+1}\bigr)\bigr) 1_{\{S_{n+1} \le T\}} | \mathcal{F}_{S_n}\bigr) \\ &\le e^K \mathbb{E}\bigl(\bigl(\bigl|f(S_{n+1}, X_{n+1}, 0, 0)\bigr| \\ &\qquad + Le^{(1+L+L')K}\bigl(|\xi| 1_{\{S_{n+2} > T\}} + v_{n+1}\bigr)\bigr) 1_{\{S_{n+1} \le T\}} | \mathcal{F}_{S_n}\bigr), \end{aligned}$$

where we have used $\mathbb{P}(S_{n+2} > T | \mathcal{F}_{S_{n+1}}) \ge e^{-K}$, which implies

$$\begin{aligned} \mathbb{E}\bigl(|\xi| 1_{\{S_{n+2} > T \ge S_{n+1}\}} | \mathcal{F}_{S_{n+1}}\bigr) &= \mathbb{E}\bigl(\bigl|u^{n+1}_{D_{n+1}}\bigr| 1_{\{S_{n+2} > T \ge S_{n+1}\}} | \mathcal{F}_{S_{n+1}}\bigr) \\ &= \bigl|u^{n+1}_{D_{n+1}}\bigr| 1_{\{T \ge S_{n+1}\}} \mathbb{P}(S_{n+2} > T | \mathcal{F}_{S_{n+1}}) \\ &\ge \bigl|u^{n+1}_{D_{n+1}}\bigr| 1_{\{T \ge S_{n+1}\}} e^{-K}. \end{aligned}$$

Iterating the estimates for v_n, and by successive conditioning, we deduce

$$
\begin{aligned}
v_n &\le (1+L)^M e^{MK(2+L+L')} \\
&\quad \times \mathbb{E}\Biggl(\sum_{i=n}^{M-1} \bigl| f(S_{i+1}, X_{i+1}, 0, 0)\bigr| 1_{\{S_{i+1}\le T\}} + L|\xi| 1_{\{S_i \le T < S_{i+1}\}} | \mathcal{F}_{S_n} \Biggr) \\
&\le (1+L)^M e^{MK(2+L+L')} \\
&\quad \times \mathbb{E}\biggl(\int_{S_n\wedge T}^{T} \int_E \bigl| f(s,x,0,0)\bigr| \mu(ds,dx) + L|\xi| 1_{\{S_n \le T\}} | \mathcal{F}_{S_n} \biggr).
\end{aligned}
$$

Since ν is the compensator of μ, this is equal to the right-hand side of (5.1), hence the result. $\square$

PROOF OF THEOREM 2. (a) We first prove the uniqueness. Let $Y \equiv (y^n)$ and $Y' \equiv (y'^n)$ be two solutions. By Lemma 7, for any $n = 0, \dots, M$ we have a subset B_n of H_n with $\Lambda_n(B_n^c) = 0$ and such that for any $D \in B_n$ both y_D^n and $y_D'^n$ satisfy (4.13) if $n < M$ and (4.14) if $n = M$.

The proof is done by downward induction. The induction hypothesis $K(n)$ is that for all $m = n, \dots, M$ we have a subset $B(n,m)$ of H_m with $\Lambda_m(B(n,m)^c) = 0$ such that $y_D^m \equiv y_D'^m$ for all $D \in B(n,m)$. That $K(M)$ holds with $B(M,M) = B_M$ is obvious, and $K(0)$ yields $Y_t = Y'_t$ a.s. for all t.

It remains to show that $K(n+1)$ for some n between 0 and $M-1$ implies $K(n)$. Assuming $K(n+1)$, we set $B(n,m) = B(n+1,m)$ for $m > n$ and let $B(n,n)$ be the intersection of B_n and of the set of all $D \in H_n$ such that $y^{n+1}_{D\cup\{(s,x)\}} = y'^{n+1}_{D\cup\{(s,x)\}}$ for ν_D^n-almost all (s,x). By virtue of the last claim in Lemma 9 applied with $C = B(n+1,n+1)^c$, plus $\Lambda_n(B_n^c) = 0$, we have $\Lambda_n(B(n,n)^c) = 0$. Then Lemma 8 yields $y_D^n = y_D'^n$ when $D \in B(n,n)$, hence $K(n)$ holds.

(b) We now turn to the existence, assuming further $A_T \le K$ and (2.18). We construct the family $(y^n_{D_n}(t))$ by downward induction on n, starting with $y_D^M(t) = u_D^M$ for all $D \in H_M$, hence (4.14) holds everywhere. Suppose now that we have a null set C_{n+1} and functions $y^m_{D_m}$ for $m = n+1, \dots, M-1$, each one satisfying (4.13) outside C_{n+1}. The assumption (2.18) and Lemma 14 imply $\mathbb{E}(v_n) < \infty$, so the set $C_n = C_{n+1} \cup \{v_n = \infty\}$ is negligible. Now, (4.13) is the same as (4.17) with $D = D_n$ and $\widehat{y}(s,x) = y^{n+1}_{D_n\cup\{(s,x)\}}(s) 1_{\{t > D^{max}\}}$, which is well defined for $G^n_{D_n}$-almost all (s,x), hence for $\nu^n_{D_n}$-almost all (s,x). Therefore, outside C_n these terms satisfy (4.18), and it follows that (4.17) has a unique solution $y^n_{D_n}$. This validates the induction, hence (2.7) has a solution, necessarily a.s. unique by part (a) above.

(c) It remains to prove the last claims. We denote by Y the (a.s. unique) solution, and recall that the associated predictable function Z can be chosen as $Z \equiv (z^n)$ with the form (4.9). Since $N_T \le M$, the last two claims amount to proving that $\mathbb{E}(U_n) < \infty$ for all $n \le M$, where $U_n = \int_{S_n\wedge T}^{S_{n+1}\wedge T} \int_E (|Y_s| + |Z(s,x)|) \nu(ds,dx)$.

Since $U_M = 0$ because $A_T = A_{T\wedge S_M}$, we restrict our attention to the case $n < M$. (4.3), (4.6) and (4.9) yield $U_n \le 2V_n + W_n$, where

$$V_n = \int_{S_n\wedge T}^{T} \big|y^n_{D_n}(s)\big|\,da^n_{D_n}(s), \qquad W_n = \int_{S_n\wedge T}^{T}\int_E \big|y^{n+1}_{D_n\cup\{(s,x)\}}(s)\big|\nu^n_{D_n}(ds,dx).$$

On the one hand, $LW_n \le v_n$, so (2.18) and (5.1) yield $\mathbb{E}(W_n) < \infty$. On the other hand, applying first (4.20) with any $\rho > L + L'$ and (5.2) and then $\mathbb{P}(S_{n+1} > T|\mathcal{F}_{S_n}) \ge e^{-K}$ and (5.1), we get

$$\begin{aligned}\mathbb{E}(V_n) &\le \overline{C}e^{K(L+L')}\mathbb{E}\big(|u^n_{D_n}|1_{\{S_n\le T\}} + v_n\big)\\ &\le \overline{C}e^{K(1+L+L')}\mathbb{E}\big(|\xi|1_{\{S_n\le T<S_{n+1}\}} + v_n\big) < \infty.\end{aligned}$$

This completes the proof. $\square$

PROOF OF THEOREM 3. (a) The uniqueness has been proved in Lemma 6. For the existence, we will "localize" the problem in the following way: for any $n \ge 1$ we set $T_n = S_n \wedge \inf(t : A_t \ge n)$ and we consider the equation

$$\begin{aligned}
&Y^{(n)}_t + \int_t^T\int_E Z^{(n)}(s,x)\mu^{(n)}(ds,dx)\\
&\qquad = \xi^{(n)} + \int_t^T\int_E f\big(s,x,Y^{(n)}_s,Z^{(n)}_s(\cdot)\big)\nu^{(n)}(ds,dx),\\
&\mu^{(n)}(ds,dx) = \mu(ds,dx)1_{\{s\le T_n\}}, \qquad \nu^{(n)}(ds,dx) = \nu(ds,dx)1_{\{s\le T_n\}},\\
&\xi^{(n)} = \xi 1_{\{T<T_n\}}.
\end{aligned}\tag{5.3}$$

Then $\nu^{(n)}$ is the compensator of $\mu^{(n)}$, relative to $(\mathcal{F}_t)$ and also to the smaller filtration $(\mathcal{F}^{(n)}_t = \mathcal{F}_{t\wedge T_n})$ generated by $\mu^{(n)}$, whereas $\xi^{(n)}$ is $\mathcal{F}^{(n)}_T$-measurable. The two marginal processes $N^{(n)}_t = \mu^{(n)}([0,t]\times E)$ and $A^{(n)}_t = \nu^{(n)}([0,t]\times E)$ satisfy $A^{(n)}_T \le n$ and $N^{(n)}_T \le n$, and (2.20) clearly implies (2.18) for $\xi^{(n)}$ and $\nu^{(n)}$. Therefore, Theorem 2 implies the existence of an a.s. unique solution $(Y^{(n)}, Z^{(n)})$ to (5.3), and the last claim of this theorem further implies that $\|(Y^{(n)}, Z^{(n)})\|^{(n)}_{\alpha,\beta} < \infty$, where the previous norm is the same as (2.19) with (A, N, ν) substituted with $(A^{(n)}, N^{(n)}, \nu^{(n)})$.

For $n' > n$, set

$$\overline{Y}^{(n,n')} = \sup_{s\in[0,T]}\big(e^{\beta A_s}\alpha^{N_s}|Y^{(n')}_s - Y^{(n)}_s|\big),$$

$$\mathcal{W}^{(n,n')}_{(s,t]} = \mathcal{W}^{\alpha,\beta}_{(s,t]}\big(Y^{(n')} - Y^{(n)}, Z^{(n')} - Z^{(n)}\big),$$

the latter being computed as in (3.2) with (A, N, ν).

We now proceed to bound these variables, and to this end we observe that

$$\begin{aligned}
&Y^{(n')}_{T_n\wedge t}+\int_t^T\int_E Z^{(n')}(s,x)\mu^{(n)}(ds,dx)\\
&\qquad = Y^{(n')}_{T_n\wedge T}+\int_t^T\int_E f\big(s,x,Y^{n'}_{T_n\wedge s},Z^{(n')}_s(\cdot)\big)\nu^{(n)}(ds,dx),
\end{aligned}$$

so $(Y^{(n')}_{T_n\wedge t}, Z^{(n')})$ is a solution of (5.3) with terminal value $Y^{(n')}_{T_n\wedge T}$ instead of $\xi^{(n)}$, and clearly has a finite $\|\cdot\|^{(n')}_{\alpha,\beta}$ norm. It then follows from (3.6) and (3.7), plus the maximal inequality for martingales, that for any $\varepsilon>0$ we have

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{t\in[0,T]} e^{\beta A_{T_n\wedge t}}\alpha^{N_{T_n\wedge t}}\big|Y^{(n')}_{T_n\wedge t}-Y^{(n)}_t\big|>\varepsilon\Big)\le\frac{\delta(n,n')}{\varepsilon},\\
&\mathbb{E}\big(\mathcal{W}^{(n,n')}_{(0,T_n\wedge T]}\big)\le C\delta\big(n,n'\big)\\
&\qquad\qquad\text{where } \delta\big(n,n'\big)=\mathbb{E}\big(\big|Y^{(n')}_{T_n\wedge T}-\xi^{(n)}\big|e^{\beta A_{T_n\wedge T}}\alpha^{N_{T_n\wedge T}}\big).
\end{aligned}$$

If $T_n>T$, we have $Y^{(n')}_{T_n\wedge T}=Y^{(n')}_T=\xi^{(n')}=\xi=\xi^{(n)}$, and otherwise $\xi^{(n)}=0$. Hence, (3.3) yields

$$\begin{aligned}
&\delta\big(n,n'\big)=\mathbb{E}\big(\big|Y^{(n')}_{T_n}\big|e^{\beta A_{T_n}}\alpha^{N_{T_n}}1_{\{T_n\le T\}}\big)\le\delta_n\\
&\text{where } \delta_n=\mathbb{E}\bigg(|\xi|e^{\beta A_T}\alpha^{N_T}1_{\{T_n\le T\}}+\int_{T_n\wedge T}^T\int_E\big|f(s,x,0,0)\big|e^{\beta A_s}\alpha^{N_s}\nu(ds,dx)\bigg).
\end{aligned}$$

If $T_n<t\le T$, we have $Y^{(n)}_t=\xi^{(n)}=0$ and we may take $Z^{(n)}(t,x)=0$, whereas if $T_{n'}\le t\le T$ we have $Y^{(n')}_t=\xi^{(n')}=0$ and we may take $Z^{(n')}(t,x)=0$, hence

$$\begin{aligned}
&\mathcal{W}^{n,n'}_{(0,T]}-\mathcal{W}^{n,n'}_{(0,T_n\wedge T]}\\
&\qquad=\int_{T_n\wedge T}^{T_{n'}\wedge T}\int_E\big(\big|Y^{(n')}_s\big|+\big|Z^{(n')}(s,x)\big|\big)e^{\beta A_{s\wedge T_{n'}}}\alpha^{N_{s\wedge T_{n'}}}\nu^{(n')}(ds,dx).
\end{aligned}$$

This and (3.4) yield $\mathbb{E}(\mathcal{W}^{n,n'}_{(0,T]}-\mathcal{W}^{n,n'}_{(0,T_n\wedge T]})\le C\delta_n$. Gathering all those partial results, we end up with

$$\mathbb{P}\big(\overline{Y}^{(n,n')}>\varepsilon\big)\le\mathbb{P}(T_n\le T)+\frac{\delta_n}{\varepsilon},\qquad \mathbb{E}\big(\mathcal{W}^{n,n'}_{(0,T]}\big)\le 2C\delta_n. \tag{5.4}$$

In view of (2.20) and the property $T_n\to\infty$ as $n\to\infty$, the dominated convergence theorem implies $\delta_n\to0$, hence both left sides in (5.4) go to 0 as $n\to\infty$, uniformly in $n'>n$. It follows that the sequence $Y^{(n)}$ is Cauchy for the convergence in probability, in the Skorokhod space $\mathbb{D}([0,T])$ endowed with the uniform metric, and that the pair $(Y^{(n)},Z^{(n)})$ is Cauchy in the space $\mathcal{L}^1_{\alpha,\beta}$. Therefore, these sequences converge in these spaces, to two limits Y and (Y',Z), with Y càdlàg adapted and

$(Y', Z) \in \mathcal{L}^1_{\alpha,\beta}$ and Z predictable and satisfying $\int_0^T \int_E |Z(s,x)|\nu(ds,dx) < \infty$ a.s.; we can of course find versions of the two limits for which $Y' = Y$ is the same process. Note that, since all Y^n are continuous outside the points S_n's, the same is true of Y.

We further deduce $\mathbb{E}(\int_0^T \int_E |Z^{(n)}(s,x) - Z(s,x)|\nu(ds,dx)) \to 0$, implying $\mathbb{E}(\int_0^T \int_E |Z^{(n)}(s,x) - Z(s,x)|\mu(ds,dx)) \to 0$, and thus $\int_t^T \int_E Z^{(n)}(s,x)\mu(ds, dx) \xrightarrow{\mathbb{P}} \int_t^T \int_E Z(s,x)\mu(ds,dx)$. Similarly, we obtain $\int_t^T \int_E f(s,x,Y_s^{(n)}, Z_s^{(n)}(\cdot))\nu(ds,dx) \xrightarrow{\mathbb{P}} \int_t^T \int_E f(s,x,Y_s,Z_s(\cdot))\nu(ds,dx)$ (we use the Lipschitz property of f here), and of course $Y_t^{(n)} \xrightarrow{\mathbb{P}} Y_t$ for each t. Since $(Y^{(n)}, Z^{(n)})$ solves (5.2), by passing to the limit we deduce that (Y, Z) solves (2.7), and it clearly belongs to $\mathcal{L}^1_{\alpha,\beta}$, thus ending the proof of the claim (a).

(b) We only need to prove that (2.21) for some $\varepsilon > 0$ implies (2.20) for all $\alpha > 0$ and $\beta \ge 0$, when $A_T \le K$ for some constant K. Since $\nu([0,T] \times E) = A_T$ and $\alpha^{N_t} \le (\alpha \vee 1)^{N_T}$ and $e^{\beta A_t} \le e^{\beta K}$, by Hölder's inequality it is clearly enough to show that α^{N_T} is in all $\mathbf{L}^p$ when $\alpha > 1$, or equivalently that $\mathbb{E}(\alpha^{N_T}) < \infty$ for all $\alpha > 1$.

We consider the nonnegative increasing process $U_t = \alpha^{N_t}$, which satisfies the equation

$$U_t = 1 + \alpha \int_0^t U_{s-}\,dN_s = 1 + \alpha \int_0^t U_{s-}\,dA_s + \alpha \int_0^t U_{s-}(dN_s - dA_s).$$

The last term is a local martingale, and a bounded martingale if we stop it at time $S_n \wedge T$, because $N_{S_n} \le n$ and $A_T \le K$ and $U_{t-} \le \alpha^{n-1}$ if $t \le S_n \wedge T$. Therefore, for any stopping time $S \le S'_n := S_n \wedge T$ we have

$$\mathbb{E}(U_{S-}) \le \mathbb{E}(U_S) = 1 + \alpha \mathbb{E}\biggl(\int_0^S U_{s-}\,dA_s\biggr).$$

Then one applies the Gronwall-type lemma (3.39) in [14] and $A_{S'_n} \le K$ to obtain that $\mathbb{E}(U_{S'_n-}) \le K'$ for a constant K' which only depends on K and α. Letting $n \to \infty$ and using the fact that $U_T \le \alpha U_{T-}$, the monotone convergence theorem yields $\mathbb{E}(U_T) \le \alpha K'$ as well, hence the result. $\square$

6. Application to a control problem. In this section, we show how what precedes can be put in use for solving a control problem. As before, we are given the multivariate point process μ of (2.1) on $(\Omega, \mathcal{F}, \mathbb{P})$, generating the filtration $(\mathcal{F}_t)$, and satisfying (A). The control problem is specified by the following data:

- a *terminal cost*, which is an $\mathcal{F}_T$-measurable random variable ξ;
- an *action* (or, decision) *space*, which is a measurable space $(U, \mathcal{U})$, and an associated predictable function r on $\Omega \times [0,T] \times E \times U$, which specifies how the control acts;
- a *running cost*, which is a predictable function l on $\Omega \times [0,T] \times U$.

These data should satisfy the following.

ASSUMPTION (B). There is a constant $C > 0$ such that, with A and N as in (2.2) and (2.3),

$$0 \le r(\omega, t, x, u) \le C, \tag{6.1}$$

$$\mathbb{E}\big(e^{A_T} C^{N_T}\big) < \infty. \tag{6.2}$$

We also have, for two constants $\alpha \in [1, \infty) \cap (C, \infty)$ and $\beta > 1 + C$,

$$\begin{aligned} \mathbb{E}\bigg(e^{\beta A_T} \alpha^{N_T} |\xi| + \int_0^T e^{\beta A_s} \alpha^{N_s} \Big|\inf_{u \in U} l(s, u)\Big| \, dA_s \\ + \int_0^T e^{A_s} C^{N_s} \sup_{u \in U} \big|l(s, u)\big| \, dA_s\bigg) < \infty. \end{aligned} \tag{6.3}$$

We denote by $\mathcal{A}$ the set of U-valued predictable processes. An element of $\mathcal{A}$ is called an *admissible control*, and it operates as follows. With $u = (u_t) \in \mathcal{A}$ we associate the probability measure $\mathbb{P}_u$ on $(\Omega, \mathcal{F})$ which is absolutely continuous with respect to $\mathbb{P}$ and admits the density process

$$L_t^u = \exp\bigg(\int_0^t \int_E (1 - r(s, x, u_s)) \nu(ds, dx)\bigg) \prod_{n \ge 1 : S_n \le t} r(S_n, X_n, u_{S_n}), \qquad t \in [0, T],$$

with the convention that an empty product equals 1. Such a $\mathbb{P}_u$ exists, because L^u is a nonnegative local martingale, satisfying $\sup_{t \le T} L_t^u \le e^{A_T} C^{N_T}$ by (6.1), and the latter variable is integrable by (6.2), so L^u is indeed a uniformly integrable martingale, with of course $\mathbb{E}(L_T^u) = 1$. By Girsanov's theorem for point processes, the predictable compensator of the measure μ under $\mathbb{P}_u$ is

$$\nu^u(dt, dx) = r(t, x, u_t) \nu(dt, dx) = r(t, x, u_t) \phi_t(dx) \, dA_t.$$

We finally define the cost associated to every $u(\cdot) \in \mathcal{A}$ as

$$J\big(u(\cdot)\big) = \mathbb{E}_u\bigg(\int_0^T l(t, u_t) \, dA_t + \xi\bigg),$$

where $\mathbb{E}_u$ denotes the expectation under $\mathbb{P}_u$.

Observe that, if $V_t = \int_0^t \sup_{u \in U} |l(s, u)| \, dA_s$, we have

$$\mathbb{E}_u\bigg(\int_0^T \big|l(t, u_t)\big| \, dA_t\bigg) \le \mathbb{E}_u\bigg(\int_0^T \sup_{u \in U} \big|l(t, u)\big| \, dA_t\bigg) = \mathbb{E}\big(L_T^u V_T\big).$$

Since L^u is a nonnegative martingale and V is continuous, adapted and increasing, we deduce

$$\mathbb{E}\big(L_T^u V_T\big) = \mathbb{E}\bigg(\int_0^T L_t^u \, dV_t\bigg) \le \mathbb{E}\bigg(\int_0^T e^{A_t} C^{N_t} \sup_{u \in U} \big|l(t, u)\big| \, dA_t\bigg) < \infty \tag{6.4}$$

by (6.3). Similarly, $\mathbb{E}_u(|\xi|) = \mathbb{E}(|\xi| L_T^u) \le \mathbb{E}(|\xi| e^{A_T} C^{N_T}) < \infty$, and we conclude that under (6.3) the cost $J(u(\cdot))$ is finite for every admissible control.

REMARK 15. Suppose that the cost functional has the form

$$J^1\big(u(\cdot)\big) = \mathbb{E}_u\Big(\sum_{n\geq 1: S_n \leq T} c(S_n, X_n, u_{S_n})\Big)$$

for some given predictable function c on $\Omega \times [0, T] \times E \times U$ which is, for instance, nonnegative. By a standard procedure, we can reduce this control problem to the previous one because

$$\begin{aligned} J^1\big(u(\cdot)\big) &= \mathbb{E}_u\Big(\int_0^T \int_E c(t, x, u_t)\mu(dt, dx)\Big) \\ &= \mathbb{E}_u\Big(\int_0^T \int_E c(t, x, u_t) r(t, x, u_t)\phi_t(dx)\, dA_t\Big). \end{aligned}$$

Thus, $J^1(u(\cdot))$ has the same form as $J(u(\cdot))$, with $\xi = 0$ and with the function l replaced by $l^1(t, u) = \int_E c(t, x, u) r(t, x, u)\phi_t(dx)$, so our forthcoming results can be applied.

Similar considerations obviously hold for cost functionals of the form $J(u(\cdot)) + J^1(u(\cdot))$.

The control problem consists in minimizing $J(u(\cdot))$ over $u(\cdot) \in \mathcal{A}$, and to this end a basic role is played by the BSDE

$$Y_t + \int_{(t,T]} \int_E Z(s, x)\mu(ds, dx) = \xi + \int_{(t,T]} f\big(s, Z_s(\cdot)\big)\, dA_s, \qquad t \in [0, T], \tag{6.5}$$

with terminal condition ξ being the terminal cost above, and with the generator f being the Hamiltonian function defined below. This is equation (2.7), with f only depending on (ω, t, ζ), and indeed it comes from an equation of type II via the transformation (2.14).

The Hamiltonian function f is defined on $\Omega \times [0, T] \times \mathcal{B}(E)$ as

$$f(\omega, t, \zeta) = \begin{cases} \inf_{u\in U}\Big(l(\omega, t, u) + \int_E \zeta(x) r(\omega, t, x, u)\phi_t(\omega, dx)\Big), \\ \qquad \text{if } \int_E |\zeta(x)|\phi_{\omega,t}(dx) < \infty, \\ 0, \qquad \text{otherwise.} \end{cases} \tag{6.6}$$

We will assume that the infimum is in fact achieved, possibly at many points. Moreover, we need to verify that the generator of the BSDE satisfies the conditions required in the previous section, in particular the measurability property, as expressed in (2.8), which does not follow from its definition. An appropriate assumption is the following one, since we will see below in Proposition 17 that it can be verified under quite general conditions.

ASSUMPTION (C). For every predictable function Z on $\Omega \times [0,T] \times E$ there exists a U-valued predictable process (i.e., an admissible control) $\underline{u}^Z$ such that, $dA_t(\omega)\mathbb{P}(d\omega)$-almost surely,

$$\begin{aligned} &f(\omega,t,Z_{\omega,t}(\cdot)) \\ &\quad = l(\omega,t,\underline{u}^Z(\omega,t)) + \int_E Z_{\omega,t}(x) r(\omega,t,x,\underline{u}^Z(\omega,t))\phi_t(\omega,dx). \end{aligned} \tag{6.7}$$

Now, it is easy to check that all the required assumptions for the solvability of the BSDE (6.5) are satisfied. Namely, using (6.1), one easily proves the inequality

$$|f(\omega,t,x,\zeta) - f(\omega,t,x,\zeta')| \le C \int_E |\zeta(y) - \zeta'(y)| \phi_{\omega,t}(dy),$$

whereas $f(\omega,t,0) = \inf_{u\in U} l(\omega,t,u)$. Then, in view of (6.3), we see that (2.8) and (2.20) are satisfied, with $L = C$ and $L' = 0$, hence $\beta > 1 + L + L'$ and $\alpha > L$. We thus conclude from Theorem 3 that the BSDE has a unique solution $(Y,Z) \in \mathcal{L}^1_{\alpha,\beta}$. The corresponding admissible control $\underline{u}^Z$, whose existence is required in Assumption (B), will be denoted as u^*.

THEOREM 16. *Assume (A), (B) and (C). Then, with (Y,Z) and u^* as above, the admissible control $u^*(\cdot)$ is optimal, and $Y_0 = J(u^*(\cdot)) = \inf_{u(\cdot)\in\mathcal{A}} J(u(\cdot))$ is the minimal cost.*

PROOF. Fix $u(\cdot) \in \mathcal{A}$. We first show that $\mathbb{E}_u \int_0^T \int_E |Z(t,x)| \nu^u(dt,dx) < \infty$. Indeed, setting $V_t = \int_0^t \int_E |Z(s,x)| r(s,x,u_s)\nu(ds,dx)$ and arguing as in (6.4),

$$\begin{aligned} &\mathbb{E}_u\left(\int_0^T \int_E |Z(t,x)|\nu^u(dt,dx)\right) \\ &\quad = \mathbb{E}_u\left(\int_0^T \int_E |Z(t,x)| r(t,x,u_t)\nu(dt,dx)\right) \\ &\quad = \mathbb{E}(L^u_T V_T) = \mathbb{E}\left(\int_0^T L^u_t\, dV_t\right) \le \mathbb{E}\left(\int_0^T e^{A_t} C^{N_t}\, dV_t\right) \\ &\quad = \mathbb{E}\left(\int_0^T \int_E e^{A_t} C^{N_t} |Z(t,x)| r(t,x,u_t)\nu(dt,dx)\right) \\ &\quad \le C\mathbb{E}\left(\int_0^T \int_E e^{\beta A_t} \alpha^{N_t} |Z_t(x)| \nu(dt,dx)\right), \end{aligned}$$

which is finite, since $(Y,Z) \in \mathcal{L}^1_{\alpha,\beta}$. By similar arguments, we also check that

$$\begin{aligned} &\mathbb{E}_u\left(\int_0^T |f(t,Z_t(\cdot))|\, dA_t\right) \\ &\quad = \mathbb{E}\left(\int_0^T L^u_t |f(t,Z_t(\cdot))|\, dA_t\right) \le \mathbb{E}\left(\int_0^T e^{A_t} C^{N_t} |f(t,Z_t(\cdot))|\, dA_t\right) \\ &\quad \le \mathbb{E}\left(\int_0^T e^{A_t} C^{N_t}\left(C\int_E |Z(t,x)|\phi_t(dx) + |f(t,0)|\right) dA_t\right) < \infty. \end{aligned}$$

Setting $t=0$ and taking the $\mathbb{P}_u$-expectation in the BSDE (6.5) we therefore obtain

$$Y_0+\mathbb{E}_u\biggl(\int_0^T\int_E Z(t,x)r(t,x,u_t)\nu(dt,dx)\biggr)=\mathbb{E}_u(\xi)+\mathbb{E}_u\biggl(\int_0^T f\bigl(t,Z_t(\cdot)\bigr)\,dA_t\biggr).$$

Adding $\mathbb{E}_u(\int_0^T l(t,u_t)\,dA_t)$ to both sides, we finally obtain the equality

$$\begin{aligned}
&Y_0+\mathbb{E}_u\biggl(\int_0^T\biggl(l(t,u_t)+\int_E Z(t,x)r(t,x,u_t)\phi_t(dx)\biggr)\,dA_t\biggr)\\
&\quad=J\bigl(u(\cdot)\bigr)+\mathbb{E}_u\biggl(\int_0^T f\bigl(t,Z_t(\cdot)\bigr)\,dA_t\biggr)\\
&\quad=J\bigl(u(\cdot)\bigr)+\mathbb{E}_u\biggl(\int_0^T\inf_{u\in U}\biggl(l(t,u)+\int_E Z(t,x)r(t,x,u_t),\phi_t(dx)\biggr)\,dA_t\biggr).
\end{aligned}$$

This implies immediately the inequality $Y_0\le J(u(\cdot))$ for every admissible control, with an equality if $u(\cdot)=u^*(\cdot)$. $\square$

Assumption (C) can be verified in specific situations when it is possible to compute explicitly the function $\underline{u}^Z$. General conditions for its validity can also be formulated using appropriate measurable selection theorems, as in the following proposition.

PROPOSITION 17. *Suppose that U is a compact metric space with its Borel σ-field $\mathcal{U}$ and that the functions $r(\omega,t,x,\cdot)$, $l(\omega,t,\cdot)$ are continuous on U for every (ω,t,x). Then if further (6.1) holds, Assumption (C) is satisfied.*

PROOF. For every predictable function Z set $G^Z=\{(\omega,t):\int_E|Z(\omega,t,x)|\phi_{\omega,t}(dx)=\infty\}$ and define a map $F^Z:\Omega\times[0,T]\times U\to\mathbb{R}$ by

$$F^Z(\omega,t,u)=\begin{cases} l(\omega,t,u)+\displaystyle\int_E Z(\omega,t,x)r(\omega,t,x,u)\phi_t(\omega,dx), & \text{if } (\omega,t)\notin G^Z,\\ 0, & \text{if } (\omega,t)\in G^Z.\end{cases}$$

Then $F^Z(\omega,t,\cdot)$ is continuous for every (ω,t) and F^Z is a predictable function on $\Omega\times[0,T]\times U$. By a classical selection theorem (see, e.g., Theorems 8.1.3 and 8.2.11 in [1] there exists a U-valued function $\underline{u}^Z$ on $\Omega\times[0,T]$ such that $F^Z(\omega,t,\underline{u}^Z(\omega,t))=\inf_{u\in U}F^Z(\omega,t,u)$ for every $(\omega,t)\in\Omega\times[0,T]$ [so that (6.7) holds true for every (ω,t)] and such that $\underline{u}^Z$ is measurable with respect to the completion of the predictable σ-algebra in $\Omega\times[0,T]$ with respect to the measure $dA_t(\omega)\mathbb{P}(d\omega)$. After modification on a null set, the function u^Z can be made predictable, and (6.7) still holds, as it is understood as an equality for $dA_t(\omega)\mathbb{P}(d\omega)$-almost all (ω,t). $\square$

REFERENCES

[1] AUBIN, J.-P. and FRANKOWSKA, H. (1990). *Set-Valued Analysis*. *Systems & Control*: *Foundations & Applications* **2**. Birkhäuser, Boston, MA. MR1048347

[2] BARLES, G., BUCKDAHN, R. and PARDOUX, E. (1997). Backward stochastic differential equations and integral-partial differential equations. *Stoch*. *Stoch*. *Rep*. **60** 57–83. MR1436432

[3] BECHERER, D. (2006). Bounded solutions to backward SDE's with jumps for utility optimization and indifference hedging. *Ann*. *Appl*. *Probab*. **16** 2027–2054. MR2288712

[4] BUCKDAHN, R. and PARDOUX, E. (1994). BSDE's with jumps and associated integral-stochastic differential equations. Preprint.

[5] CARBONE, R., FERRARIO, B. and SANTACROCE, M. (2007). Backward stochastic differential equations driven by càdlàg martingales. *Teor*. *Veroyatn*. *Primen*. **52** 375–385. MR2742510

[6] COHEN, S. N. and ELLIOTT, R. J. (2008). Solutions of backward stochastic differential equations on Markov chains. *Commun*. *Stoch*. *Anal*. **2** 251–262. MR2446692

[7] COHEN, S. N. and ELLIOTT, R. J. (2010). Comparisons for backward stochastic differential equations on Markov chains and related no-arbitrage conditions. *Ann*. *Appl*. *Probab*. **20** 267–311. MR2582649

[8] COHEN, S. N. and SZPRUCH, L. (2012). On Markovian solutions to Markov chain BSDEs. *Numer*. *Algebra Control Optim*. **2** 257–269. MR2929450

[9] CONFORTOLA, F. and FUHRMAN, M. (2013). Backward stochastic differential equations and optimal control of marked point processes. *SIAM J*. *Control Optim*. **51** 3592–3623. MR3105784

[10] CRÉPEY, S. and MATOUSSI, A. (2008). Reflected and doubly reflected BSDEs with jumps: A priori estimates and comparison. *Ann*. *Appl*. *Probab*. **18** 2041–2069. MR2462558

[11] EL KAROUI, N. and HUANG, S.-J. (1997). A general result of existence and uniqueness of backward stochastic differential equations. In *Backward Stochastic Differential Equations* (*Paris*, 1995–1996) (N. El Karoui and L. Mazliak, eds.). *Pitman Res*. *Notes Math*. *Ser*. **364** 27–36. Longman, Harlow. MR1752673

[12] EL KAROUI, N., PENG, S. and QUENEZ, M. C. (1997). Backward stochastic differential equations in finance. *Math*. *Finance* **7** 1–71. MR1434407

[13] JACOD, J. (1974/75). Multivariate point processes: Predictable projection, Radon–Nikodým derivatives, representation of martingales. *Z*. *Wahrsch*. *Verw*. *Gebiete* **31** 235–253. MR0380978

[14] JACOD, J. and MÉMIN, J. (1981). Weak and strong solutions of stochastic differential equations: Existence and stability. In *Stochastic Integrals* (*Proc*. *Sympos*., *Univ*. *Durham*, *Durham*, 1980) (R. Williams, ed.). *Lecture Notes in Math*. **851** 169–212. Springer, Berlin. MR0620991

[15] JEANBLANC, M. and RÉVEILLAC, A. (2014). A note on BSDEs with singular driver coefficients. In *Arbitrage*, *Credit and Informational Risks*. *Peking University Series in Mathematics* **5**. World Scientific, Hackensack, NJ.

[16] KHARROUBI, I. and LIM, T. (2012). Progressive enlargement of Filtrations and Backward SDEs with jumps. Preprint.

[17] PARDOUX, É. and PENG, S. G. (1990). Adapted solution of a backward stochastic differential equation. *Systems Control Lett*. **14** 55–61. MR1037747

[18] SHEN, L. and ELLIOTT, R. J. (2011). Backward stochastic differential equations for a single jump process. *Stoch*. *Anal*. *Appl*. **29** 654–673. MR2812521

[19] TANG, S. J. and LI, X. J. (1994). Necessary conditions for optimal control of stochastic systems with random jumps. *SIAM J*. *Control Optim*. **32** 1447–1475. MR1288257

[20] XIA, J. (2000). Backward stochastic differential equation with random measures. *Acta Math. Appl. Sin. Engl. Ser.* **16** 225–234. MR1779016

F. CONFORTOLA
M. FUHRMAN
POLITECNICO DI MILANO
DIPARTIMENTO DI MATEMATICA
VIA BONARDI 9
20133 MILANO
ITALY
E-MAIL: fulvia.confortola@polimi.it
marco.fuhrman@polimi.it

J. JEACOD
INSTITUT DE MATHÉMATIQUES DE JUSSIEU
CNRS-UMR 7586
UNIVERSITÉ PIERRE ET MARIE CURIE
4 PLACE JUSSIEU
75252 PARIS-CEDEX 5
FRANCE
E-MAIL: jean.jacod@upmc.fr