

## BACKWARD SDES FOR OPTIMAL CONTROL OF PARTIALLY OBSERVED PATH-DEPENDENT STOCHASTIC SYSTEMS: A CONTROL RANDOMIZATION APPROACH

BY ELENA BANDINI\*, ANDREA COSSO<sup>†</sup>, MARCO FUHRMAN<sup>‡,1</sup> AND HUYÊN PHAM<sup>§,¶</sup>

*LUISS Roma\**, *Politecnico di Milano<sup>†</sup>*, *Università di Milano<sup>‡</sup>*,  
*Université Paris Diderot<sup>§</sup>* and *CREST-ENSAE<sup>¶</sup>*

We introduce a suitable backward stochastic differential equation (BSDE) to represent the value of an optimal control problem with partial observation for a controlled stochastic equation driven by Brownian motion. Our model is general enough to include cases with latent factors in mathematical finance. By a standard reformulation based on the reference probability method, it also includes the classical model where the observation process is affected by a Brownian motion (even in presence of a correlated noise), a case where a BSDE representation of the value was not available so far. This approach based on BSDEs allows for greater generality beyond the Markovian case, in particular our model may include path-dependence in the coefficients (both with respect to the state and the control), and does not require any nondegeneracy condition on the controlled equation.

We use a randomization method, previously adopted only for cases of full observation, and consisting in a first step, in replacing the control by an exogenous process independent of the driving noise and in formulating an auxiliary (“randomized”) control problem where optimization is performed over changes of equivalent probability measures affecting the characteristics of the exogenous process. Our first main result is to prove the equivalence between the original partially observed control problem and the randomized problem. In a second step, we prove that the latter can be associated by duality to a BSDE, which then characterizes the value of the original problem as well.

**1. Introduction.** The main aim of this paper is to prove a representation formula for the value of a general class of stochastic optimal control problems with partial observation by means of an appropriate backward stochastic differential equation (backward SDE or BSDE).

To motivate our results, let us start with a classical optimal control problem with partial observation, where we consider an  $\mathbb{R}^n$ -valued controlled process  $X$  solution to

$$dX_t = b(X_t, \alpha_t) dt + \sigma^1(X_t, \alpha_t) dV_t^1 + \sigma^2(X_t, \alpha_t) dV_t^2,$$

---

Received September 2016; revised July 2017.

<sup>1</sup>Supported by the Italian MIUR-PRIN 2015 “Deterministic and stochastic evolution equations.” *MSC2010 subject classifications.* Primary 60H10; secondary 93E20.

*Key words and phrases.* Stochastic optimal control with partial observation, backward SDEs, randomization of controls, path-dependent controlled SDEs.

with initial condition  $X_0 = x_0$ , possibly random. The equation is driven by two processes  $V^1, V^2$  which are independent Wiener processes under some probability  $\mathbb{P}$ , and the coefficients depend on a control process  $\alpha$ . The aim is to maximize a reward functional of the form

$$J(\alpha) = \mathbb{E} \left[ \int_0^T f(X_s, \alpha_s) ds + g(X_T) \right],$$

where  $\mathbb{E}$  denotes the expectation under  $\bar{\mathbb{P}}$ . In the partial observation problem the control  $\alpha$  is constrained to being adapted to the filtration  $\mathbb{F}^W = (\mathcal{F}_t^W)_{t \geq 0}$  generated by another process  $W$ , called the observation process. A standard model, widely used in applications, consists in assuming that  $W$  is defined by the formula

$$dW_t = h(X_t, \alpha_t) dt + dV_t^2, \quad W_0 = 0.$$

In this problem,  $b, \sigma^1, \sigma^2, f, g, h$  are given data satisfying appropriate assumptions. We also introduce the value

$$U_0 = \sup_{\alpha} J(\alpha),$$

where  $\alpha$  ranges in the class of admissible control processes, that is,  $\mathbb{F}^W$ -progressive processes with values in some set  $A$  of control actions. A very effective approach to this problem is the so-called reference probability method, which consists in introducing, by means of a Girsanov transformation, a probability  $\mathbb{P}$  under which  $V^1$  and  $W$  are independent Wiener processes. Explicitly, one defines  $d\mathbb{P} = Z_T^{-1} d\bar{\mathbb{P}}$  where the density process  $Z$  satisfies

$$dZ_t = Z_t h(X_t, \alpha_t) dW_t, \quad Z_0 = 1.$$

Next, one introduces the process of unnormalized conditional distributions defined for every test function  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$  by the formula

$$\rho_t(\phi) = \mathbb{E}[\phi(X_t) Z_t \mid \mathcal{F}_t^W]$$

and proves that  $\rho$  is a solution to the so-called controlled Zakai equation:

$$(1.1) \quad d\rho_t(\phi) = \rho_t(\mathcal{L}^{\alpha_t} \phi) dt + \rho_t(h(\cdot, \alpha_t)\phi + \mathcal{M}^{\alpha_t} \phi) dW_t,$$

where  $\mathcal{L}^a \phi = \frac{1}{2} \text{Tr}(\sigma \sigma^T(\cdot, a) D^2 \phi) + D\phi b(\cdot, a)$ ,  $\mathcal{M}^a \phi = D\phi \sigma^2(\cdot, a)$ ,  $\sigma = (\sigma^1, \sigma^2)$  and initial condition  $\rho_0$  equal to the law of  $x_0$ . The reward functional to be maximized can be rewritten as

$$(1.2) \quad J(\alpha) = \mathbb{E} \left[ \int_0^T \rho_t(f(\cdot, \alpha_t)) dt + \rho_T(g(\cdot)) \right].$$

Under appropriate assumptions, for every admissible control process, the equation (1.1) has a unique solution  $\rho$  in some class of  $\mathbb{F}^W$ -progressive processes with values in the set of nonnegative Borel measures on  $\mathbb{R}^n$ , and thus (1.1)–(1.2) can be seen as an optimal control problem with full observation, called the separated

problem, having the same value  $\mathcal{U}_0$  as the original one (properly reformulated). Often, conditions are given so that  $\rho_t(dx)$  admits a density  $\eta_t(x)$  with respect to the Lebesgue measure on  $\mathbb{R}^n$  and the controlled Zakai equation is then written as an equation for  $\eta$ , considered as a process with values in some Hilbert space, for instance the space  $L^2(\mathbb{R}^n)$  or some weighted  $L^2$ -space. We also mention that, as an alternative to the Zakai equation, one could use the (controlled version of the) Kushner–Stratonovich equation and repeat similar considerations.

However, the new controlled state  $\rho$ , or equivalently  $\eta$ , is now an infinite-dimensional process, which makes the separated control problem rather challenging, and the subject of intensive study. It is not possible to give here a satisfactory description of the obtained results and we will limit ourselves to a brief sketch of the possible approaches and refer the reader to the treatises [27] and [6] for more complete results and references. A first method is the stochastic maximum principle, in the sense of Pontryagin, which provides necessary conditions for the optimality of a control process in terms of an adjoint equation and can be used to solve successfully the problem in a number of cases; see [6] and [32]. A second approach is the dynamic programming method to (1.1)–(1.2), which leads to a Hamilton–Jacobi–Bellman (HJB) equation in infinite dimension, and has been studied by viscosity methods in [26], later significantly extended in [18]. The issue of existence of an optimal control is addressed for instance in [13].

It is a remarkable fact that the use of BSDEs in this context is limited to the adjoint equations in the stochastic maximum principle cited above, in spite of the fact that BSDEs are used extensively and successfully in many areas of stochastic optimization to represent directly the value function. The reason for this can be explained by looking at the simple case when  $\sigma^1 = I$ ,  $\sigma^2 = 0$ ,  $h(x, a) = h(x)$ , that is, we have the partially observed controlled system

$$dX_t = b(X_t, \alpha_t) dt + dV_t^1, \quad dW_t = h(X_t) dt + dV_t^2,$$

and the corresponding controlled Zakai equation for the density process  $\eta$  is the following stochastic PDE in  $\mathbb{R}^n$ :

$$(1.3) \quad d\eta_t(x) = \frac{1}{2} \Delta \eta_t(x) dt - \operatorname{div}(b(x, \alpha_t) \eta_t(x)) dt + \eta_t(x) h(x) dW_t.$$

The standard method to represent the value based on BSDEs fails for such a control problem, since the diffusion coefficient is degenerate and the drift lacks the required structural condition; see for instance [14] for the infinite-dimensional case. Roughly speaking, an associated BSDE could immediately be written for equations having the form

$$d\eta_t(x) = \frac{1}{2} \Delta \eta_t(x) dt + \eta_t(x) h(x) [r(\eta_t(\cdot), t, x, \alpha_t) dt + dW_t]$$

for some coefficient  $r$ . This often implies, by an application of the Girsanov theorem, that the laws of the controlled processes  $\eta$  (depending on the control process

$\alpha$ ) are all absolutely continuous with respect to the law of the uncontrolled process corresponding to  $r = 0$ , but this property fails in general for the solutions to (1.3). The same difficulty can also be seen at the level of the corresponding HJB equation: in the simple case, we are addressing this equation is of semilinear type, but it does not fall into the class of PDEs whose solution can be represented by means of an associated BSDE; see for instance [28] or [14]. The problem of representing the value function of the classical partially observed control problem by means of a suitable BSDE becomes even more difficult in the general case when the HJB equation is fully nonlinear, and has remained unsolved so far.

It is the purpose of this paper to fill this gap in the existing literature, by introducing a suitable BSDE whose solution provides such a representation formula. As a motivation, we note that methods based on BSDEs have the advantage that they easily generalize beyond the Markovian framework. As a matter of fact, we will be able to treat control problems where the coefficients in the state equation and in the reward functional exhibit memory effects both with respect to state and to control, that is, their value at some time  $t$  may depend in a rather general way on the whole trajectory of the state and control processes on the time interval  $[0, t]$ . Various models with delay effects, or hereditary systems, are thus included in our treatment. For these models, there is no direct application of methods which exploit Markovianity, such as the HJB equation (although the Markovian character can be retrieved in a number of cases after an appropriate reformulation, which requires however nontrivial efforts and often introduces additional assumptions).

Another motivation for introducing BSDEs is the fact that their solutions can be approximated numerically. This way our result opens perspectives to finding an effective way to approximate the value of a partially observed problem, which is a difficult task due to the infinite-dimensional character of the Zakai equation and its corresponding HJB equation.

To perform our program, we first formulate a general control problem of the form

$$(1.4) \quad dX_t^\alpha = b_t(X^\alpha, \alpha) dt + \sigma_t(X^\alpha, \alpha) dB_t, \quad X_0^\alpha = x_0,$$

for  $t \in [0, T]$ , with reward functional and value defined by

$$(1.5) \quad J(\alpha) = \mathbb{E} \left[ \int_0^T f_t(X^\alpha, \alpha) dt + g(X^\alpha) \right], \quad \mathcal{V}_0 = \sup_\alpha J(\alpha),$$

where the coefficients  $b, \sigma, f, g$  depend on the whole trajectories of  $X$  and  $\alpha$  in a nonanticipative way. The partial observation character is modeled as follows: in the Wiener process  $B$ , we distinguish two components (possibly multidimensional) and write it in the form  $B = (V, W)$ . We call  $W$  the observation process and we require the control process  $\alpha$  to be adapted to the filtration generated by  $W$  and taking values in some set  $A$ , so that the supremum in (1.5) is taken over such controls. In Section 2.3.2, we prove that this model includes the classical partial observation problem described above as a special case and, moreover, that this

formulation is general enough to include large classes of optimization models with latent factors of interest in mathematical finance; see Section 2.3.1 below.

To tackle this problem, we will use a *randomization method*, introduced in [24] for classical Markovian models, but earlier considered in [23] in connection with impulse control and in [8, 12] on optimal switching problems. The idea of using the randomization method was inspired by the fact that it allows to represent (or construct) viscosity solutions to some classes of fully nonlinear PDEs. Other methods yield similar results, for instance those based on the notion of second-order BSDEs [31] or the theory of  $G$ -expectations [29]. It is likely that they might also be successfully applied to optimal control problems with partial observation. We also note that the randomization method has already been applied to a variety of situations; see [15] (compare also Remark 3.3 below), [2, 5, 9, 11, 16] in addition to the references given above. In order to present this method applied to the problem (1.4)–(1.5), we assume for simplicity that  $A$  is a subset of a Euclidean space and we take a finite measure  $\lambda$  on  $A$  with full support. Then, enlarging the original probability space if needed, we introduce a Poisson random measure  $\mu$  on  $\mathbb{R}_+ \times A$  with intensity  $\lambda(da)$  and independent of the Brownian motion  $B$ . Then we consider the stepwise process  $I$  associated with  $\mu$  and replace the control process  $\alpha$  by  $I$ , thus arriving at the following dynamics:

$$\begin{cases} dX_t = b_t(X, I) dt + \sigma_t(X, I) dB_t, \\ I_t = a_0 + \int_0^t \int_A (a - I_{s-}) \mu(ds da). \end{cases}$$

Next, we consider an auxiliary optimization problem, called the *randomized* or *dual* problem (in contrast to the starting optimal control problem with partial observation which we refer to also as *primal problem*), which consists in optimizing among equivalent changes of probability measures which only affect the intensity measure of  $\mu$  but not the law of  $W$ . In the randomized problem, an admissible control is a bounded positive map  $v$  defined on  $\Omega \times \mathbb{R}_+ \times A$ , which is predictable with respect to the filtration  $\mathbb{F}^{W, \mu}$  generated by  $W$  and  $\mu$ . Given  $v$ , by means of an absolutely continuous change of measure of Girsanov-type, we construct a probability measure  $\mathbb{P}^v$  such that the compensator of  $\mu$  is given by  $v_t(a)\lambda(da) dt$  and  $B$  remains a Brownian motion under  $\mathbb{P}^v$ . Then we introduce the reward and the value as

$$J^{\mathcal{R}}(v) = \mathbb{E}^v \left[ \int_0^T f_t(X, I) dt + g(X) \right], \quad \nu_0^{\mathcal{R}} = \sup_v J^{\mathcal{R}}(v),$$

where  $\mathbb{E}^v$  denotes the expectation under  $\mathbb{P}^v$ . One of our main results (see Theorem 3.1) states that the two control problems presented above are equivalent, in the sense that

$$(1.6) \quad \nu_0 = \nu_0^{\mathcal{R}}.$$

The reason for this construction is that, as shown in Section 5, the randomized control problem is associated to the following BSDE with a sign constraint, which then also characterizes the value function of the initial control problem (1.5). For any bounded measurable functional  $\varphi$  on the space of continuous paths with values in  $\mathbb{R}^n$ , define

$$\rho_t(\varphi) = \mathbb{E}[\varphi(X_{\cdot \wedge t}) \mid \mathcal{F}_t^{W, \mu}],$$

and consider the BSDE:

$$(1.7) \quad \begin{cases} Y_t = \rho_T(g) + \int_t^T \rho_s(f_s(\cdot, I)) ds + K_T - K_t \\ \quad - \int_t^T Z_s dW_s - \int_t^T \int_A U_s(a) \mu(ds da), \\ U_t(a) \leq 0. \end{cases}$$

In Theorem 5.1, which is another of our main results, we prove that there exists a unique minimal solution  $(Y, Z, U, K)$  to (1.7) (i.e., among all solutions we take the minimal one in terms of the  $Y$ -component) in a suitable space of stochastic processes adapted to the filtration  $\mathbb{F}^{W, \mu}$ , and moreover,

$$(1.8) \quad \begin{aligned} Y_0 = \mathcal{U}_0^{\mathcal{R}} = \mathcal{U}_0 \quad \text{and more generally} \\ Y_t = \operatorname{ess\,sup}_v \mathbb{E}^v \left[ \int_t^T \rho_s(f_s(\cdot, I)) ds + \rho_T(g) \mid \mathcal{F}_t^{W, \mu} \right]. \end{aligned}$$

The BSDE (1.7) is called the *randomized equation*, and corresponds to the HJB equation of the classical Markovian framework. Note that the introduction of the measure-valued process  $\rho$  and its occurrence in the generator and the terminal condition of the BSDE is reminiscent of the separated problem in classical optimal control with partial observation. We study in a companion paper [4] how one can also derive such kind of HJB equation in the context of partially observed Markovian control problems.

We note that probabilistic numerical methods have already been designed for BSDEs with constraints similar to (1.7) in [21] and [22]. We shall postpone for a future work the investigation, in the Markovian case, of an approximation scheme for (1.7), and hence for the value of the partially observed control problem.

We would like to point out that in our approach the original partially observed optimal control problem is formulated in the strong form, that is, with a fixed probability space. This is probably a more natural setting, especially in connection with modeling applications, and it is customary for the stochastic maximum principle and for other classes of optimization problems like optimal stopping and switching. However, almost all applications of BSDE techniques to the search of an optimal continuous control process are set in the weak formulation, since this avoids some difficulties (one exception may be found for instance in [17]). In spite of that, in the present paper we have chosen to adopt the strong formulation, at

the expense of additional technical difficulties. Moreover, we note that our main results are stated in a fairly general framework, allowing for locally Lipschitz coefficients with linear growth and without any nondegeneracy condition imposed on the diffusion coefficient  $\sigma$ . In particular, when  $\sigma = 0$ , this includes the case of the deterministic control problem with a path-dependent state dynamics and delay on control. Finally, when the diffusion coefficient of the Brownian motion  $V$  is zero, meaning that the dynamics of  $X$  is driven only by  $W$ , we are reduced to the case of full observation control problem. Therefore, we have provided a general equivalence and representation result in a unifying framework embedding several classical cases in optimal control theory and the proofs we present are almost entirely self-contained. We end this **Introduction** noting that the randomization method has already been applied to a variety of situations; see [2, 5, 9–11, 15, 16] in addition to the references given above.

The rest of the paper is organized as follows. In Section 2, we formulate the general optimal control problem (1.4)–(1.5) (the primal problem) with partial observation and path-dependence in the state and the control. We then present two motivating particular cases: a general optimization model with latent factors and uncontrolled observation process, which finds usual applications in mathematical finance, and the classical optimal control problem with partial observation discussed above (but including also path-dependence). Then, in Section 3, we implement the randomization method and formulate the randomized optimal control problem associated with the primal problem. We state in Theorem 3.1 the basic equivalence result between the primal and the randomized problem. Section 4 is entirely devoted to the proof of Theorem 3.1, which requires for both inequalities sharp approximation results and suitable constructions with marked point processes. In Section 5, we show a separation principle for the randomized control problem using nonlinear filtering arguments, and then relate by duality the separated randomized problem to a constrained BSDE, which may be viewed consequently as the randomized equation for the primal control problem.

## 2. General formulation and applications.

2.1. *Basic notation and assumptions.* In the following, we will consider controlled stochastic equations of the form

$$(2.1) \quad dX_t^\alpha = b_t(X^\alpha, \alpha) dt + \sigma_t(X^\alpha, \alpha) dB_t,$$

for  $t \in [0, T]$ , where  $T > 0$  is a fixed deterministic and finite terminal time, and gain functionals

$$J(\alpha) = \mathbb{E} \left[ \int_0^T f_t(X^\alpha, \alpha) dt + g(X^\alpha) \right].$$

The initial condition in (2.1) is  $X_0^\alpha = x_0$ , a given random variable with law denoted  $\rho_0$ . Before formulating precise assumptions let us explain informally the meaning

of several terms occurring in these expressions. The controlled process  $X^\alpha$  takes values in  $\mathbb{R}^n$  while  $B$  is a Wiener process in  $\mathbb{R}^{m+d}$ . We write  $B = (V, W)$  when we need to distinguish the first  $m$  components of  $B$  from the other  $d$  components. The control process, denoted by  $\alpha$ , takes values in a set  $A$  of control actions. The partial observation available to the controller will be described by imposing that the control process should be adapted to the filtration generated by the process  $W$  alone. Our formulation includes path-dependent (or hereditary) systems, that is, it allows for the presence of memory effects both on the state and the control. Indeed, the coefficients  $b, \sigma, f, g$  depend on the whole trajectory of  $X^\alpha$  and  $\alpha$ . The dependence will be nonanticipative, in the sense that their values at time  $t$  depend on the values  $X_s^\alpha$  and  $\alpha_s$  for  $s \in [0, t]$ : this is expressed below in a standard way by requiring that they should be progressive with respect to the canonical filtration on the space of paths.

Now let us come to precise assumptions and notation. Let us denote by  $\mathbf{C}_n$  the space of continuous paths from  $[0, T]$  to  $\mathbb{R}^n$ , equipped with the usual supremum norm  $\|x\|_\infty = x_T^*$ , where we set  $x_t^* := \sup_{s \in [0, t]} |x(s)|$ , for  $t \in [0, T]$  and  $x \in \mathbf{C}_n$ . We define the filtration  $(\mathcal{C}_t^n)_{t \in [0, T]}$ , where  $\mathcal{C}_t^n$  is the  $\sigma$ -algebra generated by the canonical coordinate maps  $\mathbf{C}_n \rightarrow \mathbb{R}^n, x(\cdot) \mapsto x(s), 0 \leq s \leq t$ :

$$\mathcal{C}_t^n := \sigma \{x(\cdot) \mapsto x(s) : s \in [0, t]\},$$

and we denote  $\text{Prog}(\mathbf{C}_n)$  the progressive  $\sigma$ -algebra on  $[0, T] \times \mathbf{C}_n$  with respect to  $(\mathcal{C}_t^n)$ .

We will require that the space of control actions  $A$  is a Borel space. We recall that a Borel space  $A$  is a topological space homeomorphic to a Borel subset of a Polish space. When needed,  $A$  will be endowed with its Borel  $\sigma$ -algebra  $\mathcal{B}(A)$ . We denote by  $\mathbf{M}_A$  the space of Borel measurable paths  $a : [0, T] \rightarrow A$ , we introduce the filtration  $(\mathcal{M}_t^A)_{t \in [0, T]}$ , where  $\mathcal{M}_t^A$  is the  $\sigma$ -algebra

$$\mathcal{M}_t^A := \sigma \{a(\cdot) \mapsto a(s) : s \in [0, t]\}$$

and we denote  $\text{Prog}(\mathbf{C}_n \times \mathbf{M}_A)$  the progressive  $\sigma$ -algebra on  $[0, T] \times \mathbf{C}_n$  with respect to the filtration  $(\mathcal{C}_t^n \otimes \mathcal{M}_t^A)_{t \in [0, T]}$ .

**(A1)**

- (i)  $A$  is a Borel space.
- (ii) The functions  $b, \sigma, f$  are defined on  $[0, T] \times \mathbf{C}_n \times \mathbf{M}_A$  with values in  $\mathbb{R}^n, \mathbb{R}^{n \times (m+d)}$  and  $\mathbb{R}$ , respectively, are assumed to be  $\text{Prog}(\mathbf{C}_n \times \mathbf{M}_A)$ -measurable (see also Remark 2.1 below).
- (iii) The function  $g$  is continuous on  $\mathbf{C}_n$ , with respect to the supremum norm. The functions  $b, \sigma$  and  $f$  are assumed to satisfy the following sequential continuity condition: once  $x_k, x \in \mathbf{C}_n, a_k, a \in \mathbf{M}_A, \|x_k - x\|_\infty \rightarrow 0, a_k(t) \rightarrow a(t)$  for  $dt$ -a.e.  $t \in [0, T]$  as  $k \rightarrow \infty$  we have

$$\begin{aligned} b_t(x_k, a_k) &\rightarrow b_t(x, a), & \sigma_t(x_k, a_k) &\rightarrow \sigma_t(x, a), \\ f_t(x_k, a_k) &\rightarrow f_t(x, a) & \text{for } dt\text{-a.e. } t \in [0, T]. \end{aligned}$$



(iv) There exist nonnegative constants  $L$  and  $r$  such that

$$(2.2) \quad |b_t(x, a) - b_t(x', a)| + |\sigma_t(x, a) - \sigma_t(x', a)| \leq L(x - x')^*_t,$$

$$(2.3) \quad |b_t(0, a)| + |\sigma_t(0, a)| \leq L,$$

$$(2.4) \quad |f_t(x, a)| + |g(x)| \leq L(1 + \|x\|_\infty^r),$$

for all  $(t, x, x', a) \in [0, T] \times \mathbf{C}_n \times \mathbf{C}_n \times \mathbf{M}_A$ .

(v)  $\rho_0$  is a probability measure on the Borel subsets of  $\mathbb{R}^n$  satisfying  $\int_{\mathbb{R}^n} |x|^p \rho_0(dx) < \infty$  for some  $p \geq \max(2, 2r)$ .

REMARK 2.1. The measurability condition **(A1)**(ii) is assumed because it guarantees the following property, which is easily deduced:

(ii)' Whenever  $(\Omega, \mathcal{F}, \mathbb{P})$  is a probability space with a filtration  $\mathbb{F}$ , and  $\alpha$  and  $X^\alpha$  are  $\mathbb{F}$ -progressive processes with values in  $A$  and  $\mathbb{R}^n$ , respectively, then the process  $(b_t(X^\alpha, \alpha), \sigma_t(X^\alpha, \alpha), f_t(X^\alpha, \alpha))_{t \in [0, T]}$  is also  $\mathbb{F}$ -progressive.

All the results in this paper still hold, with the same proofs, if property (ii)' is assumed to hold instead of (ii). There are cases when (ii)' is easy to be checked directly.

REMARK 2.2. We mention that the global Lipschitz condition in **(A1)**(iv) can be weakened to the case of locally Lipschitz coefficients with linear growth, which is needed to include the motivating examples presented in Sections 2.3.1 and 2.3.2 under the scope of our results. For the sake of conciseness, we postpone this extension to the longer version [3] of our paper (see Section 4.3 in [3]). We finally note that the function  $g$ , being continuous, is also  $\mathcal{C}_T^n$ -measurable.

REMARK 2.3. Assumption **(A1)** allows us to model various memory effects of the control on the state process, including important and usual cases of delay in the control. For instance, suppose that  $A$  is a bounded Borel subset of a Banach space and  $\bar{b} : A \rightarrow \mathbb{R}^n$  is Lipschitz continuous. Then we may consider a weighted combination of pure delays:

$$b_t(x, a) = \bar{b}\left(\sum_{i=1}^q \pi_i(t)a(t - \delta_i)\right),$$

where  $0 < \delta_1 < \dots < \delta_q < T$ ,  $\pi_i$  are bounded measurable real-valued functions and we use the convention that  $\alpha_t = \bar{\alpha}$  (a fixed element of  $A$ ) if  $t < 0$ . We may also allow the delays  $\delta_i$  to depend on  $t$  in an appropriate way. Alternatively, we may have

$$b_t(x, \alpha) = \bar{b}\left(\int_0^t \pi(t, s)a(s) ds\right),$$

with  $\pi$  bounded measurable and real-valued. Note that in the latter case the measurability condition **(A1)**(ii) fails in general, since the  $\sigma$ -algebras  $\mathcal{M}_t^A$  are determined by a countable number of times, but the property (ii)' in Remark 2.1 is easy to verify.

Clearly, we may address more complicated situations which are combinations of the two previous cases and may also include a dependence on the path  $x$ .

**REMARK 2.4.** We mention that no nondegeneracy assumption on the diffusion coefficient  $\sigma$  is imposed, and in particular, some lines or columns of  $\sigma$  may be equal to zero. We can then consider a priori more general model than (2.1) by adding dependence of the coefficients  $b, \sigma$  on another diffusion process  $M$ , for example, an unobserved and uncontrolled factor (see Application in Section 2.3.1). This generality is only apparent since it can be embedded in a standard way in our framework by considering the enlarged state process  $(X, M)$ .

**REMARK 2.5.** The requirement that  $p \geq \max(2, 2r)$  in **(A1)**(v) can be weakened for specific results in the sequel. For instance, Theorem 3.1 below still holds provided we only require  $p \geq \max(2, r)$ .

*2.2. Formulation of the partially observed control problem.* We assume that  $A, b, \sigma, f, g, \rho_0$  are given and satisfy the assumptions **(A1)**. We formulate a control problem fixing a setting  $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F}, V, W, x_0)$ , where  $(\Omega, \mathcal{F}, \mathbb{P})$  is a complete probability space with a right-continuous and  $\mathbb{P}$ -complete filtration  $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ ,  $V$  and  $W$  are processes with values in  $\mathbb{R}^m$  and  $\mathbb{R}^d$ , respectively, such that  $B = (V, W)$  is an  $\mathbb{R}^{m+d}$ -valued standard Wiener process with respect to  $\mathbb{F}$  and  $\mathbb{P}$ , and  $x_0$  is an  $\mathbb{R}^n$ -valued random variable, with law  $\rho_0$  under  $\mathbb{P}$ , which is assumed to be  $\mathcal{F}_0$ -measurable. Note that  $V$  and  $W$  are also standard Wiener processes and that  $V, W, x_0$  are all independent.

Let us denote  $\mathbb{F}^W = (\mathcal{F}_t^W)_{t \geq 0}$  the right-continuous and  $\mathbb{P}$ -complete filtration generated by  $W$ . An admissible control process is any  $\mathbb{F}^W$ -progressive process  $\alpha$  with values in  $A$ . The set of admissible control processes is denoted by  $\mathcal{A}^W$ . The controlled equation has the form

$$(2.5) \quad dX_t^\alpha = b_t(X^\alpha, \alpha) dt + \sigma_t(X^\alpha, \alpha) dB_t$$

on the interval  $[0, T]$  with initial condition  $X_0^\alpha = x_0$ , and the gain functional is

$$(2.6) \quad J(\alpha) = \mathbb{E} \left[ \int_0^T f_t(X^\alpha, \alpha) dt + g(X^\alpha) \right].$$

Since we assume that **(A1)** holds, by standard results (see, e.g., [30], Theorem V.11.2, or [20], Theorem 14.23), there exists a unique  $\mathbb{F}$ -adapted strong solution  $X^\alpha = (X_t^\alpha)_{0 \leq t \leq T}$  to (2.5) with continuous trajectories and such that (with the same  $p$  for which  $\mathbb{E}|x_0|^p < \infty$ )

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |X_t^\alpha|^p \right] < \infty.$$

The stochastic optimal control problem under partial observation consists in maximizing  $J(\alpha)$  over all  $\alpha \in \mathcal{A}^W$ :

$$(2.7) \quad \mathcal{U}_0 = \sup_{\alpha \in \mathcal{A}^W} J(\alpha).$$

REMARK 2.6. Let  $\mathbb{F}^B = (\mathcal{F}_t^B)_{t \geq 0}$  be the right-continuous and  $\mathbb{P}$ -complete filtration generated by  $B$ . Then  $B$  is clearly a  $\mathbb{F}^B$ -Brownian motion, the processes  $\alpha$  and  $X^\alpha$  are  $\mathbb{F}^B$ -progressive and the filtration  $\mathbb{F}$  does not play any role in determining  $J(\alpha)$  and  $\mathcal{U}_0$ . So we might assume from the beginning that  $\mathbb{F} = \mathbb{F}^B$  and even that  $\mathcal{F} = \mathcal{F}_\infty^B$  whenever convenient, but in the sequel we keep the previous framework unless explicitly mentioned.

2.3. *Two basic applications.* In this paragraph, we address two classical optimal control problems with partial observation, and we show that they can be recast in the form outlined in the previous subsection.

2.3.1. *Model with latent factors and uncontrolled observation process.* Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a complete probability space with a right-continuous and  $\mathbb{P}$ -complete filtration  $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ . Let  $V, \bar{W}$  be independent standard Wiener processes with respect to  $\mathbb{F}$ , with values in  $\mathbb{R}^m$  and  $\mathbb{R}^d$ , respectively. We assume that a controller, for instance a financial agent, wants to optimize her/his position, described by an  $\bar{n}$ -dimensional stochastic process  $\bar{X}^\alpha$  solution on the interval  $[0, T]$  to an equation of the form

$$(2.8) \quad \begin{aligned} d\bar{X}_t^\alpha &= \bar{b}_t(\bar{X}^\alpha, M, O, \alpha) dt + \bar{\sigma}_t^1(\bar{X}^\alpha, M, O, \alpha) dV_t \\ &\quad + \bar{\sigma}_t^2(\bar{X}^\alpha, M, O, \alpha) d\bar{W}_t \end{aligned}$$

with coefficients  $\bar{b}, \bar{\sigma}^1, \bar{\sigma}^2$  defined on  $[0, T] \times \mathbf{C}_{\bar{n}+\bar{m}+d} \times \mathbf{M}_A$  valued in  $\mathbb{R}^{\bar{n}}, \mathbb{R}^{\bar{n} \times m}, \mathbb{R}^{\bar{n} \times d}$ , respectively, and  $\text{Prog}(\mathbf{C}_{\bar{n}+\bar{m}+d} \times \mathbf{M}_A)$ -measurable. Here, the process  $M$ , valued in  $\mathbb{R}^{\bar{m}}$ , represents a latent factor that can influence the dynamics of  $\bar{X}^\alpha$  and is governed by a dynamics of the form

$$(2.9) \quad dM_t = \bar{\beta}_t(M) dt + \gamma_t^1(M) dV_t + \gamma_t^2(M) d\bar{W}_t,$$

for some coefficients  $\bar{\beta}, \gamma^1, \gamma^2$  defined on  $[0, T] \times \mathbf{C}_{\bar{m}}$  valued in  $\mathbb{R}^{\bar{m}}, \mathbb{R}^{\bar{m} \times m}, \mathbb{R}^{\bar{m} \times d}$ , respectively, and  $\text{Prog}(\mathbf{C}_{\bar{m}})$ -measurable. The process  $M$  is not directly observed, and actually the agent takes her/his decisions based on a noisy observation represented by a process  $O$  in  $\mathbb{R}^d$  solution to an equation of the form

$$(2.10) \quad dO_t = h_t(M, O) dt + k_t(O) d\bar{W}_t, \quad t \in [0, T],$$

for some coefficients  $h$  and  $k$  defined on  $[0, T] \times \mathbf{C}_{\bar{m}+d}$  and  $[0, T] \times \mathbf{C}_d$ ,  $\text{Prog}(\mathbf{C}_{\bar{m}+d})$ -measurable and  $\text{Prog}(\mathbf{C}_d)$ -measurable, valued in  $\mathbb{R}^d$  and  $\mathbb{R}^{d \times d}$ , re-

spectively. For instance,  $O_t$  may be related to the market price of financial risky assets at time  $t$ . We denote  $\mathbb{F}^O = (\mathcal{F}_t^O)_{t \geq 0}$  the right-continuous and  $\bar{\mathbb{P}}$ -complete filtration generated by  $O$ . An admissible control process, representing for instance the agent’s investment strategy, is any  $\mathbb{F}^O$ -progressive process  $\alpha$  with values in the Borel space  $A$ .

The agent wishes to maximize, among all admissible control processes, a gain functional of the form

$$J(\alpha) = \bar{\mathbb{E}} \left[ \int_0^T \bar{f}_t(\bar{X}^\alpha, M, O, \alpha) dt + \bar{g}(\bar{X}^\alpha, M, O) \right],$$

where  $\bar{\mathbb{E}}$  denotes expectation with respect to  $\bar{\mathbb{P}}$ , for real-valued coefficients  $\bar{f}, \bar{g}$  defined on  $[0, T] \times \mathbf{C}_{\bar{n}+\bar{m}+d} \times \mathbf{M}_A$  and  $\mathbf{C}_{\bar{n}+\bar{m}+d}, \text{Prog}(\mathbf{C}_{\bar{n}+\bar{m}+d} \times \mathbf{M}_A)$ -measurable and  $\mathcal{C}_T^{\bar{n}+\bar{m}+d}$ -measurable, respectively.

In order to put this problem in the form addressed in the previous subsection, we make a change of probability measure and pass from the “physical” probability  $\bar{\mathbb{P}}$  to a “reference” probability  $\mathbb{P}$ . Assuming that  $k_t(y)$  is invertible for all  $t \in [0, T]$  and  $y \in \mathbf{C}_d$ , and that the process  $\{k_t^{-1}(O)h_t(M, O), 0 \leq t \leq T\}$  is bounded, we define a process  $Z$  setting

$$Z_t^{-1} = \exp \left( - \int_0^t k_s(O)^{-1} h_s(M, O) d\bar{W}_s - \frac{1}{2} \int_0^t |k_s(O)^{-1} h_s(M, O)|^2 ds \right), \quad t \in [0, T].$$

The process  $Z^{-1}$  is a martingale under  $\bar{\mathbb{P}}$ , and by the Girsanov theorem, under the probability  $\mathbb{P}(d\omega) = Z_T(\omega)^{-1} \bar{\mathbb{P}}(d\omega)$  the pair  $(V, W)$  is a standard Wiener process in  $\mathbb{R}^{d+m}$  with respect to  $\mathbb{F}$ , where  $W_t = \bar{W}_t + \int_0^t k_s(O)^{-1} h_s(M, O) ds, t \in [0, T]$ . We denote by  $\mathbb{F}^W = (\mathcal{F}_t^W)_{t \in [0, T]}$  the right-continuous and  $\mathbb{P}$ -complete filtration generated by  $W$ , and see that the observation process  $O$  is a solution under  $\mathbb{P}$  to the equation:

$$(2.11) \quad dO_t = k_t(O) dW_t.$$

Assuming a Lipschitz condition on  $k$ , that is, there exists a constant  $K$  such that

$$|k_t(y) - k_t(y^1)| \leq K(y - y^1)_t^*,$$

for all  $(t, y, y^1) \in [0, T] \times \mathbf{C}_d \times \mathbf{C}_d$ , we deduce that  $O$  must be  $\mathbb{F}^W$ -adapted and, therefore, that  $\mathcal{F}_t^O \subset \mathcal{F}_t^W$  for  $t \in [0, T]$ . On the other hand, since  $W_t = \int_0^t k_s(O)^{-1} dO_s$ , the opposite inclusion also holds and we conclude that  $\mathbb{F}^O = \mathbb{F}^W$ . Moreover, it is easily checked that  $Z$  is a  $\mathbb{P}$ -martingale satisfying the equation

$$(2.12) \quad dZ_t = Z_t k_t(O)^{-1} h_t(M, O) dW_t,$$

and that the equations (2.8)–(2.9) for  $(\bar{X}^\alpha, M)$  can be rewritten under  $\mathbb{P}$  as

$$(2.13) \quad \begin{aligned} d\bar{X}_t^\alpha &= [\bar{b}_t(\bar{X}^\alpha, M, O, \alpha) - \bar{\sigma}_t^2(\bar{X}^\alpha, M, O, \alpha)k_t(O)^{-1}h_t(M, O)] dt \\ &\quad + \bar{\sigma}_t^1(\bar{X}^\alpha, M, O, \alpha)dV_t + \bar{\sigma}_t^2(\bar{X}^\alpha, M, O, \alpha)dW_t, \end{aligned}$$

$$(2.14) \quad \begin{aligned} dM_t &= [\bar{\beta}_t(M) - \gamma_t^2(M)k_t(O)^{-1}h_t(M, O)] dt \\ &\quad + \gamma_t^1(M)dV_t + \gamma_t^2(M)dW_t, \end{aligned}$$

while the gain functional is rewritten as an expectation under  $\mathbb{P}$  from the Bayes formula:

$$(2.15) \quad J(\alpha) = \mathbb{E} \left[ \int_0^T Z_t \bar{f}_t(\bar{X}^\alpha, M, O, \alpha) dt + Z_T \bar{g}(\bar{X}^\alpha, M, O) \right].$$

Now let us define the four-component process  $X^\alpha = (\bar{X}^\alpha, M, O, Z)$  and note that the equations (2.11), (2.12), (2.13), (2.14) specify a controlled stochastic equation for  $X^\alpha$  of the form (2.5) (with the obvious choice of  $b$  and  $\sigma$  in that equation). Similarly, the gain functional (2.15) can be put in the form (2.6) (with the obvious choice of  $f$  and  $g$ ).

EXAMPLE 2.1. As an example of financial application, let us mention the case of a risky asset whose price  $S_t$  satisfies

$$dS_t = S_t(\rho(M_t) dt + \sigma_t(S) d\bar{W}_t)$$

for a scalar Brownian motion  $\bar{W}$ , a volatility which is a functional of the past values of  $S$ , and an unobserved return process  $M$  governed by (2.9). We assume that  $\rho$ ,  $\sigma_t(\cdot)$  and  $\sigma_t^{-1}(\cdot)$  are bounded functions. The wealth  $\bar{X}_t^\alpha$  of an investor that invests a fraction  $\alpha_t$  of her/his wealth in this asset (and the rest in a risk-free asset with interest rate  $r$ ) evolves according to the self-financing equation:

$$(2.16) \quad \begin{aligned} d\bar{X}_t^\alpha &= \alpha_t \bar{X}_t^\alpha \frac{dS_t}{S_t} + (1 - \alpha_t) \bar{X}_t^\alpha r dt \\ &= \bar{X}_t^\alpha [r + \alpha_t(\rho(M_t) - r)] dt + \bar{X}_t^\alpha \alpha_t \sigma_t(S) d\bar{W}_t. \end{aligned}$$

The investor typically observes the risky price process or equivalently the log-price process  $O_t := \log S_t$  that solves the equation

$$dO_t = \left( \rho(M_t) - \frac{\sigma_t(S)^2}{2} \right) dt + \sigma_t(S) d\bar{W}_t,$$

which can be put in the form (2.10) setting  $k_t(y) = \sigma_t(\exp(y))$  and  $h_t(z, y) = \rho(z) - k_t(y)^2/2$ . Notice that the wealth process is  $\mathbb{F}^O$ -adapted, since it is solution to equation (2.16). Therefore, when choosing the investment strategy  $\alpha$  the investor gains no additional information by observing the wealth process, and so it is reasonable to impose the condition that  $\alpha$  should be adapted to the filtration  $\mathbb{F}^O$  alone, rather than to the one generated by  $O$  and  $\bar{X}^\alpha$ .

2.3.2. *A classical partially observed control problem.* In the previous example, the observed process  $O$  was not affected by the choice of the control. We next remove this restriction, adopting a classical approach which consists in starting with the “reference” probability  $\mathbb{P}$  and introducing the “physical” probability later, as presented, for example, in the book [6].

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a complete probability space with a right-continuous and  $\mathbb{P}$ -complete filtration  $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ . Let  $V, W$  be independent standard Wiener processes with respect to  $\mathbb{F}$ , with values in  $\mathbb{R}^m$  and  $\mathbb{R}^d$ , respectively, and consider the observation process solution to the equation in  $\mathbb{R}^d$

$$(2.17) \quad dO_t = k_t(O) dW_t,$$

where  $k_t(y)$  is defined on  $[0, T] \times \mathbf{C}_d$ ,  $\text{Prog}(\mathbf{C}_d)$ -measurable, Lipschitz in  $y$  and invertible with bounded inverse. Similarly, as in the previous paragraph, we see that  $\mathbb{F}^W = \mathbb{F}^O$ , and an admissible control process is any  $\mathbb{F}^W$ -progressive process  $\alpha$  with values in a Borel space  $A$ .

We are given coefficients  $\bar{b}, h, \bar{\sigma}^1, \bar{\sigma}^2$  defined on  $[0, T] \times \mathbf{C}_{\bar{n}+d} \times \mathbf{M}_A$ , valued in  $\mathbb{R}^{\bar{n}}, \mathbb{R}^d, \mathbb{R}^{\bar{n} \times m}, \mathbb{R}^{\bar{n} \times d}$ , respectively, and  $\text{Prog}(\mathbf{C}_{\bar{n}+d} \times \mathbf{M}_A)$ -measurable. Then, for any admissible control process  $\alpha$ , let the process  $\bar{X}^\alpha$  be defined as the solution to the equation in  $\mathbb{R}^{\bar{n}}$ :

$$(2.18) \quad \begin{aligned} d\bar{X}_t^\alpha &= [\bar{b}_t(\bar{X}^\alpha, O, \alpha) - \bar{\sigma}_t^2(\bar{X}^\alpha, O, \alpha)k_t(O)^{-1}h_t(\bar{X}^\alpha, O, \alpha)] dt \\ &+ \bar{\sigma}_t^1(\bar{X}^\alpha, O, \alpha) dV_t + \bar{\sigma}_t^2(\bar{X}^\alpha, O, \alpha) dW_t. \end{aligned}$$

We introduce the gain functional  $J(\alpha)$  associated to a control  $\alpha$  by means of a change of probability in the following way. Assuming that the function  $k^{-1}h$  is bounded, let us define for any admissible control process  $\alpha$ , the  $\mathbb{P}$ -martingale:

$$Z_t^\alpha = \exp\left(\int_0^t k_s(O)^{-1}h_s(\bar{X}^\alpha, O, \alpha) dW_s - \frac{1}{2} \int_0^t |k_s(O)^{-1}h_s(\bar{X}^\alpha, O, \alpha)|^2 ds\right),$$

solution to the equation

$$(2.19) \quad dZ_t^\alpha = Z_t^\alpha k_t(O)^{-1}h_t(\bar{X}^\alpha, O, \alpha) dW_t,$$

and introduce the “physical” probability  $\mathbb{P}^\alpha$  setting  $\mathbb{P}^\alpha(d\omega) = Z_T^\alpha(\omega)\mathbb{P}(d\omega)$ . Given real-valued coefficients  $\bar{f}, \bar{g}$  defined on  $[0, T] \times \mathbf{C}_{\bar{n}+d} \times \mathbf{M}_A$  and  $\mathbf{C}_{\bar{n}+d}$ ,  $\text{Prog}(\mathbf{C}_{\bar{n}+d} \times \mathbf{M}_A)$ -measurable and  $\mathcal{C}_T^{\bar{n}+d}$ -measurable, respectively, the gain functional is then defined as

$$J(\alpha) = \mathbb{E}^\alpha \left[ \int_0^T \bar{f}_t(\bar{X}^\alpha, O, \alpha) dt + \bar{g}(\bar{X}^\alpha, O) \right].$$

The interpretation of this formulation is the following. By defining the process  $W^\alpha$  as

$$W_t^\alpha = W_t - \int_0^t k_s(O)^{-1}h_s(\bar{X}^\alpha, O, \alpha) ds, \quad t \in [0, T],$$

for any admissible control process  $\alpha$ , we see by the Girsanov theorem that the pair  $(V, W^\alpha)$  is a standard Wiener process in  $\mathbb{R}^{m+d}$  under the probability  $\mathbb{P}^\alpha$  and with respect to  $\mathbb{F}$ . Moreover, the dynamics of  $(\bar{X}^\alpha, O)$  is written under  $\mathbb{P}^\alpha$  as

$$\begin{aligned} d\bar{X}_t^\alpha &= \bar{b}_t(\bar{X}^\alpha, O, \alpha) dt + \bar{\sigma}_t^1(\bar{X}^\alpha, O, \alpha) dV_t + \bar{\sigma}_t^2(\bar{X}^\alpha, O, \alpha) dW_t^\alpha, \\ dO_t &= h_t(\bar{X}^\alpha, O, \alpha) dt + k_t(O) dW_t^\alpha. \end{aligned}$$

We then obtain a classical controlled state equation, and an observation process perturbed by noise and also affected by the choice of the control.

Finally, we notice that this problem is recast in the framework of Section 2.2 by rewriting from Bayes' formula and the  $\mathbb{P}$ -martingale property of  $Z^\alpha$ , the gain functional as an expectation under  $\mathbb{P}$ :

$$(2.20) \quad J(\alpha) = \mathbb{E} \left[ \int_0^T Z_t^\alpha \bar{f}_t(\bar{X}^\alpha, O, \alpha) dt + Z_T^\alpha \bar{g}(\bar{X}^\alpha, O) \right].$$

Thus, by defining the three-component process  $X^\alpha = (\bar{X}^\alpha, Z^\alpha, O)$ , we see that the equations (2.17), (2.18), (2.19) specify a controlled stochastic equation for  $X^\alpha$  of the form (2.5), and the gain functional (2.20) can be put in the form (2.6).

**3. The randomized stochastic optimal control problem.** We still assume that  $A, b, \sigma, f, g, \rho_0$  are given and satisfy the assumptions **(A1)**. We implement the randomization method and formulate the randomized stochastic optimal control problem associated with the control problem of Section 2.2. To this end, we suppose we are also given  $\lambda, a_0$  satisfying the following conditions, which are assumed to hold from now on:

**(A2)**

- (i)  $\lambda$  is a finite positive measure on  $(A, \mathcal{B}(A))$  with full topological support.
- (ii)  $a_0$  is a fixed, deterministic point in  $A$ .

We anticipate that  $\lambda$  will play the role of an intensity measure and  $a_0$  will be the starting point of some auxiliary process introduced later. Notice that the initial problem (2.7) does not depend on  $\lambda, a_0$ , which only appear in order to give a randomized representation of the partially observed control problem. In this sense, **(A2)** is not a restriction imposed on the original problem and we have the choice to fix  $a_0 \in A$  and an intensity measure  $\lambda$  satisfying this condition.

3.1. *Formulation of the randomized control problem.* The randomized control problem is formulated fixing a setting  $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}, \hat{V}, \hat{W}, \hat{\mu}, \hat{x}_0)$ , where  $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$  is an arbitrary complete probability space with independent random elements  $\hat{V}, \hat{W}, \hat{\mu}, \hat{x}_0$ . The random variable  $\hat{x}_0$  is  $\mathbb{R}^n$ -valued, with law  $\rho_0$  under  $\hat{\mathbb{P}}$ . The process  $\hat{B} := (\hat{V}, \hat{W})$  is a standard Wiener process in  $\mathbb{R}^{m+d}$  under  $\hat{\mathbb{P}}$ .  $\hat{\mu}$  is a Poisson random measure on  $A$  with intensity  $\lambda(da)$  under  $\hat{\mathbb{P}}$ ; thus,  $\hat{\mu}$  is a sum of Dirac measures

of the form  $\hat{\mu} = \sum_{n \geq 1} \delta_{(\hat{S}_n, \hat{\eta}_n)}$ , where  $(\hat{\eta}_n)_{n \geq 1}$  is a sequence of  $A$ -valued random variables and  $(\hat{S}_n)_{n \geq 1}$  is a strictly increasing sequence of random variables with values in  $(0, \infty)$ , and for any  $C \in \mathcal{B}(A)$  the process  $\hat{\mu}((0, t] \times C) - t\lambda(C)$ ,  $t \geq 0$ , is a  $\hat{\mathbb{P}}$ -martingale. We also define the  $A$ -valued process

$$(3.1) \quad \hat{I}_t = \sum_{n \geq 0} \hat{\eta}_n 1_{[\hat{S}_n, \hat{S}_{n+1})}(t), \quad t \geq 0,$$

where we use the convention that  $\hat{S}_0 = 0$  and  $\hat{I}_0 = a_0$ , the point in assumption **(A2)(ii)**. Notice that the formal sum in (3.1) makes sense even if there is no addition operation defined in  $A$  and that, when  $A$  is a subset of a linear space, formula (3.1) can be written as

$$\hat{I}_t = a_0 + \int_0^t \int_A (a - \hat{I}_{s-}) \hat{\mu}(ds da), \quad t \geq 0.$$

Let  $\hat{X}$  be the solution to the equation

$$(3.2) \quad d\hat{X}_t = b_t(\hat{X}, \hat{I}) dt + \sigma_t(\hat{X}, \hat{I}) dB_t,$$

for  $t \in [0, T]$ , starting from  $\hat{X}_0 = \hat{x}_0$ . We define two filtrations  $\mathbb{F}^{\hat{W}, \hat{\mu}} = (\mathcal{F}_t^{\hat{W}, \hat{\mu}})_{t \geq 0}$  and  $\mathbb{F}^{\hat{x}_0, \hat{B}, \hat{\mu}} = (\mathcal{F}_t^{\hat{x}_0, \hat{B}, \hat{\mu}})_{t \geq 0}$  setting

$$(3.3) \quad \begin{aligned} \mathcal{F}_t^{\hat{W}, \hat{\mu}} &= \sigma(\hat{W}_s, \hat{\mu}((0, s] \times C) : s \in [0, t], C \in \mathcal{B}(A)) \vee \mathcal{N}, \\ \mathcal{F}_t^{\hat{x}_0, \hat{B}, \hat{\mu}} &= \sigma(\hat{x}_0, \hat{B}_s, \hat{\mu}((0, s] \times C) : s \in [0, t], C \in \mathcal{B}(A)) \vee \mathcal{N}, \end{aligned}$$

where  $\mathcal{N}$  denotes the family of  $\hat{\mathbb{P}}$ -null sets of  $\hat{\mathcal{F}}$ . We denote  $\mathcal{P}(\mathbb{F}^{\hat{W}, \hat{\mu}})$ ,  $\mathcal{P}(\mathbb{F}^{\hat{x}_0, \hat{B}, \hat{\mu}})$  the corresponding predictable  $\sigma$ -algebras.

Under **(A1)** it is well known (see, e.g., Theorem 14.23 in [20]) that there exists a unique  $\mathbb{F}^{\hat{x}_0, \hat{B}, \hat{\mu}}$ -adapted strong solution  $\hat{X} = (\hat{X}_t)_{0 \leq t \leq T}$  to (3.2), satisfying  $\hat{X}_0 = \hat{x}_0$ , with continuous trajectories and such that (with the same  $p$  for which  $\hat{\mathbb{E}}|\hat{x}_0|^p < \infty$ )

$$(3.4) \quad \hat{\mathbb{E}} \left[ \sup_{t \in [0, T]} |\hat{X}_t|^p \right] < \infty.$$

We can now define the randomized optimal control problem as follows: the set  $\hat{\mathcal{V}}$  of admissible controls consists of all  $\hat{v} = \hat{v}_t(\hat{\omega}, a) : \hat{\Omega} \times \mathbb{R}_+ \times A \rightarrow (0, \infty)$ , which are  $\mathcal{P}(\mathbb{F}^{\hat{W}, \hat{\mu}}) \otimes \mathcal{B}(A)$ -measurable and bounded. Then the Doléans exponential process

$$(3.5) \quad \begin{aligned} \kappa_t^{\hat{v}} &= \mathcal{E}_t \left( \int_0^t \int_A (\hat{v}_s(a) - 1) (\hat{\mu}(ds da) - \lambda(da) ds) \right) \\ &= \exp \left( \int_0^t \int_A (1 - \hat{v}_s(a)) \lambda(da) ds \right) \prod_{0 < \hat{S}_n \leq t} v_{\hat{S}_n}(\hat{\eta}_n), \quad t \geq 0, \end{aligned}$$



is a martingale with respect to  $\hat{\mathbb{P}}$  and  $\mathbb{F}^{\hat{W}, \hat{\mu}}$ , and we can define a new probability setting  $\hat{\mathbb{P}}^{\hat{v}}(d\hat{\omega}) = \kappa_T^{\hat{v}}(\hat{\omega})\hat{\mathbb{P}}(d\hat{\omega})$ . From the Girsanov theorem for multivariate point processes ([19]), it follows that under  $\hat{\mathbb{P}}^{\hat{v}}$  the  $\mathbb{F}^{\hat{W}, \hat{\mu}}$ -compensator of  $\hat{\mu}$  on the set  $[0, T] \times A$  is the random measure  $\hat{v}_t(a)\lambda(da) dt$ . Notice that  $\hat{B}$  remains a Brownian motion under  $\hat{\mathbb{P}}^{\hat{v}}$ , and using (2.2)–(2.3) we can generalize estimate (3.4) as follows:

$$(3.6) \quad \sup_{\hat{v} \in \mathcal{V}} \hat{\mathbb{E}}^{\hat{v}} \left[ \sup_{t \in [0, T]} |\hat{X}_t|^p \right] < \infty,$$

where  $\hat{\mathbb{E}}^{\hat{v}}$  denotes the expectation with respect to  $\hat{\mathbb{P}}^{\hat{v}}$ . We finally introduce the gain functional of the randomized control problem:

$$(3.7) \quad J^{\mathcal{R}}(\hat{v}) = \hat{\mathbb{E}}^{\hat{v}} \left[ \int_0^T f_t(\hat{X}, \hat{I}) dt + g(\hat{X}) \right].$$

The randomized stochastic optimal control problem consists in maximizing  $J^{\mathcal{R}}(\hat{v})$  over all  $\hat{v} \in \hat{\mathcal{V}}$ . Its value is defined as

$$(3.8) \quad \mathcal{U}_0^{\mathcal{R}} = \sup_{\hat{v} \in \hat{\mathcal{V}}} J^{\mathcal{R}}(\hat{v}).$$

REMARK 3.1. Let us define  $\hat{\mathcal{V}}_{\text{inf}>0} = \{\hat{v} \in \hat{\mathcal{V}} : \inf_{\hat{\Omega} \times [0, T] \times A} \hat{v} > 0\}$ . Then

$$(3.9) \quad \mathcal{U}_0^{\mathcal{R}} = \sup_{\hat{v} \in \hat{\mathcal{V}}_{\text{inf}>0}} J^{\mathcal{R}}(\hat{v}).$$

Indeed, given  $\hat{v} \in \hat{\mathcal{V}}$  and  $\epsilon > 0$ , define  $\hat{v}^\epsilon = \hat{v} \vee \epsilon \in \hat{\mathcal{V}}_{\text{inf}>0}$  and write the gain (3.7) in the form

$$J^{\mathcal{R}}(\hat{v}^\epsilon) = \hat{\mathbb{E}} \left[ \kappa_T^{\hat{v}^\epsilon} \left( \int_0^T f_t(\hat{X}, \hat{I}) dt + g(\hat{X}) \right) \right].$$

It is easy to see that  $J^{\mathcal{R}}(\hat{v}^\epsilon) \rightarrow J^{\mathcal{R}}(\hat{v})$  as  $\epsilon \rightarrow 0$ , which implies

$$\mathcal{U}_0^{\mathcal{R}} = \sup_{\hat{v} \in \hat{\mathcal{V}}} J^{\mathcal{R}}(\hat{v}) \leq \sup_{\hat{v} \in \hat{\mathcal{V}}_{\text{inf}>0}} J^{\mathcal{R}}(\hat{v}).$$

The other inequality being obvious, we obtain (3.9).

REMARK 3.2. We end this section noting that a randomized control problem can be constructed starting from the initial control problem with partial observation. Indeed, let  $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F}, V, W, x_0)$  be the setting for the stochastic optimal control problem formulated in Section 2.2. Suppose that  $(\Omega', \mathcal{F}', \mathbb{P}')$  is another probability space where a Poisson random measure  $\mu$  with intensity  $\lambda$  is defined (for instance by a classical result, see [34], Theorem 2.3.1, we may take  $\Omega' = [0, 1]$ ,  $\mathcal{F}'$  the corresponding Borel sets and  $\mathbb{P}'$  the Lebesgue measure). Then we define  $\bar{\Omega} = \Omega \times \Omega'$ , we denote by  $\bar{\mathcal{F}}$  the completion of  $\mathcal{F} \otimes \mathcal{F}'$  with respect

to  $\mathbb{P} \otimes \mathbb{P}'$  and by  $\bar{\mathbb{P}}$  the extension of  $\mathbb{P} \otimes \mathbb{P}'$  to  $\bar{\mathcal{F}}$ . The random elements  $V, W, x_0$  in  $\Omega$  and the random measure  $\mu$  in  $\Omega'$  have obvious extensions to  $\bar{\Omega}$ , that will be denoted by the same symbols. Clearly,  $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}}, V, W, \mu, x_0)$  is a setting for a randomized control problem as formulated before that we call *product extension* of the setting  $(\Omega, \mathcal{F}, \mathbb{P}, V, W, x_0)$  for the initial control problem (2.7).

We note that the initial formulation of a randomized setting  $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}, \hat{V}, \hat{W}, \hat{\mu}, \hat{x}_0)$  was more general, since it was not required that  $\hat{\Omega}$  should be a product space  $\Omega \times \Omega'$  and, even if it were the case, it was not required that the process  $\hat{B} = (\hat{V}, \hat{W})$  should depend only on  $\omega \in \Omega$  while the random measure  $\hat{\mu}$  should depend only on  $\omega' \in \Omega'$ .

3.2. *The value of the randomized control problem.* In this section, it is our purpose to show that the value  $\mathcal{U}_0^{\mathcal{R}}$  of the randomized control problem defined in (3.8) does not depend on the specific setting  $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}, \hat{V}, \hat{W}, \hat{\mu}, \hat{x}_0)$ , so that it is just a functional of the (deterministic) elements  $A, b, \sigma, f, g, \rho_0, \lambda, a_0$ . Later on, in Theorem 3.1, we will prove that in fact  $\mathcal{U}_0^{\mathcal{R}}$  does not depend on the choice of  $\lambda$  and  $a_0$  either.

So let now  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}, \tilde{V}, \tilde{W}, \tilde{\mu}, \tilde{x}_0)$  be another setting for the randomized control problem, as in Section 3.1, and let  $\mathbb{F}^{\tilde{W}, \tilde{\mu}}, \mathbb{F}^{\tilde{x}_0, \tilde{B}, \tilde{\mu}}, \tilde{X}, \tilde{I}, \tilde{\mathcal{V}}$  be defined in analogy with what was done before. So, for any  $\tilde{v} \in \tilde{\mathcal{V}}$ , we also define  $\kappa^{\tilde{v}}$  and the probability  $d\tilde{\mathbb{P}}^{\tilde{v}} = \kappa_T^{\tilde{v}} d\tilde{\mathbb{P}}$  as well as the gain and the value

$$\tilde{J}^{\mathcal{R}}(\tilde{v}) = \tilde{\mathbb{E}}^{\tilde{v}} \left[ \int_0^T f_t(\tilde{X}, \tilde{I}) dt + g(\tilde{X}) \right], \quad \tilde{\mathcal{U}}_0^{\mathcal{R}} = \sup_{\tilde{v} \in \tilde{\mathcal{V}}} \tilde{J}^{\mathcal{R}}(\tilde{v}).$$

We recall that the gain functional and value for the setting  $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}, \hat{V}, \hat{W}, \hat{\mu}, \hat{x}_0)$  was defined in (3.7) and (3.8) and denoted by  $J^{\mathcal{R}}$  and  $\mathcal{U}_0^{\mathcal{R}}$  rather than  $\hat{J}^{\mathcal{R}}$  and  $\hat{\mathcal{U}}_0^{\mathcal{R}}$ , to simplify the notation.

PROPOSITION 3.1. *With the previous notation, we have  $\mathcal{U}_0^{\mathcal{R}} = \tilde{\mathcal{U}}_0^{\mathcal{R}}$ . In other words,  $\mathcal{U}_0^{\mathcal{R}}$  only depends on the objects  $A, b, \sigma, f, g, \rho_0, \lambda, a_0$  appearing in the assumptions (A1) and (A2).*

PROOF. It is enough to prove that  $\mathcal{U}_0^{\mathcal{R}} \leq \tilde{\mathcal{U}}_0^{\mathcal{R}}$ , since the opposite inequality is established by the same arguments. Writing the gain  $J^{\mathcal{R}}(\hat{v})$  defined in (3.7) in the form

$$J^{\mathcal{R}}(\hat{v}) = \hat{\mathbb{E}} \left[ \kappa_T^{\hat{v}} \left( \int_0^T f_t(\hat{X}, \hat{I}) dt + g(\hat{X}) \right) \right],$$

recalling the definition (3.5) of the process  $\kappa^{\hat{v}}$  and noting that the process  $\hat{I}$  is completely determined by  $\hat{\mu}$ , we see that  $J^{\mathcal{R}}(\hat{v})$  only depends on the (joint) law of  $(\hat{X}, \hat{\mu}, \hat{v})$  under  $\hat{\mathbb{P}}$ . Since, however,  $\hat{X}$  is the solution to equation (3.2) with

initial condition  $\hat{X}_0 = \hat{x}_0$ , it is easy to check that under our assumptions the law of  $(\hat{X}, \hat{\mu}, \hat{\nu})$  only depends on the law of  $(\hat{x}_0, \hat{V}, \hat{W}, \hat{\mu}, \hat{\nu})$ . Since  $\hat{x}_0, \hat{V}$  and  $(\hat{W}, \hat{\mu}, \hat{\nu})$  are all independent, and the laws of  $\hat{x}_0$  and  $\hat{V}$  are fixed (since  $\hat{V}$  is a standard Wiener process and  $\hat{x}_0$  has law  $\rho_0$ ) we conclude that  $J^{\mathcal{R}}(\hat{\nu})$  only depends on the law of  $(\hat{W}, \hat{\mu}, \hat{\nu})$  under  $\hat{\mathbb{P}}$ . Similarly,  $\tilde{J}^{\mathcal{R}}(\tilde{\nu})$  only depends on the law of  $(\tilde{W}, \tilde{\mu}, \tilde{\nu})$  under  $\tilde{\mathbb{P}}$ .

Next, we claim that, given  $\hat{\nu} \in \hat{\mathcal{V}}$  there exists  $\tilde{\nu} \in \tilde{\mathcal{V}}$  such that the law of  $(\hat{W}, \hat{\mu}, \hat{\nu})$  under  $\hat{\mathbb{P}}$  is the same as the law of  $(\tilde{W}, \tilde{\mu}, \tilde{\nu})$  under  $\tilde{\mathbb{P}}$ . Assuming the claim for a moment, it follows from the previous discussion that for this choice of  $\tilde{\nu}$  we have

$$J^{\mathcal{R}}(\hat{\nu}) = \tilde{J}^{\mathcal{R}}(\tilde{\nu}) \leq \tilde{\mathcal{U}}_0^{\mathcal{R}},$$

and taking the supremum over  $\hat{\nu} \in \hat{\mathcal{V}}$  we deduce that  $\mathcal{U}_0^{\mathcal{R}} \leq \tilde{\mathcal{U}}_0^{\mathcal{R}}$ , which proves the result.

It only remains to prove the claim. By a monotone class argument, we may suppose that  $\hat{\nu}_t(a) = k(a)\phi_t\psi_t$ , where  $k$  is a  $\mathcal{B}(A)$ -measurable,  $\phi$  is  $\mathbb{F}^{\hat{W}}$ -predictable and  $\psi$  is  $\mathbb{F}^{\hat{\mu}}$ -predictable (where these filtrations are the ones generated by  $\hat{W}$  and  $\hat{\mu}$ , resp.). We may further suppose that  $\phi_t = 1_{(t_0, t_1]}(t)\phi_0(\hat{W}_{s_1}, \dots, \hat{W}_{s_h})$ , for an integer  $h$  and deterministic times  $0 \leq s_1 \leq \dots \leq s_h \leq t_0 < t_1$  and a Borel function  $\phi_0$  on  $\mathbb{R}^h$ , since this class of processes generates the predictable  $\sigma$ -algebra of  $\mathbb{F}^{\hat{W}}$ , and that  $\psi_t = 1_{(\hat{s}_n, \hat{s}_{n+1}]}(t)\psi_0(\hat{S}_1, \dots, \hat{S}_n, \hat{\eta}_1, \dots, \hat{\eta}_n, t)$ , for an integer  $n \geq 1$  and a Borel function  $\psi_0$  on  $\mathbb{R}^{2n+1}$ , since this class of processes generates the predictable  $\sigma$ -algebra of  $\mathbb{F}^{\hat{\mu}}$  [see [19], Lemma (3.3)]. It is immediate to verify that the required process  $\tilde{\nu}$  can be defined setting

$$\begin{aligned} \tilde{\nu}_t(a) &= k(a)1_{(t_0, t_1]}(t)\phi_0(\tilde{W}_{s_1}, \dots, \tilde{W}_{s_h})1_{(\tilde{s}_n, \tilde{s}_{n+1}]}(t) \\ &\quad \times \psi_0(\tilde{S}_1, \dots, \tilde{S}_n, \tilde{\eta}_1, \dots, \tilde{\eta}_n, t), \end{aligned}$$

where  $(\tilde{S}_n, \tilde{\eta}_n)_{n \geq 1}$  are associated to the measure  $\tilde{\mu}$ , that is,  $\tilde{\mu} = \sum_{n \geq 1} \delta_{(\tilde{S}_n, \tilde{\eta}_n)}$ .  $\square$

3.3. *Equivalence of the partially observed and the randomized control problem.*  
 We can now state one of the main results of the paper.

**THEOREM 3.1.** *Assume that (A1) and (A2) are satisfied. Then the values of the partially observed control problem and of the randomized control problem are equal:*

$$(3.10) \quad \mathcal{U}_0 = \mathcal{U}_0^{\mathcal{R}},$$

where  $\mathcal{U}_0$  and  $\mathcal{U}_0^{\mathcal{R}}$  are defined by (2.7) and (3.8), respectively. This common value only depends on the objects  $A, b, \sigma, f, g, \rho_0$  appearing in assumption (A1).

The last sentence follows immediately from Proposition 3.1, from the equality  $\mathcal{U}_0 = \mathcal{U}_0^{\mathcal{R}}$  and from the obvious fact that  $\mathcal{U}_0$  cannot depend on  $\lambda$ ,  $a_0$  introduced in assumption (A2). The proof of the equality is contained in the next section.

Before giving the proof of Theorem 3.1, let us discuss the significance of this equivalence result. The randomized control problem involves an uncontrolled state process  $(X, I)$  solution to (3.1)–(3.2), and the optimization is done over a set of equivalent probability measures whose effect is to change the characteristics (the intensity) of the auxiliary randomized process  $I$  without impacting on the Brownian motion  $B$  driving  $X$ . Therefore, the equivalence result (3.10) means that by performing such optimization in the randomized problem, we achieve the same value as in the original control problem where controls affect directly the drift and diffusion of the state process. As explained in the Introduction, such equivalence result has important implications that will be addressed in Section 5 where it is shown that the randomized control problem is associated by duality to a backward stochastic differential equation (with nonpositive jumps), called the *randomized equation*, which then also characterizes the value function of the initial control problem (2.7).

REMARK 3.3. We mention that in the article [15] an equivalence result similar to Theorem 3.1 was proved. However, in [15] only the case of full observation was addressed and there was no memory effect with respect to the control, whereas path-dependence in the state variable was allowed. But the main difference with respect to our setting is that in [15] the primal problem was formulated in a weak form, that is, taking the supremum of the gain functional (1.5) also over all possible choices of the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . This simplifies many arguments, and in particular makes the inequality  $\mathcal{U}_0 \geq \mathcal{U}_0^{\mathcal{R}}$  trivial.

**4. Proof of Theorem 3.1.** The proof is split into two parts, corresponding to the inequalities  $\mathcal{U}_0^{\mathcal{R}} \leq \mathcal{U}_0$  and  $\mathcal{U}_0 \leq \mathcal{U}_0^{\mathcal{R}}$ . In the sequel, (A1) and (A2) are always assumed to hold. However, instead of the inequality  $p \geq \max(2, 2r)$ , in (A1)(v) it is enough to suppose that  $p \geq \max(2, r)$ .

Before starting with the rigorous proof, let us have a look at the main points:

- $\mathcal{U}_0^{\mathcal{R}} \leq \mathcal{U}_0$ . First, we prove that the value of the primal problem  $\mathcal{U}_0$  does not change if we reformulate it on the enlarged probability space where the randomized problem lives, taking the supremum over  $\mathcal{A}^{W, \mu'}$ , which is the set of controls  $\bar{\alpha}$  progressively measurable with respect to the filtration generated by  $W$  and the Poisson random measure  $\mu'$  (Lemma 4.1; actually, we take  $\bar{\alpha}$  progressively measurable with respect to an even larger filtration, denoted  $\mathbb{F}^{W, \mu'_{\infty}}$ ). Second, we prove that for every  $\nu \in \mathcal{V}_{\text{inf}>0}$  there exists  $\bar{\alpha}^{\nu} \in \mathcal{A}^{W, \mu'}$  such that  $\mathcal{L}_{\mathbb{P}^{\nu}}(x_0, B, I) = \mathcal{L}_{\mathbb{P}}(x_0, B, \bar{\alpha}^{\nu})$  (Proposition 4.2). This result is a direct consequence of the key Lemma 4.3. From  $\mathcal{L}_{\mathbb{P}^{\nu}}(x_0, B, I) = \mathcal{L}_{\mathbb{P}}(x_0, B, \bar{\alpha}^{\nu})$ , we obtain

that  $J^{\mathcal{R}}(\nu) = \bar{J}(\bar{\alpha}^\nu)$ , namely

$$\mathcal{U}_0^{\mathcal{R}} := \sup_{\nu \in \mathcal{V}_{\inf>0}^{\mathcal{R}}} J^{\mathcal{R}}(\nu) = \sup_{\nu \in \mathcal{V}_{\inf>0}^{\bar{\alpha}^\nu}} \bar{J}(\bar{\alpha}^\nu).$$

Since  $\mathcal{U}_0 = \sup_{\bar{\alpha} \in \mathcal{A}^{W, \mu'}} \bar{J}(\bar{\alpha})$  by Lemma 4.1, and every  $\bar{\alpha}^\nu$  belongs to  $\mathcal{A}^{W, \mu'}$ , we easily obtain the inequality  $\mathcal{U}_0^{\mathcal{R}} \leq \mathcal{U}_0$ .

- $\mathcal{U}_0 \leq \mathcal{U}_0^{\mathcal{R}}$ . The proof of this inequality is based on a “density” result in the spirit of Lemma 3.2.6 in [25], which is Proposition 4.1 below. Before, we need to introduce some notation.

We take an auxiliary probability space denoted  $(\Omega', \mathcal{F}', \mathbb{P}')$ , where appropriate random objects are defined [see the Appendix of the longer version [3] of our paper for the precise requirements on  $(\Omega', \mathcal{F}', \mathbb{P}')$ ]. Then we construct the product space  $(\hat{\Omega}, \hat{\mathcal{F}}, \mathbb{Q})$ :

$$\hat{\Omega} = \Omega \times \Omega', \quad \hat{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}', \quad \mathbb{Q} = \mathbb{P} \otimes \mathbb{P}'.$$

The random variable  $x_0$  and the processes  $\alpha$  and  $B = (V, W)$  are extended to  $\hat{\Omega}$  in a natural way. We denote  $\hat{x}_0$  and  $\hat{\alpha}$  the extensions of  $x_0$  and  $\alpha$ . The extension of  $B$ , denoted  $\hat{B} = (\hat{V}, \hat{W})$ , remains a Wiener process under  $\mathbb{Q}$ . The filtration  $\mathbb{F}^W$  can also be canonically extended to a filtration in  $(\hat{\Omega}, \hat{\mathcal{F}})$ , which coincides with the filtration  $\mathbb{F}^{\hat{W}}$  generated by  $\hat{W}$ .

Following [25], for any pair  $\alpha^1, \alpha^2 : \hat{\Omega} \times [0, T] \rightarrow A$  of measurable processes in  $(\hat{\Omega}, \hat{\mathcal{F}}, \mathbb{Q})$  we define a distance  $\tilde{\rho}(\alpha^1, \alpha^2)$  setting

$$(4.1) \quad \tilde{\rho}(\alpha^1, \alpha^2) = \mathbb{E}^{\mathbb{Q}} \left[ \int_0^T \rho(\alpha_t^1, \alpha_t^2) dt \right],$$

where  $\mathbb{E}^{\mathbb{Q}}$  denotes the expectation under  $\mathbb{Q}$ , and  $\rho$  is a metric in  $A$  satisfying  $\rho < 1$ .

**PROPOSITION 4.1.** *Let  $A$  be a Borel space, let  $\lambda$  and  $a_0$  satisfy **(A2)** and let  $(\hat{\Omega}, \hat{\mathcal{F}}, \mathbb{Q})$  be the product space defined above. Then, for any  $\mathbb{F}^W$ -progressive  $A$ -valued process  $\alpha$  and for any  $\delta > 0$ , there exists a marked point process  $(\hat{S}_n, \hat{\eta}_n)_{n \geq 1}$  defined in  $(\hat{\Omega}, \hat{\mathcal{F}}, \mathbb{Q})$  satisfying the following conditions:*

1. *setting*

$$\hat{S}_0 = 0, \quad \hat{\eta}_0 = a_0, \quad \hat{I}_t = \sum_{n \geq 0} \hat{\eta}_n 1_{[\hat{S}_n, \hat{S}_{n+1})}(t),$$

*the process  $\hat{I}$  satisfies*

$$(4.2) \quad \tilde{\rho}(\hat{I}, \hat{\alpha}) = \mathbb{E}^{\mathbb{Q}} \left[ \int_0^T \rho(\hat{I}_t, \hat{\alpha}_t) dt \right] < \delta;$$

2. denoting  $\hat{\mu} = \sum_{n \geq 1} \delta_{(\hat{S}_n, \hat{\eta}_n)}$  the random measure associated to  $(\hat{S}_n, \hat{\eta}_n)_{n \geq 1}$ ,  $\mathbb{F}^{\hat{\mu}} = (\mathcal{F}_t^{\hat{\mu}})_{t \geq 0}$  the natural filtration of  $\hat{\mu}$  and  $\mathbb{F}^{\hat{W}} \vee \mathbb{F}^{\hat{\mu}} = (\mathcal{F}_t^{\hat{W}} \vee \mathcal{F}_t^{\hat{\mu}})_{t \geq 0}$ , then the  $\mathbb{F}^{\hat{W}} \vee \mathbb{F}^{\hat{\mu}}$ -compensator of  $\hat{\mu}$  under  $\mathbb{Q}$  is absolutely continuous with respect to  $\lambda(da) dt$  and it can be written in the form

$$(4.3) \quad \hat{v}_t(\hat{\omega}, a) \lambda(da) dt$$

for some nonnegative  $\mathcal{P}(\mathbb{F}^{\hat{W}} \vee \mathbb{F}^{\hat{\mu}}) \otimes \mathcal{B}(A)$ -measurable function  $\hat{v}$  satisfying

$$(4.4) \quad \inf_{\hat{\Omega} \times [0, T] \times A} \hat{v} > 0, \quad \sup_{\hat{\Omega} \times [0, T] \times A} \hat{v} < \infty.$$

Roughly speaking, we prove that the class  $\{\bar{\alpha}^v : v \in \mathcal{V}_{\text{inf} > 0}\}$  is dense in  $\mathcal{A}^{W, \mu'}$ , with respect to the metric  $\tilde{\rho}$  defined in (4.1) (the same metric used in Lemma 3.2.6 in [25]). The proof of Proposition 4.1 is quite technical, and is postponed in the Appendix of the longer version [3] of our paper. Then the inequality  $\mathcal{V}_0 \leq \mathcal{V}_0^{\mathcal{R}}$  follows from the stability Lemma 4.4, which states that, under Assumption (A1), the gain functional is continuous with respect to the metric  $\tilde{\rho}$ .

4.1. *Proof of the inequality  $\mathcal{V}_0^{\mathcal{R}} \leq \mathcal{V}_0$ .* We note at the outset that the requirement that  $\lambda$  has full support will not be used in the proof of the inequality  $\mathcal{V}_0^{\mathcal{R}} \leq \mathcal{V}_0$ .

Let  $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F}, V, W, x_0)$  be a setting for the stochastic optimal control problem with partial observation formulated in Section 2.2. We construct a setting for a randomized control problem in the form of a product extension as described at the end of Section 3.1.

Let  $\lambda$  be a Borel measure on  $A$  satisfying (A2). As a first step, we need to construct a suitable surjective measurable map  $\pi : \mathbb{R} \rightarrow A$  and to introduce a properly chosen measure  $\lambda'$  on the Borel subsets of the real line such that in particular  $\lambda = \lambda' \circ \pi^{-1}$ . We also recall that the space of control actions  $A$  is assumed to be a Borel space and it is known that any such space is either finite or countable (with the discrete topology) or isomorphic, as a measurable space, to the real line [or equivalently to the half line  $(0, \infty)$ ]: see, for example, [7], Corollary 7.16.1. Let us denote by  $A_c$  the subset of  $A$  consisting of all points  $a \in A$  such that  $\lambda(\{a\}) > 0$ , and let  $A_{nc} = A \setminus A_c$ . Since  $\lambda$  is finite, the set  $A_c$  is either empty or countable, and it follows in particular that both  $A_c$  and  $A_{nc}$  are also Borel spaces. In the construction of  $\lambda'$ , we distinguish three cases:

1.  $A_c = \emptyset$ , so that  $A = A_{nc}$  is uncountable. Then, as recalled above, there exists a bijection  $\pi : \mathbb{R} \rightarrow A$  such that  $\pi$  and its inverse are both Borel measurable. We define a measure  $\lambda'$  on  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  setting  $\lambda'(B) = \lambda(\pi(B))$  for  $B \in \mathcal{B}(\mathbb{R})$ . Even if we cannot guarantee that  $\lambda'$  has full support, it clearly holds that  $\lambda'(\{r\}) = 0$  for every  $r \in \mathbb{R}$ . Basically, in this case we are identifying  $A$  with  $\mathbb{R}$  and  $\lambda$  with its image measure  $\lambda'$ .

2.  $A_{nc} = \emptyset$ , so that  $A = A_c$  is countable, with the discrete topology. For every  $j \in A$ , choose a (nontrivial) interval  $\mathcal{I}_j \subset \mathbb{R}$  in such a way that  $\{\mathcal{I}_j, j \in A\}$  is a partition of  $\mathbb{R}$ . Choose an arbitrary nonatomic finite measure on  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  with full support (say, the standard Gaussian measure, denoted by  $\gamma$ ) and denote by  $\lambda'$  the unique positive measure on  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  such that

$$\lambda'(B) = \lambda(\{j\})\gamma(B)/\gamma(\mathcal{I}_j) \quad \text{for every } B \subset \mathcal{I}_j, B \in \mathcal{B}(\mathbb{R}), j \in A.$$

Notice that  $\lambda'$  is a finite measure ( $\lambda'(\mathbb{R}) = \lambda(A)$ ), satisfying  $\lambda'(\mathcal{I}_j) = \lambda(\{j\})$  for every  $j \in A$  and  $\lambda'(\{r\}) = 0$  for every  $r \in \mathbb{R}$ . We also define the projection  $\pi : \mathbb{R} \rightarrow A$  given by

$$(4.5) \quad \pi(r) = j \quad \text{if } r \in \mathcal{I}_j \text{ for some } j \in A.$$

Clearly,  $\lambda = \lambda' \circ \pi^{-1}$ .

3.  $A_c \neq \emptyset$  and  $A_{nc} \neq \emptyset$ . For every  $j \in A_c$  choose a (nontrivial) interval  $\mathcal{I}_j \subset (-\infty, 0]$  in such a way that  $\{\mathcal{I}_j, j \in A_c\}$  is a partition of  $(-\infty, 0]$ . Moreover, there exists a bijection  $\pi_1 : (0, \infty) \rightarrow A_{nc}$  such that  $\pi_1$  and its inverse are both Borel measurable. Denote by  $\lambda'$  the unique positive measure on  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  such that

$$\begin{aligned} \lambda'(B) &= \lambda(\{j\})\gamma(B)/\gamma(\mathcal{I}_j) && \text{for every } B \subset \mathcal{I}_j, B \in \mathcal{B}(\mathbb{R}), j \in A_c, \\ \lambda'(B) &= \lambda(\pi_1(B)) && \text{for every } B \subset (0, \infty), B \in \mathcal{B}(\mathbb{R}). \end{aligned}$$

Again,  $\lambda'$  is a finite measure satisfying  $\lambda'(\mathcal{I}_j) = \lambda(\{j\})$  for every  $j \in A_c$  and  $\lambda'(\{r\}) = 0$  for every  $r \in \mathbb{R}$ . We also define the projection  $\pi : \mathbb{R} \rightarrow A$  given by

$$(4.6) \quad \pi(r) = \begin{cases} j & \text{if } r \in \mathcal{I}_j \text{ for some } j \in A_c, \\ \pi_1(r) & \text{if } r \in (0, \infty), \end{cases}$$

so that in particular  $\lambda = \lambda' \circ \pi^{-1}$ .

Now let  $(\Omega', \mathcal{F}', \mathbb{P}')$  denote the canonical probability space of a nonexplosive Poisson point process on  $\mathbb{R}_+ \times \mathbb{R}$  with intensity  $\lambda'$ . Thus,  $\Omega'$  is the set of sequences  $\omega' = (t_n, r_n)_{n \geq 1} \subset (0, \infty) \times \mathbb{R}$  with  $t_n < t_{n+1} \nearrow \infty$ ,  $(T_n, R_n)_{n \geq 1}$  is the canonical marked point process (i.e.,  $T_n(\omega') = t_n$ ,  $R_n(\omega') = r_n$ ), and  $\mu' = \sum_{n \geq 1} \delta_{(T_n, R_n)}$  is the corresponding random measure. Let  $\mathcal{F}'$  denote the smallest  $\sigma$ -algebra such that all the maps  $T_n, R_n$  are measurable and  $\mathbb{P}'$  the unique probability on  $\mathcal{F}'$  such that  $\mu'$  is a Poisson random measure with intensity  $\lambda'$  (since  $\lambda'$  is a finite measure, this probability actually exists). We will also use the completion of the space  $(\Omega', \mathcal{F}', \mathbb{P}')$ , still denoted by the same symbol by abuse of notation. In all the cases considered above, setting

$$A_n = \pi(R_n), \quad \mu = \sum_{n \geq 1} \delta_{(T_n, A_n)},$$

it is easy to verify that  $\mu$  is a Poisson random measure on  $(0, \infty) \times A$  with intensity  $\lambda$ , defined in  $(\Omega', \mathcal{F}', \mathbb{P}')$ . Then, following (3.1), we associate to this Poisson random measure on  $(0, \infty) \times A$ , the  $A$ -valued process

$$I_t = \sum_{n \geq 0} A_n 1_{[T_n, T_{n+1})}(t), \quad t \geq 0,$$

where we use the convention that  $T_0 = 0$  and  $I_0 = a_0$  the point in assumption **(A2)**(ii). In  $(\Omega', \mathcal{F}')$ , we define the natural filtrations  $\mathbb{F}^\mu = (\mathcal{F}_t^\mu)_{t \geq 0}$ ,  $\mathbb{F}^{\mu'} = (\mathcal{F}_t^{\mu'})_{t \geq 0}$  given by

$$\begin{aligned} \mathcal{F}_t^\mu &= \sigma(\mu((0, s] \times C) : s \in [0, t], C \in \mathcal{B}(A)) \vee \mathcal{N}', \\ \mathcal{F}_t^{\mu'} &= \sigma(\mu'((0, s] \times B) : s \in [0, t], B \in \mathcal{B}(\mathbb{R})) \vee \mathcal{N}', \end{aligned}$$

where  $\mathcal{N}'$  denotes the family of  $\mathbb{P}'$ -null sets of  $\mathcal{F}'$ . We denote by  $\mathcal{P}(\mathbb{F}^\mu)$ ,  $\mathcal{P}(\mathbb{F}^{\mu'})$  the corresponding predictable  $\sigma$ -algebras. Note that  $\mathcal{F}_t^\mu \subset \mathcal{F}_t^{\mu'}$  and  $\mathcal{F}_\infty^{\mu'} = \mathcal{F}'$ .

Then we define  $\bar{\Omega} = \Omega \times \Omega'$ ; we denote by  $\bar{\mathcal{F}}$  the completion of  $\mathcal{F} \otimes \mathcal{F}'$  with respect to  $\mathbb{P} \otimes \mathbb{P}'$  and by  $\bar{\mathbb{P}}$  the extension of  $\mathbb{P} \otimes \mathbb{P}'$  to  $\bar{\mathcal{F}}$ . The random elements  $V, W, x_0$  in  $\Omega$  and the random measures  $\mu, \mu'$  in  $\Omega'$  have obvious extensions to  $\bar{\Omega}$ , that will be denoted by the same symbols. Then  $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}}, V, W, \mu, x_0)$  is a setting for a randomized control problem as formulated in Section 3.1. Recall that  $\mathbb{F}^W$  denotes the  $\mathbb{P}$ -completed filtration in  $(\Omega, \mathcal{F})$  generated by the Wiener process  $W$ . All filtrations  $\mathbb{F}^W, \mathbb{F}^\mu, \mathbb{F}^{\mu'}$  can also be lifted to filtrations in  $(\bar{\Omega}, \bar{\mathcal{F}})$ , and  $\bar{\mathbb{P}}$ -completed. In the sequel, it should be clear from the context whether they are considered as filtrations in  $(\bar{\Omega}, \bar{\mathcal{F}})$  or in their original spaces. As in Section 3.1, we define the filtration  $\mathbb{F}^{W, \mu} = (\mathcal{F}_t^{W, \mu})_{t \geq 0}$  in  $(\bar{\Omega}, \bar{\mathcal{F}})$  by

$$\mathcal{F}_t^{W, \mu} = \mathcal{F}_t^W \vee \mathcal{F}_t^\mu \vee \mathcal{N}$$

( $\mathcal{N}$  denotes the family of  $\bar{\mathbb{P}}$ -null sets of  $\bar{\mathcal{F}}$ ), we introduce the classes  $\mathcal{V}, \mathcal{V}_{\inf > 0}$  and, for any admissible control  $v \in \mathcal{V}$ , the corresponding martingale  $\kappa^v$ , the probability  $\mathbb{P}^v(d\omega d\omega') = \kappa_t^v(\omega, \omega') \bar{\mathbb{P}}(d\omega d\omega')$  and the gain  $J^{\mathcal{R}}(v)$ . For technical purposes, we need to introduce the set  $\mathcal{V}'$  of elements  $v' = v'_t(\omega', a) : \Omega' \times \mathbb{R}_+ \times A \rightarrow (0, \infty)$ , which are  $\mathcal{P}(\mathbb{F}^\mu) \otimes \mathcal{B}(A)$ -measurable and bounded. We also define another filtration  $\mathbb{F}^{W, \mu'_\infty} = (\mathcal{F}_t^{W, \mu'_\infty})_{t \geq 0}$  in  $(\bar{\Omega}, \bar{\mathcal{F}})$  setting

$$\mathcal{F}_t^{W, \mu'_\infty} = \mathcal{F}_t^W \vee \mathcal{F}' \vee \mathcal{N}$$

(here  $\mathcal{F}'$  denotes a  $\sigma$ -algebra in  $(\bar{\Omega}, \bar{\mathcal{F}})$ , namely  $\{\Omega \times B : B \in \mathcal{F}'\}$ ).

In order to prove the inequality  $\mathcal{U}_0^{\mathcal{R}} \leq \mathcal{U}_0$ , we first prove two technical lemmata. In particular, in Lemma 4.1 we show that the primal problem is equivalent to a new primal problem with  $\mathbb{F}^{W, \bar{\mu}_\infty}$ -progressive controls on the enlarged space  $(\bar{\Omega}, \bar{\mathcal{F}})$ .



LEMMA 4.1. We have  $\mathcal{U}_0 = \sup_{\bar{\alpha} \in \mathcal{A}^{W, \mu'}} \bar{J}(\bar{\alpha})$ , where

$$\bar{J}(\bar{\alpha}) = \bar{\mathbb{E}} \left[ \int_0^T f_t(X^{\bar{\alpha}}, \bar{\alpha}) dt + g(X^{\bar{\alpha}}) \right],$$

and  $\mathcal{A}^{W, \mu'}$  is the set of all  $\mathbb{F}^{W, \mu'}$ -progressive processes  $\bar{\alpha}$  with values in  $A$ . Moreover,  $X^{\bar{\alpha}} = (X_t^{\bar{\alpha}})_{0 \leq t \leq T}$  is the strong solution to (2.5) (with  $\bar{\alpha}$  in place of  $\alpha$ ) satisfying  $X_0^{\bar{\alpha}} = x_0$ , which is unique in the class of continuous processes adapted to the filtration  $(\mathcal{F}_t^B \vee \sigma(x_0) \vee \mathcal{F}' \vee \mathcal{N})_{t \geq 0}$ .

PROOF. The inequality  $\mathcal{U}_0 \leq \sup_{\bar{\alpha} \in \mathcal{A}^{W, \mu'}} \bar{J}(\bar{\alpha})$  is immediate, since every control  $\alpha \in \mathcal{A}^W$  also lies in  $\mathcal{A}^{W, \mu'}$  and  $J(\alpha) = \bar{J}(\alpha)$ , whence  $J(\alpha) \leq \sup_{\bar{\alpha} \in \mathcal{A}^{W, \mu'}} \bar{J}(\bar{\alpha})$  and so  $\mathcal{U}_0 = \sup_{\alpha \in \mathcal{A}^W} J(\alpha) \leq \sup_{\bar{\alpha} \in \mathcal{A}^{W, \mu'}} \bar{J}(\bar{\alpha})$ .

Let us prove the opposite inequality. Fix  $\tilde{\alpha} \in \mathcal{A}^{W, \mu'}$  and consider the (un-completed) filtration  $\mathbb{F}'' := (\mathcal{F}_t^W \vee \mathcal{F}')_{t \geq 0}$ . Then we can find an  $A$ -valued  $\mathbb{F}''$ -progressive process  $\bar{\alpha}$  such that  $\bar{\alpha} = \tilde{\alpha} \bar{\mathbb{P}}(d\bar{\omega})$   $dt$ -almost surely, so that in particular  $\bar{J}(\bar{\alpha}) = \bar{J}(\tilde{\alpha})$ . It is easy to verify that, for every  $\omega' \in \Omega'$ , the process  $\alpha^{\omega'}$ , defined by  $\alpha_t^{\omega'}(\omega) := \bar{\alpha}_t(\omega, \omega')$ , is  $\mathbb{F}^W$ -progressive. Consider now the controlled equation on  $[0, T]$

$$\begin{aligned} X_t &= x_0 + \int_0^t b_s(X, \alpha^{\omega'}) ds + \int_0^t \sigma_s(X, \alpha^{\omega'}) dB_s \\ (4.7) \quad &= x_0 + \int_0^t b_s(X, \bar{\alpha}(\cdot, \omega')) ds + \int_0^t \sigma_s(X, \bar{\alpha}(\cdot, \omega')) dB_s. \end{aligned}$$

From the first line of (4.7), we see that, under Assumption (A1), for every  $\omega'$  there exists a unique (up to indistinguishability) continuous process  $X^{\alpha^{\omega'}} = (X_t^{\alpha^{\omega'}})_{0 \leq t \leq T}$  strong solution to (4.7), adapted to the filtration  $(\mathcal{F}_t^B \vee \sigma(x_0) \vee \mathcal{N})_{t \geq 0}$ . On the other hand, from the second line of (4.7), it follows that the process  $X^{\bar{\alpha}}(\cdot, \omega') = (X_t^{\bar{\alpha}}(\cdot, \omega'))_{0 \leq t \leq T}$  solves the above equation. From the path-wise uniqueness of strong solutions to equation (4.7), it follows that  $X_t^{\alpha^{\omega'}}(\omega) = X_t^{\bar{\alpha}}(\omega, \omega')$ , for all  $t \in [0, T]$ ,  $\mathbb{P}(d\omega)$ -a.s. By the Fubini theorem,

$$\bar{J}(\tilde{\alpha}) = \bar{J}(\bar{\alpha}) = \int_{\Omega'} \mathbb{E} \left[ \int_0^T f_t(X^{\alpha^{\omega'}}, \alpha^{\omega'}) dt + g(X^{\alpha^{\omega'}}) \right] \mathbb{P}'(d\omega').$$

Since the inner expectation equals the gain  $J(\alpha^{\omega'})$ , it cannot exceed  $V$  and it follows that  $\bar{J}(\tilde{\alpha}) \leq \mathcal{U}_0$ . The claim follows from the arbitrariness of  $\tilde{\alpha}$ .  $\square$

The next result provides a decomposition of any element  $\nu \in \mathcal{V}$ , that is,  $\mathcal{P}(\mathbb{F}^{W, \mu}) \otimes \mathcal{B}(A)$ -measurable and bounded.

LEMMA 4.2. (i) Let  $v \in \mathcal{V}$ , then there exists a  $\bar{\mathbb{P}}$ -null set  $\bar{N} \in \mathcal{N}$  such that  $v$  admits the following representation:

$$\begin{aligned}
 v_t(\omega, \omega', a) &= v_t^{(0)}(\omega, a) \mathbf{1}_{\{0 < t \leq T_1(\omega')\}} \\
 &\quad + \sum_{n=1}^{\infty} v_t^{(n)}(\omega, (T_1(\omega'), A_1(\omega')), \dots, (T_n(\omega'), A_n(\omega')), a) \\
 &\quad \times \mathbf{1}_{\{T_n(\omega') < t \leq T_{n+1}(\omega')\}},
 \end{aligned}$$

for all  $(\omega, \omega', t, A) \in \bar{\Omega} \times \mathbb{R}_+ \times A$ ,  $(\omega, \omega') \notin \bar{N}$ , for some maps  $v^{(n)} : \Omega \times \mathbb{R}_+ \times (\mathbb{R}_+ \times A)^n \times A \rightarrow (0, \infty)$ ,  $n \geq 1$ , [resp.,  $v^{(0)} : \Omega \times \mathbb{R}_+ \times A \rightarrow (0, \infty)$ ], which are  $\mathcal{P}(\mathbb{F}^W) \otimes \mathcal{B}((\mathbb{R}_+ \times A)^n) \otimes \mathcal{B}(A)$ -measurable (resp.,  $\mathcal{P}(\mathbb{F}^W) \otimes \mathcal{B}(A)$ -measurable) and uniformly bounded with respect to  $n$ . Moreover, if  $v \in \mathcal{V}_{\inf > 0}$  then  $\inf_{\bar{\Omega} \times [0, T] \times A} v^{(n)} > 0$  as well, for every  $n \geq 0$ .

(ii) Let  $v \in \mathcal{V}$ , then there exists  $\tilde{N} \in \mathcal{F}$ , with  $\mathbb{P}(\tilde{N}) = 0$ , such that the map  $v^\omega = v_t^\omega(\omega', a) : \Omega' \times \mathbb{R}_+ \times A \rightarrow (0, \infty)$ , defined by

$$v_t^\omega(\omega', a) := v_t(\omega, \omega', a), \quad (\omega', t, a) \in \Omega' \times \mathbb{R}_+ \times A,$$

belongs to  $\mathcal{V}'$  whenever  $\omega \notin \tilde{N}$ . Moreover, for every  $\omega \notin \tilde{N}$  there exists  $N_\omega \in \mathcal{N}'$  such that

$$\begin{aligned}
 v_t^\omega(\omega', a) &= v_t^{(0)}(\omega, a) \mathbf{1}_{\{0 < t \leq T_1(\omega')\}} \\
 (4.8) \quad &\quad + \sum_{n=1}^{\infty} v_t^{(n)}(\omega, (T_1(\omega'), A_1(\omega')), \dots, (T_n(\omega'), A_n(\omega')), a) \\
 &\quad \times \mathbf{1}_{\{T_n(\omega') < t \leq T_{n+1}(\omega')\}},
 \end{aligned}$$

for all  $(\omega', t, a) \in \Omega' \times \mathbb{R}_+ \times A$ ,  $\omega' \notin N'_\omega$ , where, clearly,  $v^{(n)}(\omega, \cdot)$  [resp.,  $v^{(0)}(\omega, \cdot)$ ] is  $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}((\mathbb{R}_+ \times A)^n) \otimes \mathcal{B}(A)$ -measurable (resp.  $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(A)$ -measurable).

PROOF. The proof is an extension of the results in [19], Lemma 3.3; it is based on monotone class arguments and is left to the reader.  $\square$

By Lemma 4.2(ii), given  $v \in \mathcal{V}$ , consider the process  $v^\omega \in \mathcal{V}'$ , with corresponding  $\mathbb{P}$ -null set  $\tilde{N} \in \mathcal{F}$ . Define the Doléans exponential process  $\kappa^{v^\omega}$  by formula (3.5) with  $v^\omega$  in place of  $v$ . Notice that by Lemma 4.2(ii) we have  $\kappa_t^{v^\omega}(\omega') = \kappa_t^v(\omega, \omega')$ , for all  $(\omega', t) \in \Omega' \times \mathbb{R}_+$ , whenever  $\omega \notin \tilde{N}$ . Moreover, for  $\omega \notin \tilde{N}$ ,  $(\kappa_t^{v^\omega})_{t \geq 0}$  is a martingale with respect to  $\mathbb{P}'$  and  $\mathbb{F}^\mu$ . We claim that there exists a unique probability measure  $\mathbb{P}^{v^\omega}$  on  $(\Omega', \mathcal{F}_\infty^\mu)$  such that  $\mathbb{P}^{v^\omega}(d\omega') = \kappa_t^{v^\omega}(\omega') \mathbb{P}'(d\omega')$  on each  $\sigma$ -algebra  $\mathcal{F}_t^\mu$ , and by the Girsanov theorem, the  $\mathbb{F}^\mu$ -compensator of  $\mu$  under  $\mathbb{P}^{v^\omega}$  is given by the right-hand side of (4.8).

The verification of the claim is a standard argument: using the boundedness of  $\nu$ , one first verifies that

$$\kappa_{t \wedge T_n}^{\nu^\omega}(\omega') \leq a_n e^{bT_n(\omega')}$$

for some constants  $a_n, b$ , which implies that  $(\kappa_{t \wedge T_n}^\nu)_{t \geq 0}$  is a uniformly integrable martingale with respect to  $\mathbb{P}'$  and  $\mathbb{F}^\mu$ . Then the probabilities  $\mathbb{P}_n^{\nu^\omega}$  defined on  $\mathcal{F}_{T_n}^\mu$  setting  $\mathbb{P}_n^{\nu^\omega}(d\omega') = \kappa_{T_n}^{\nu^\omega}(\omega')\mathbb{P}'(d\omega')$  satisfy the compatibility condition:  $\mathbb{P}_{n+1}^{\nu^\omega} = \mathbb{P}_n^{\nu^\omega}$  on  $\mathcal{F}_{T_n}^\mu$  for every  $n$ . Arguing as in Theorem 3.6 in [19], by the Kolmogorov extension theorem there exists a unique probability  $\mathbb{P}^{\nu^\omega}$  on  $(\Omega', \mathcal{F}_\infty^\mu)$  such that  $\mathbb{P}^{\nu^\omega} = \mathbb{P}_n^{\nu^\omega}$  on each  $\mathcal{F}_{T_n}^\mu$ , and  $\mathbb{P}^{\nu^\omega}$  has the required properties.

We can now state the following key result (Lemma 4.3) from which the required conclusion of this subsection follows readily (see Proposition 4.2). Recall that  $(\Omega', \mathcal{F}', \mathbb{P}')$  denotes the canonical probability space constructed above.

LEMMA 4.3. *Given  $\nu \in \mathcal{V}_{\text{inf}>0}$ , there exist a sequence  $(T_n^\nu, A_n^\nu)_{n \geq 1}$  on  $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$  and a  $\mathbb{P}$ -null set  $N \in \mathcal{F}$ , with  $\tilde{N} \subset N$  ( $\tilde{N}$  is the set appearing in Lemma 4.2(ii)), such that:*

- (i) *for every  $n \geq 1$ ,  $(T_n^\nu, A_n^\nu)$  takes values in  $(0, \infty) \times A$  and  $T_n^\nu < T_{n+1}^\nu$ ;*
- (ii) *for every  $n \geq 1$ ,  $T_n^\nu$  is an  $\mathbb{F}^{W, \mu'_\infty}$ -stopping time and  $A_n^\nu$  is  $\mathcal{F}_{T_n^\nu}^{W, \mu'_\infty}$ -measurable;*
- (iii)  *$\lim_{n \rightarrow \infty} T_n^\nu = \infty$ ;*
- (iv) *for every  $\omega \notin N$ , we have*

$$\mathcal{L}_{\mathbb{P}'}((T_n^\nu(\omega, \cdot), A_n^\nu(\omega, \cdot))_{n \geq 1}) = \mathcal{L}_{\mathbb{P}^{\nu^\omega}}((T_n, A_n)_{n \geq 1}).$$

Finally, let  $\bar{\alpha}_t^\nu = a_0 1_{[0, T_1^\nu)} + \sum_{n=1}^\infty A_n^\nu 1_{[T_n^\nu, T_{n+1}^\nu)}(t)$  be the step process associated with  $(T_n^\nu, A_n^\nu)_{n \geq 1}$ . Then  $\bar{\alpha}^\nu \in \mathcal{A}^{W, \mu'}$  and  $\mathcal{L}_{\mathbb{P}'}(\bar{\alpha}^\nu(\omega, \cdot)) = \mathcal{L}_{\mathbb{P}^{\nu^\omega}}(I), \omega \notin N$ .

PROOF. Suppose that we have already constructed a multivariate point process  $(T_n^\nu, A_n^\nu)_{n \geq 1}$  satisfying points (i)–(ii)–(iii)–(iv) of the theorem. Then, by (ii), it follows that  $\bar{\alpha}^\nu$  is càdlàg and  $\mathbb{F}^{W, \mu'_\infty}$ -adapted, hence progressive. Moreover, by (iii), for every  $(\bar{\omega}, t) \in \bar{\Omega} \times [0, T]$  the series  $\sum_{n=1}^\infty A_n^\nu(\bar{\omega}) 1_{[T_n^\nu(\bar{\omega}), T_{n+1}^\nu)}(t)$  is a finite sum, and thus  $\bar{\alpha}^\nu \in \mathcal{A}^{W, \mu'}$ . Furthermore, by (iv), we see that  $\mathcal{L}_{\mathbb{P}'}(\bar{\alpha}^\nu(\omega, \cdot)) = \mathcal{L}_{\mathbb{P}^{\nu^\omega}}(I), \omega \notin N$ .

Let us now construct  $(T_n^\nu, A_n^\nu)_{n \geq 1}$  satisfying points (i)–(ii)–(iii)–(iv). Fix  $\nu \in \mathcal{V}_{\text{inf}>0}$  and let  $\tilde{N} \in \mathcal{F}$  be as in Lemma 4.2. In particular, recall that formula (4.8) holds for some maps  $\nu^{(n)}, n \geq 0$ , satisfying  $0 < \inf \nu^{(n)} \leq \sup \nu^{(n)} \leq M^\nu$ , for some constant  $M^\nu > 0$ , independent of  $n$ . Next, recall the construction of the map  $\pi : \mathbb{R} \rightarrow A$  and the measure  $\lambda'$ . Accordingly, we split the rest of the proof into two cases.

Case I:  $A_c = \emptyset$ , so that  $A = A_{nc}$  is uncountable. In this case,  $\pi : \mathbb{R} \rightarrow A$  is a Borel isomorphism, so to shorten notation we identify  $A$  with  $\mathbb{R}$  and use the notation  $A, \lambda, A_n, \mu, \mathbb{F}^{W, \mu_\infty} = (\mathcal{F}_t^{W, \mu_\infty})_{t \geq 0}$  instead of  $\mathbb{R}, \lambda', R_n, \bar{\mu}, \mathbb{F}^{W, \mu'_\infty} = (\mathcal{F}_t^{W, \mu'_\infty})_{t \geq 0}$ . Since we are treating the case  $A_c = \emptyset$ , we have  $\lambda(\{a\}) = 0$  for every  $a \in A$ . We construct by induction on  $n \geq 1$  a sequence  $(T_n^\nu, A_n^\nu)_{n \geq 1}$  and a  $\mathbb{P}$ -null set  $N \in \mathcal{F}$ , with  $\tilde{N} \subset N$ , such that  $(T_n^\nu, A_n^\nu)_{n \geq 1}$  satisfies properties (i) and (ii) of the theorem, and also the following properties:

- (iii)' for every  $n \geq 1$ , we have  $T_n^\nu \geq T_n/M^\nu$ ;
- (iv)' for every  $n \geq 1$  and  $\omega \notin N$ , we have

$$(4.9) \quad \begin{aligned} &\mathcal{L}_{\mathbb{P}^\nu}(T_1^\nu(\omega, \cdot), A_1^\nu(\omega, \cdot), \dots, T_n^\nu(\omega, \cdot), A_n^\nu(\omega, \cdot)) \\ &= \mathcal{L}_{\mathbb{P}^\nu}^\omega(T_1, A_1, \dots, T_n, A_n). \end{aligned}$$

Notice that (iv)' is equivalent to (iv). Moreover, since  $\lim_{n \rightarrow \infty} T_n = \infty$ , we see that (iii)' implies property (iii).

Step 1: the case  $n = 1$ . Define

$$(4.10) \quad \theta_t^{(1)}(\omega) := \frac{1}{\lambda(A)} \int_0^t \int_A v_s^{(0)}(\omega, a) \lambda(da) ds.$$

Since  $0 < \inf v^{(0)} \leq \sup v^{(0)} \leq M^\nu$ , we see that, for every  $\omega \in \Omega$ , the map  $t \mapsto \theta_t^{(1)}(\omega)$  is continuous, strictly increasing,  $\theta_0^{(1)}(\omega) = 0, \theta_t^{(1)}(\omega) \leq M^\nu t$ , and  $\theta_t^{(1)}(\omega) \nearrow \infty$  as  $t$  goes to infinity. Then there exists a unique  $T_1^\nu : \bar{\Omega} \rightarrow \mathbb{R}_+$  such that

$$\theta_{T_1^\nu(\bar{\omega})}^{(1)}(\omega) = T_1(\omega').$$

Notice that  $T_1^\nu \geq T_1/M_1^\nu$ . Moreover, since  $T_1 > 0$ , we also have  $T_1^\nu > 0$ . Let  $\bar{E}_{T_1} := \{(\bar{\omega}, t) \in \bar{\Omega} \times \mathbb{R}_+ : \theta_t^{(1)}(\omega) = T_1(\omega')\}$ . Since the process  $(\bar{\omega}, t) \mapsto (\theta_t^{(1)}(\omega), T_1(\omega'))$  is  $\mathbb{F}^{W, \mu_\infty}$ -adapted and continuous,  $\bar{E}_{T_1}$  is an  $\mathbb{F}^{W, \mu_\infty}$ -optional set (in fact, predictable). Since  $T_1^\nu(\bar{\omega}) = \inf\{t \in \mathbb{R}_+ : (\bar{\omega}, t) \in \bar{E}_{T_1}\}$  is the debut of  $\bar{E}_{T_1}$ , from Theorem 1.14 of [20] it follows that  $T_1^\nu$  is an  $\mathbb{F}^{W, \mu_\infty}$ -stopping time. In particular,  $T_1^\nu$  is  $\bar{\mathcal{F}}$ -measurable, therefore there exists a  $\mathbb{P}$ -null set  $N_{T_1^\nu} \in \mathcal{F}$  such that  $T_1^\nu(\omega, \cdot)$  is  $\mathcal{F}'$ -measurable, whenever  $\omega \notin N_{T_1^\nu}$ .

Now define

$$F_b := \mathbb{P}'(A_1 \leq b) = \frac{\lambda((-\infty, b])}{\lambda(A)}, \quad F_b^{(1)}(\bar{\omega}) := \frac{\int_{-\infty}^b v_{T_1^\nu(\bar{\omega})}^{(0)}(\omega, a) \lambda(da)}{\int_{-\infty}^{+\infty} v_{T_1^\nu(\bar{\omega})}^{(0)}(\omega, a) \lambda(da)}.$$

Since  $\inf v^{(0)} > 0$  and  $\lambda(\{a\}) = 0$  for any  $a \in A$ , we see that, for every  $\bar{\omega} \in \bar{\Omega}$ , the map  $b \mapsto F_b^{(1)}(\bar{\omega})$  is continuous, strictly increasing, valued in  $(0, 1)$ , and

$\lim_{b \rightarrow -\infty} F_b^{(1)}(\bar{\omega}) = 0$ ,  $\lim_{b \rightarrow +\infty} F_b^{(1)}(\bar{\omega}) = 1$ . Then there exists a unique  $A_1^v : \bar{\Omega} \rightarrow \mathbb{R}$  such that

$$F_{A_1^v(\bar{\omega})}^{(1)}(\bar{\omega}) = F_{A_1(\omega')}.$$

We note that the process

$$(\bar{\omega}, t) \mapsto \frac{\int_{-\infty}^b v_t^{(0)}(\omega, a)\lambda(da)}{\int_{-\infty}^{+\infty} v_t^{(0)}(\omega, a)\lambda(da)}$$

is predictable with respect to  $\mathbb{F}^W$ , hence it is also  $\mathbb{F}^{W, \mu_\infty}$ -progressive. Substituting  $t$  with  $T_1^v(\bar{\omega})$ , we conclude that  $F_b^{(1)}$  is  $(\mathcal{F}_{T_1^v}^{W, \mu_\infty})$ -measurable. Since  $A_1$  is clearly  $\mathcal{F}'$ -measurable and  $\mathcal{F}' \subset \mathcal{F}_0^{W, \mu_\infty} \subset \mathcal{F}_{T_1^v}^{W, \mu_\infty}$ ,  $A_1$  is also  $(\mathcal{F}_{T_1^v}^{W, \mu_\infty})$ -measurable. Recalling the continuity of  $b \mapsto F_b^{(1)}(\bar{\omega})$ , it is easy to conclude that  $A_1^v$  is  $(\mathcal{F}_{T_1^v}^{W, \mu_\infty})$ -measurable. This implies that  $A_1^v \vee a$  is also  $\mathcal{F}_{T_1^v}^{W, \mu_\infty}$ -measurable. From the arbitrariness of  $a$ , we deduce that  $A_1^v$  is  $\mathcal{F}_{T_1^v}^{W, \mu_\infty}$ -measurable. In particular,  $A_1^v$  is  $\bar{\mathcal{F}}$ -measurable, therefore, there exists a  $\mathbb{P}$ -null set  $N_{A_1^v} \in \mathcal{F}$  such that  $A_1^v(\omega, \cdot)$  is  $\mathcal{F}'$ -measurable, whenever  $\omega \notin N_{A_1^v}$ .

In order to conclude the proof of the case  $n = 1$ , let us prove that (4.9) holds for  $n = 1$ , whenever  $\omega \notin N_1 := \tilde{N} \cup N_{T_1^v} \cup N_{A_1^v}$ . We begin recalling that, for every  $\omega \notin \tilde{N}$ , the  $\mathbb{F}^\mu$ -compensator of  $\mu$  under  $\mathbb{P}^{\nu^\omega}$  is given by the right-hand side of (4.8), so that in particular we have

$$\begin{aligned} \mathbb{P}^{\nu^\omega}(T_1 > t) &= \exp\left(-\int_0^t \int_A v_s^{(0)}(\omega, a)\lambda(da) ds\right) \\ &= \exp(-\lambda(A)\theta_t^{(1)}(\omega)). \end{aligned}$$

Notice that

$$\begin{aligned} \mathbb{P}'(T_1^v(\omega, \cdot) > t) &= \mathbb{P}'(\theta_{T_1^v(\omega, \cdot)}^{(1)}(\omega) > \theta_t^{(1)}(\omega)) \\ &= \mathbb{P}'(T_1 > \theta_t^{(1)}(\omega)) = \exp(-\lambda(A)\theta_t^{(1)}(\omega)), \end{aligned}$$

for every  $\omega \notin N_{T_1^v}$ , where for the last equality we used the formula  $\mathbb{P}'(T_1 > t) = \exp(-\lambda(A)t)$ . Therefore,  $\mathcal{L}_{\mathbb{P}'}(T_1^v(\omega, \cdot)) = \mathcal{L}_{\mathbb{P}^{\nu^\omega}}(T_1)$ , for every  $\omega \notin \tilde{N} \cup N_{T_1^v}$ . Now, recall that, for every  $\omega \notin \tilde{N}$ , we have,  $\mathbb{P}'$ -a.s.,

$$\mathbb{P}^{\nu^\omega}(A_1 \leq b \mid \sigma(T_1)) = \frac{\int_{-\infty}^b v_{T_1}^{(0)}(\omega, a)\lambda(da)}{\int_{-\infty}^{+\infty} v_{T_1}^{(0)}(\omega, a)\lambda(da)}.$$

On the other hand, for every  $\omega \notin N_{T_1^v} \cup N_{A_1^v}$ ,  $\mathbb{P}'$ -a.s.,

$$\begin{aligned} \mathbb{P}'(A_1^v(\omega, \cdot) \leq b \mid \sigma(T_1^v(\omega, \cdot))) &= \mathbb{P}'(F_{A_1^v(\omega, \cdot)}^{(1)}(\omega, \cdot) \leq F_b^{(1)}(\omega, \cdot) \mid \sigma(T_1^v(\omega, \cdot))) \\ &= \mathbb{P}'(F_{A_1} \leq F_b^{(1)}(\omega, \cdot) \mid \sigma(T_1^v(\omega, \cdot))). \end{aligned}$$

Since  $A_1$  is independent of  $T_1$  under  $\mathbb{P}'$  and  $T_1^v(\omega, \cdot)$  is  $\sigma(T_1)$ -measurable, it follows that  $A_1$  is also independent of  $T_1^v(\omega, \cdot)$ . Moreover, by definition we see that  $F_b^{(1)}(\omega, \cdot)$  is  $\sigma(T_1^v(\omega, \cdot))$ -measurable. Therefore, for every  $\omega \notin N_{T_1^v} \cup N_{A_1^v}$ , we have,  $\mathbb{P}'$ -a.s.,

$$\mathbb{P}'(F_{A_1} \leq F_b^{(1)}(\omega, \cdot) \mid \sigma(T_1^v(\omega, \cdot))) = \mathbb{P}'(F_{A_1} \leq a) \Big|_{a=F_b^{(1)}(\omega, \cdot)} = F_b^{(1)}(\omega, \cdot),$$

where we used the fact that  $F_{A_1}$  is uniformly distributed in  $(0, 1)$  under  $\mathbb{P}'$ . As a consequence, recalling that  $\mathcal{L}_{\mathbb{P}'}(T_1^v(\omega, \cdot)) = \mathcal{L}_{\mathbb{P}^{v\omega}}(T_1)$ , for every  $\omega \notin \tilde{N} \cup N_{T_1^v}$ , we deduce that  $\mathcal{L}_{\mathbb{P}'}(T_1^v(\omega, \cdot), A_1^v(\omega, \cdot)) = \mathcal{L}_{\mathbb{P}^{v\omega}}(T_1, A_1)$ , whenever  $\omega \notin N_1$ . This concludes the proof of the base case  $n = 1$ .

*Step 2: the inductive step.* Fix  $n \geq 1$  and suppose we are given  $(T_1^v, A_1^v), \dots, (T_n^v, A_n^v)$  satisfying points (i) and (ii) of the theorem. Suppose also that (4.9) holds for the fixed  $n$ , whenever  $\omega \notin N_n$ , for some  $\mathbb{P}$ -null set  $N_n \in \mathcal{F}$  in place of  $N$ , with  $\tilde{N} \subset N_n$ .

Given  $\theta^{(1)}$  as in (4.10), we define recursively, for  $i = 1, \dots, n$ ,

$$(4.11) \quad \begin{aligned} \theta_t^{(i+1)}(\bar{\omega}) &:= \theta_{T_i^v(\bar{\omega}) \wedge t}^{(i)}(\bar{\omega}) \\ &+ \frac{1}{\lambda(A)} \int_{T_i^v(\bar{\omega})}^{T_i^v(\bar{\omega}) \vee t} \int_A v_s^{(i)}(\omega, (T_1^v(\bar{\omega}), A_1^v(\bar{\omega})), \dots, \\ &(T_i^v(\bar{\omega}), A_i^v(\bar{\omega})), a) \lambda(da) ds. \end{aligned}$$

Since  $0 < \inf v^{(i)} \leq \sup v^{(i)} \leq M^v$ , we see that, for every  $\bar{\omega} \in \bar{\Omega}$  and  $i = 1, \dots, n$ , the map  $t \mapsto \theta_t^{(i+1)}(\bar{\omega})$  is continuous, strictly increasing,  $\theta_0^{(i+1)}(\bar{\omega}) = 0$ ,  $\theta_t^{(i+1)}(\bar{\omega}) \leq M^v t$ , and  $\theta_t^{(i+1)}(\bar{\omega}) \nearrow \infty$  as  $t$  goes to infinity. Then there exists a unique  $T_{n+1}^v : \bar{\Omega} \rightarrow \mathbb{R}_+$  such that

$$\theta_{T_{n+1}^v(\bar{\omega})}^{(n+1)}(\bar{\omega}) = T_{n+1}(\omega').$$

Notice that  $T_{n+1}^v \geq T_{n+1}/M^v$ . Moreover, since  $T_{n+1} > T_n$ , we also have  $T_{n+1}^v > T_n^v$ . Indeed, arguing by contradiction, suppose that  $T_{n+1}^v(\bar{\omega}) \leq T_n^v(\bar{\omega})$  for some  $\bar{\omega} \in \bar{\Omega}$ . Then

$$\theta_{T_n^v(\bar{\omega})}^{(n)}(\bar{\omega}) = T_n(\omega') < T_{n+1}(\omega') = \theta_{T_{n+1}^v(\bar{\omega})}^{(n+1)}(\bar{\omega}) = \theta_{T_{n+1}^v(\bar{\omega})}^{(n)}(\bar{\omega}),$$

where the last equality follows from (4.11). From the monotonicity of  $\theta^{(n)}$ , we get  $T_n^v(\bar{\omega}) < T_{n+1}^v(\bar{\omega})$ , which yields a contradiction.

Reasoning in the same way as for  $T_1^v$ , since  $T_{n+1}^v$  is the d ebut of  $\bar{E}_{T_{n+1}^v} := \{(\bar{\omega}, t) \in \bar{\Omega} \times \mathbb{R}_+ : \theta_t^{(n+1)}(\bar{\omega}) = T_{n+1}(\omega')\}$ , it is an  $\mathbb{F}^{W, \mu_\infty}$ -stopping time. In particular,  $T_{n+1}^v$  is  $\mathcal{F}$ -measurable, so that there exists a  $\mathbb{P}$ -null set  $N_{T_{n+1}^v} \in \mathcal{F}$  such that  $T_{n+1}^v(\omega, \cdot)$  is  $\mathcal{F}'$ -measurable, whenever  $\omega \notin N_{T_{n+1}^v}$ .

Now define  $F_b = \mathbb{P}'(A_1 \leq b) = \lambda((-\infty, b]) / \lambda(A)$  as before and

$$F_b^{(n+1)}(\bar{\omega}) := \frac{\int_{-\infty}^b v_{T_{n+1}^v(\bar{\omega})}^{(n)}(\omega, (T_1^v(\bar{\omega}), A_1^v(\bar{\omega})), \dots, (T_n^v(\bar{\omega}), A_n^v(\bar{\omega})), a) \lambda(da)}{\int_{-\infty}^{+\infty} v_{T_{n+1}^v(\bar{\omega})}^{(n)}(\omega, (T_1^v(\bar{\omega}), A_1(\bar{\omega})), \dots, (T_n^v(\bar{\omega}), A_n^v(\bar{\omega})), a) \lambda(da)}.$$

Since  $\inf v^{(n)} > 0$  and  $\lambda(\{a\}) = 0$  for any  $a \in A$ , we see that, for every  $\bar{\omega} \in \bar{\Omega}$ , the map  $b \mapsto F_b^{(n+1)}(\bar{\omega})$  is continuous, strictly increasing, valued in  $(0, 1)$ , and  $\lim_{b \rightarrow -\infty} F_b^{(n+1)}(\bar{\omega}) = 0, \lim_{b \rightarrow +\infty} F_b^{(n+1)}(\bar{\omega}) = 1$ . Then, proceeding along the same lines as for the construction of  $A_1^v$ , we see that there exists a unique  $\mathcal{F}_{T_{n+1}^v}^{W, \mu_\infty}$ -measurable map  $A_{n+1}^v : \bar{\Omega} \rightarrow \mathbb{R}$  such that

$$F_{A_{n+1}^v(\bar{\omega})}^{(n+1)}(\bar{\omega}) = F_{A_{n+1}(\omega')}.$$

In particular,  $A_{n+1}^v$  is  $\bar{\mathcal{F}}$ -measurable, therefore, there exists a  $\mathbb{P}$ -null set  $N_{A_{n+1}^v} \in \mathcal{F}$  such that  $A_{n+1}^v(\omega, \cdot)$  is  $\mathcal{F}'$ -measurable, whenever  $\omega \notin N_{A_{n+1}^v}$ .

In order to conclude the proof of the inductive step, let us prove that (4.9) holds for  $n + 1$ , whenever  $\omega \notin N_{n+1} := N_n \cup N_{T_{n+1}^v} \cup N_{A_{n+1}^v}$ . Set  $S_{n+1} = T_{n+1} - T_n$  and recall that, for every  $\omega \notin \tilde{N}$ , the  $\mathbb{F}^\mu$ -compensator of  $\mu$  under  $\mathbb{P}^{v^\omega}$  is given by the right-hand side of (4.8), so that in particular we have,  $\mathbb{P}'$ -a.s.,

$$\begin{aligned} & \mathbb{P}^{v^\omega}(S_{n+1} > t \mid \sigma(T_1, A_1, \dots, T_n, A_n)) \\ (4.12) \quad &= \exp\left(-\int_{T_n}^{T_n+t} \int_A v_s^{(n)}(\omega, (T_1, A_1), \dots, (T_n, A_n), a) \lambda(da) ds\right). \end{aligned}$$

Define  $S_{n+1}^v := T_{n+1}^v - T_n^v$  and observe that, whenever  $\omega \notin N_n \cup N_{T_{n+1}^v}$ ,  $\mathbb{P}'$ -a.s.,

$$\begin{aligned} & \mathbb{P}'(S_{n+1}^v(\omega, \cdot) > t \mid \sigma(T_1^v(\omega, \cdot), A_1^v(\omega, \cdot), \dots, T_n^v(\omega, \cdot), A_n^v(\omega, \cdot))) \\ &= \mathbb{P}'(\theta_{T_{n+1}^v(\omega, \cdot)+t}^{(n+1)}(\omega, \cdot) > \theta_{T_n^v(\omega, \cdot)+t}^{(n+1)}(\omega, \cdot) \mid \\ & \quad \sigma(T_1^v(\omega, \cdot), A_1^v(\omega, \cdot), \dots, T_n^v(\omega, \cdot), A_n^v(\omega, \cdot))) \\ &= \mathbb{P}'\left(T_{n+1}^v > \frac{1}{\lambda(A)} \int_{T_n^v(\omega, \cdot)}^{T_n^v(\omega, \cdot)+t} \int_A v_s^{(n)}(\omega, T_1^v(\omega, \cdot), A_1^v(\omega, \cdot), \dots, \right. \\ & \quad \left. T_n^v(\omega, \cdot), A_n^v(\omega, \cdot), a) \lambda(da) ds \right. \\ & \quad \left. + \theta_{T_n^v(\omega, \cdot)}^{(n)}(\omega, \cdot) \mid \sigma(T_1^v(\omega, \cdot), A_1^v(\omega, \cdot), \dots, T_n^v(\omega, \cdot), A_n^v(\omega, \cdot))\right) \\ &= \mathbb{P}'\left(S_{n+1}^v > \frac{1}{\lambda(A)} \int_{T_n^v(\omega, \cdot)}^{T_n^v(\omega, \cdot)+t} \int_A v_s^{(n)}(\omega, \dots, a) \lambda(da) ds \mid \right. \\ & \quad \left. \sigma(T_1^v(\omega, \cdot), \dots, A_n^v(\omega, \cdot))\right). \end{aligned}$$

Recall now that  $S_{n+1}$  is independent of  $(T_1, A_1, \dots, T_n, A_n)$  under  $\mathbb{P}'$ , and note that, by construction,  $(T_1^v(\omega, \cdot), A_1^v(\omega, \cdot), \dots, T_n^v(\omega, \cdot), A_n^v(\omega, \cdot))$  is  $\sigma(T_1, A_1, \dots, T_n, A_n)$ -measurable. Therefore,  $S_{n+1}$  is also independent of  $(T_1^v(\omega, \cdot), A_1^v(\omega, \cdot), \dots, T_n^v(\omega, \cdot), A_n^v(\omega, \cdot))$ . As a consequence, for every  $\omega \notin N_n \cup N_{T_{n+1}^v}$ ,  $\mathbb{P}'$ -a.s.,

$$\begin{aligned} & \mathbb{P}'\left(S_{n+1} > \frac{1}{\lambda(A)} \int_{T_n^v(\omega, \cdot)}^{T_n^v(\omega, \cdot)+t} \int_A v_s^{(n)}(\omega, \dots, a) \lambda(da) ds \mid \right. \\ & \quad \left. \sigma(T_1^v(\omega, \cdot), \dots, A_n^v(\omega, \cdot))\right) \\ &= \mathbb{P}'(S_{n+1} > r) \Big|_{r=\frac{1}{\lambda(A)} \int_{T_n^v(\omega, \cdot)}^{T_n^v(\omega, \cdot)+t} \int_A v_s^{(n)}(\omega, \dots, a) \lambda(da) ds} \\ &= \exp\left(- \int_{T_n^v(\omega, \cdot)}^{T_n^v(\omega, \cdot)+t} \int_A v_s^{(n)}(\omega, (T_1^v(\omega, \cdot), A_1^v(\omega, \cdot), \dots, \right. \\ & \quad \left. T_n^v(\omega, \cdot), A_n^v(\omega, \cdot)), a) \lambda(da) ds\right), \end{aligned}$$

where for the last equality we used the formula  $\mathbb{P}'(S_{n+1} > r) = \exp(-\lambda(A)r)$ . Comparing with (4.12), we see that the conditional distribution of  $S_{n+1}^v$  given  $T_1^v(\omega, \cdot), A_1^v(\omega, \cdot), \dots, T_n^v(\omega, \cdot), A_n^v(\omega, \cdot)$  under  $\mathbb{P}'$  is the same as the conditional distribution of  $S_{n+1}$  given  $T_1, A_1, \dots, T_n, A_n$  under  $\mathbb{P}^{\nu^\omega}$ . Together with the inductive assumption (4.9) this proves that

$$(4.13) \quad \begin{aligned} & \mathcal{L}_{\mathbb{P}'}(T_1^v(\omega, \cdot), \dots, T_n^v(\omega, \cdot), A_n^v(\omega, \cdot), T_{n+1}^v(\omega, \cdot)) \\ &= \mathcal{L}_{\mathbb{P}^{\nu^\omega}}(T_1, \dots, T_n, A_n, T_{n+1}), \end{aligned}$$

for every  $\omega \notin N_n \cup N_{T_{n+1}^v}$ . Now, recall that, for every  $\omega \notin \tilde{N}$ ,  $\mathbb{P}'$ -a.s.,

$$(4.14) \quad \begin{aligned} & \mathbb{P}^{\nu^\omega}(A_{n+1} \leq b \mid \sigma(T_1, \dots, A_n, T_{n+1})) \\ &= \frac{\int_{-\infty}^b v_{T_{n+1}}^{(n)}(\omega, (T_1, A_1), \dots, (T_n, A_n), a) \lambda(da)}{\int_{-\infty}^{+\infty} v_{T_{n+1}}^{(n)}(\omega, (T_1, A_1), \dots, (T_n, A_n), a) \lambda(da)}. \end{aligned}$$

On the other hand, for every  $\omega \notin N_{n+1}$ , we have,  $\mathbb{P}'$ -a.s.,

$$\begin{aligned} & \mathbb{P}'(A_{n+1}^v(\omega, \cdot) \leq b \mid \sigma(T_1^v(\omega, \cdot), \dots, A_n^v(\omega, \cdot), T_{n+1}^v(\omega, \cdot))) \\ &= \mathbb{P}'(F_{A_{n+1}^v(\omega, \cdot)}^{(n+1)}(\omega, \cdot) \leq F_b^{(n+1)}(\omega, \cdot) \mid \\ & \quad \sigma(T_1^v(\omega, \cdot), \dots, A_n^v(\omega, \cdot), T_{n+1}^v(\omega, \cdot))) \\ &= \mathbb{P}'(F_{A_{n+1}} \leq F_b^{(n+1)}(\omega, \cdot) \mid \sigma(T_1^v(\omega, \cdot), \dots, A_n^v(\omega, \cdot), T_{n+1}^v(\omega, \cdot))). \end{aligned}$$

Since  $A_{n+1}$  is independent of  $(T_1, \dots, A_n, T_{n+1})$  under  $\mathbb{P}'$  and  $(T_1^v(\omega, \cdot), \dots, A_n^v(\omega, \cdot), T_{n+1}^v(\omega, \cdot))$  is  $\sigma(T_1, \dots, A_n, T_{n+1})$ -measurable, it follows that  $A_{n+1}$



is also independent of  $(T_1^v(\omega, \cdot), \dots, A_n^v(\omega, \cdot), T_{n+1}^v(\omega, \cdot))$ . Moreover, by definition we see that  $F_b^{(n+1)}(\omega, \cdot)$  is  $\sigma(T_1^v(\omega, \cdot), \dots, A_n^v(\omega, \cdot), T_{n+1}^v(\omega, \cdot))$ -measurable. Therefore, for every  $\omega \notin N_{n+1}$ , we have  $\mathbb{P}'$ -a.s.,

$$\begin{aligned} &\mathbb{P}'(F_{A_{n+1}} \leq F_b^{(n+1)}(\omega, \cdot) \mid \sigma(T_1^v(\omega, \cdot), \dots, A_n^v(\omega, \cdot), T_{n+1}^v(\omega, \cdot))) \\ &= \mathbb{P}'(F_{A_{n+1}} \leq a) \Big|_{a=F_b^{(n+1)}(\omega, \cdot)} = F_b^{(n+1)}(\omega, \cdot), \end{aligned}$$

where we used the fact that  $F_{A_{n+1}}$  is uniformly distributed in  $(0, 1)$  under  $\mathbb{P}'$ . Comparing with (4.14), we see that the conditional distribution of  $A_{n+1}^v$  given  $T_1^v(\omega, \cdot), A_1^v(\omega, \cdot), \dots, T_n^v(\omega, \cdot), A_n^v(\omega, \cdot), T_{n+1}^v(\omega, \cdot)$  under  $\mathbb{P}'$  is the same as the conditional distribution of  $A_{n+1}$  given  $T_1, A_1, \dots, T_n, A_n, T_{n+1}$  under  $\mathbb{P}^{\omega}$ . Therefore, by (4.13) we deduce that (4.9) holds for  $n + 1$ , whenever  $\omega \notin N_{n+1}$ , which concludes the proof of the inductive step and also the proof of Case I.

*Case II:*  $A_c \neq \emptyset$ . Let  $\pi : \mathbb{R} \rightarrow A$  be the canonical projection (4.5) or (4.6) according whether  $A_{nc} = \emptyset$  or  $A_{nc} \neq \emptyset$ . The idea of the proof is to construct a random sequence with values in  $(0, \infty) \times \mathbb{R}$  using the Case I previously addressed, and obtain the required sequence  $(T_n^v, A_n^v)_{n \geq 1}$  by projecting the second component onto  $A$ . The detailed construction and proof is presented below in the case  $A_{nc} = \emptyset$ , the other one being simpler and entirely analogous.

Given  $\nu \in \mathcal{V}_{\text{inf} > 0}$ , define  $\bar{\nu} = \bar{\nu}_t(\omega, \omega', r) : \bar{\Omega} \times \mathbb{R}_+ \times \mathbb{R} \rightarrow (0, \infty)$  by

$$\bar{\nu}_t(\omega, \omega', r) := \nu_t(\omega, \omega', \pi(r)).$$

By a monotone class argument, we see that  $\bar{\nu}$  is  $\mathcal{P}(\mathbb{F}^{W, \mu'}) \otimes \mathcal{B}(\mathbb{R})$ -measurable. Then we can perform the construction presented in step I, with  $\mathbb{R}, \lambda', \bar{\nu}, R_n, \mu', \mathbb{F}^{W, \mu'_\infty} = (\mathcal{F}_t^{W, \mu'_\infty})_{t \geq 0}$  instead of  $A, \lambda, \nu, A_n, \mu, \mathbb{F}^{W, \mu_\infty} = (\mathcal{F}_t^{W, \mu_\infty})_{t \geq 0}$ , respectively. This way we obtain a  $\mathbb{P}$ -null set  $N \in \mathcal{F}$  and a sequence  $(\bar{T}_n^{\bar{\nu}}, \bar{R}_n^{\bar{\nu}})_{n \geq 1}$  with values in  $(0, \infty) \times \mathbb{R}$  such that  $\bar{T}_n^{\bar{\nu}} < \bar{T}_{n+1}^{\bar{\nu}} \nearrow \infty$ ,  $\bar{T}_n^{\bar{\nu}}$  is an  $\mathbb{F}^{W, \bar{\mu}_\infty}$ -stopping time and  $\bar{R}_n^{\bar{\nu}}$  is  $\mathcal{F}_{\bar{T}_n^{\bar{\nu}}}^{W, \bar{\mu}_\infty}$ -measurable, and

$$(4.15) \quad \mathcal{L}_{\mathbb{P}'}((\bar{T}_n^{\bar{\nu}}(\omega, \cdot), \bar{R}_n^{\bar{\nu}}(\omega, \cdot))_{n \geq 1}) = \mathcal{L}_{\mathbb{P}^{\bar{\nu}\omega}}((T_n, R_n)_{n \geq 1})$$

for every  $\omega \notin N$ . We define the required sequence  $(T_n^v, A_n^v)_{n \geq 1}$  setting

$$T_n^v := \bar{T}_n^{\bar{\nu}}, \quad A_n^v := \pi(\bar{R}_n^{\bar{\nu}}).$$

Clearly, conditions (i)–(ii)–(iii) of the theorem hold true and, recalling the notation  $A_n = \pi(R_n)$ , from (4.15) it follows that

$$(4.16) \quad \mathcal{L}_{\mathbb{P}'}((T_n^v(\omega, \cdot), A_n^v(\omega, \cdot))_{n \geq 1}) = \mathcal{L}_{\mathbb{P}^{\bar{\nu}\omega}}((T_n, A_n)_{n \geq 1}).$$

Next, note that, by the definition of  $\bar{v}$ ,

$$\begin{aligned} \int_{\mathbb{R}} (1 - \bar{v}_t(r))\lambda'(dr) &= \int_{(0,\infty)} (1 - \bar{v}_t(r))\lambda'(dr) + \sum_{j \in A_c} \int_{I_j} (1 - \bar{v}_t(r))\lambda'(dr) \\ &= \int_{A_{nc}} (1 - v_t(s))\lambda(da) + \sum_{j \in A_c} (1 - v_t(j))\lambda(\{j\}) \\ &= \int_A (1 - v_s(a))\lambda(da) \end{aligned}$$

and  $\bar{v}_{T_n}(R_n) = v_{T_n}(\pi(R_n)) = v_{T_n}(A_n)$ , so that

$$\begin{aligned} \kappa_t^{\bar{v}} &= \exp\left(\int_0^t \int_{\mathbb{R}} (1 - \bar{v}_s(r))\lambda'(dr) ds\right) \prod_{T_n \leq t} \bar{v}_{T_n}(R_n) \\ &= \exp\left(\int_0^t \int_A (1 - v_s(a))\lambda(da) ds\right) \prod_{T_n \leq t} v_{T_n}(A_n) \\ &= \kappa_t^v. \end{aligned}$$

Therefore, we have for  $\omega \notin N$  on every  $\sigma$ -algebra  $\mathcal{F}_t^\mu$ ,

$$\mathbb{P}^{\bar{v}^\omega}(d\omega') = \kappa_t^{\bar{v}}(\omega, \omega')\mathbb{P}'(d\omega') = \kappa_t^v(\omega, \omega')\mathbb{P}'(d\omega') = \mathbb{P}^{v^\omega}(d\omega')$$

and property (iv) of the theorem follows from (4.16). The proof is completed.  $\square$

We can now prove the main result of this subsection and conclude the proof of the inequality  $\mathcal{U}_0^{\mathcal{R}} \leq \mathcal{U}_0$ .

PROPOSITION 4.2. *For every  $v \in \mathcal{V}_{\text{inf}>0}$ , there exists  $\bar{\alpha}^v \in \mathcal{A}^{W,\mu'}$  such that*

$$(4.17) \quad \mathcal{L}_{\mathbb{P}^v}(x_0, B, I) = \mathcal{L}_{\bar{\mathbb{P}}}^v(x_0, B, \bar{\alpha}^v).$$

*In particular,  $V$  and  $W$  are standard Wiener processes,  $V, W, x_0$  are all independent under  $\mathbb{P}^v$ , and we have*

$$(4.18) \quad \mathcal{L}_{\mathbb{P}^v}(X, I) = \mathcal{L}_{\bar{\mathbb{P}}}^v(X^{\bar{\alpha}^v}, \bar{\alpha}^v), \quad J^{\mathcal{R}}(v) = \bar{J}(\bar{\alpha}^v).$$

*Finally,  $\mathcal{U}_0^{\mathcal{R}} := \sup_{v \in \mathcal{V}} J^{\mathcal{R}}(v) \leq \sup_{\bar{\alpha} \in \mathcal{A}^{W,\mu'}} \bar{J}(\bar{\alpha}) = \sup_{\alpha \in \mathcal{A}^W} J(\alpha) =: \mathcal{U}_0$ .*

PROOF. Suppose that (4.17) holds. Then, from equation (2.5) and Assumption (A1) it is well-known that the first equality in (4.18) holds as well, and this implies the second equality. From the arbitrariness of  $v \in \mathcal{V}_{\text{inf}>0}$ , we deduce that  $\sup_{v \in \mathcal{V}_{\text{inf}>0}} J^{\mathcal{R}}(v) \leq \sup_{\bar{\alpha} \in \mathcal{A}^{W,\mu'}} \bar{J}(\bar{\alpha})$ . Since  $\sup_{v \in \mathcal{V}_{\text{inf}>0}} J^{\mathcal{R}}(v) = \sup_{v \in \mathcal{V}} J^{\mathcal{R}}(v)$  by (3.9), we conclude that  $\sup_{v \in \mathcal{V}} J^{\mathcal{R}}(v) \leq \sup_{\bar{\alpha} \in \mathcal{A}^{W,\mu'}} \bar{J}(\bar{\alpha}) = \sup_{\alpha \in \mathcal{A}^W} J(\alpha)$ , where the last equality follows from Lemma 4.1.

Let us now prove (4.17). Fix  $\nu \in \mathcal{V}_{\text{inf}>0}$  and consider the process  $\bar{\alpha}^\nu$  given by Lemma 4.3. In order to prove (4.17), we have to show that

$$(4.19) \quad \mathbb{E}[\kappa_T^\nu \chi(x_0) \psi(B) \phi(I)] = \mathbb{E}[\chi(x_0) \psi(B) \phi(\bar{\alpha}^\nu)],$$

for any  $\chi \in B_b(\mathbb{R}^n)$  (the space of bounded Borel measurable real functions on  $\mathbb{R}^n$ ), for any  $\psi \in B_b(\mathbf{C}_{m+d})$  (the space of bounded Borel measurable real functions on  $\mathbf{C}_{m+d}$ , which denotes the space of continuous paths from  $[0, T]$  to  $\mathbb{R}^{m+d}$  endowed with the supremum norm) and any  $\phi \in B_b(\mathbf{D}_A)$  (the space of bounded Borel measurable real functions on  $\mathbf{D}_A$ , which denotes the space of càdlàg paths from  $[0, T]$  to  $A$  endowed with the supremum norm). By the Fubini theorem, (4.19) can be rewritten as

$$\begin{aligned} & \int_{\Omega} \chi(x_0(\omega)) \psi(B(\omega)) \left( \int_{\Omega'} \kappa_T^\nu(\omega, \omega') \phi(I(\omega')) \mathbb{P}'(d\omega') \right) \mathbb{P}(d\omega) \\ &= \int_{\Omega} \chi(x_0(\omega)) \psi(B(\omega)) \left( \int_{\Omega'} \phi(\bar{\alpha}^\nu(\omega, \omega')) \mathbb{P}'(d\omega') \right) \mathbb{P}(d\omega). \end{aligned}$$

Let  $\tilde{N} \in \mathcal{F}$  be as in Lemma 4.2. Then we have to prove that  $\mathbb{E}'[\kappa_T^\nu(\omega, \cdot) \phi(I)] = \mathbb{E}'[\phi(\bar{\alpha}^\nu(\omega, \cdot))]$ , whenever  $\omega \notin \tilde{N}$ , or, equivalently by definition of  $\mathbb{P}^{\nu\omega}$ :

$$\mathbb{E}^{\nu\omega}[\phi(I)] = \mathbb{E}'[\phi(\bar{\alpha}^\nu(\omega, \cdot))] \quad \text{whenever } \omega \notin \tilde{N}.$$

This is a direct consequence of the last statement of Lemma 4.3.  $\square$

4.2. *Proof of the inequality  $\mathcal{U}_0 \leq \mathcal{U}_0^{\mathcal{R}}$ .* In this proof, we borrow some constructions from Proposition 4.1 in [15], but we need to obtain improved results and we simplify considerably some arguments.

Suppose we are given a setting  $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F}, V, W, x_0)$  for the optimal control problem with partial observation, satisfying the conditions in Section 2.2, and consider the controlled equation (2.5) and the gain (2.6). We fix an  $\mathbb{F}^W$ -progressive process  $\alpha$  with values in  $A$ . We will show how to construct a sequence of settings  $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}_k, \hat{V}, \hat{W}, \hat{\mu}_k, \hat{x}_0)_k$  for the randomized control problem of Section 3.1, and a sequence  $(\hat{\nu}^k)_k$  of corresponding admissible controls (both sequences depending on  $\alpha$ ), such that for the corresponding gains, defined by (3.7), we have

$$(4.20) \quad J^{\mathcal{R}}(\hat{\nu}^k) \rightarrow J(\alpha) \quad \text{as } k \rightarrow \infty.$$

Admitting for a moment that this has been done, the proof is easily concluded by the following arguments. By (4.20), we can find, for any  $\varepsilon > 0$ , some  $k$  such that  $J^{\mathcal{R}}(\hat{\nu}^k) > J(\alpha) - \varepsilon$ . Since  $J^{\mathcal{R}}(\hat{\nu}^k)$  is a gain of a randomized control problem, it can not exceed the value  $\mathcal{U}_0^{\mathcal{R}}$  defined in (3.8) which, by Proposition 3.1, does not depend on  $\varepsilon$  nor on  $\alpha$ . It follows that

$$\mathcal{U}_0^{\mathcal{R}} > J(\alpha) - \varepsilon$$

and by the arbitrariness of  $\varepsilon$  and  $\alpha$ , we obtain the required inequality  $\mathcal{U}_0^{\mathcal{R}} \geq \mathcal{U}_0$ .

In order to construct the sequences  $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}_k, \hat{V}, \hat{W}, \hat{\mu}_k, \hat{x}_0)_k$  and  $(\hat{v}^k)_k$  satisfying (4.20), we apply Proposition 4.1 above. Then, for any integer  $k \geq 1$ , there exists a marked point process  $(\hat{S}_n^k, \hat{\eta}_n^k)_{n \geq 1}$  defined in  $(\hat{\Omega}, \hat{\mathcal{F}}, \mathbb{Q})$  satisfying the following conditions:

1. Setting  $\hat{S}_0^k = 0, \hat{\eta}_0^k = a_0, \hat{I}_t^k = \sum_{n \geq 0} \hat{\eta}_n^k 1_{[\hat{S}_n^k, \hat{S}_{n+1}^k)}(t)$ , we have  $\tilde{\rho}(\hat{I}^k, \hat{\alpha}) < 1/k$ .

2. Denote  $\hat{\mu}_k = \sum_{n \geq 1} \delta_{(\hat{S}_n^k, \hat{\eta}_n^k)}$  the random measure associated to  $(\hat{S}_n^k, \hat{\eta}_n^k)_{n \geq 1}$  and  $\mathbb{F}^{\hat{\mu}_k} = (\mathcal{F}_t^{\hat{\mu}_k})_{t \geq 0}$  the natural filtration of  $\hat{\mu}_k$ ; then the compensator of  $\hat{\mu}_k$  under  $\mathbb{Q}$  with respect to the filtration  $(\mathcal{F}_t^{\hat{W}} \vee \mathcal{F}_t^{\hat{\mu}_k})_{t \geq 0}$  is absolutely continuous with respect to  $\lambda(da) dt$  and it can be written in the form

$$\hat{v}_t^k(\hat{\omega}, a) \lambda(da) dt$$

for some nonnegative  $\mathcal{P}(\mathbb{F}^{\hat{W}, \hat{\mu}_k}) \otimes \mathcal{B}(A)$ -measurable function  $\hat{v}^k$  satisfying

$$\inf_{\hat{\Omega} \times [0, T] \times A} \hat{v}^k > 0 \quad \text{and} \quad \sup_{\hat{\Omega} \times [0, T] \times A} \hat{v}^k < \infty.$$

We note that  $\hat{\mu}_k$  (and so also  $\hat{I}^k$  and  $\hat{v}^k$ ) depend on  $\alpha$  as well, but we do not make it explicit in the notation. Let us now consider the completion of the probability space  $(\hat{\Omega}, \hat{\mathcal{F}}, \mathbb{Q})$ , that will be denoted by the same symbol for simplicity of notation, and let  $\mathcal{N}$  denote the family of  $\mathbb{Q}$ -null sets of the completion. Then the filtration  $(\mathcal{F}_t^{\hat{W}} \vee \mathcal{F}_t^{\hat{\mu}_k} \vee \mathcal{N})_{t \geq 0}$  coincides with the filtration previously denoted by  $\mathbb{F}^{\hat{W}, \hat{\mu}_k} = (\mathcal{F}_t^{\hat{W}, \hat{\mu}_k})_{t \geq 0}$  [compare with formula (3.3) in Section 3.1]. It is easy to see that  $\hat{v}_t^k(\hat{\omega}, a) \lambda(da) dt$  is the compensator of  $\hat{\mu}_k$  with respect to  $\mathbb{F}^{\hat{W}, \hat{\mu}_k}$  and the extended probability  $\mathbb{Q}$  as well.

Using the Girsanov theorem for point processes (see, e.g., [19]), we next construct an equivalent probability under which  $\hat{\mu}_k$  becomes a Poisson random measure with intensity  $\lambda$ . Since  $\hat{v}^k$  is a strictly positive  $\mathcal{P}(\mathbb{F}^{\hat{W}, \hat{\mu}_k}) \otimes \mathcal{B}(A)$ -measurable random field with bounded inverse, the Doléans exponential process

$$\begin{aligned} (4.21) \quad M_t^k &:= \exp\left(\int_0^t \int_A (1 - \hat{v}_s^k(a)^{-1}) \hat{v}_s^k(a) \lambda(da) ds\right) \\ &\times \prod_{\hat{S}_n^k \leq t} \hat{v}_{\hat{S}_n^k}^k(\hat{\eta}_n^k)^{-1}, \quad t \in [0, T], \end{aligned}$$

is a strictly positive martingale (with respect to  $\mathbb{F}^{\hat{W}, \hat{\mu}_k}$  and  $\mathbb{Q}$ ), and we can define an equivalent probability  $\hat{\mathbb{P}}_k$  on the space  $(\hat{\Omega}, \hat{\mathcal{F}})$  setting  $\hat{\mathbb{P}}_k(d\hat{\omega}) = M_T^k(\hat{\omega}) \mathbb{Q}(d\hat{\omega})$ . The expectation under  $\hat{\mathbb{P}}_k$  will be denoted  $\hat{\mathbb{E}}_k$ . By the Girsanov theorem, the restriction of  $\hat{\mu}_k$  to  $(0, T] \times A$  has  $(\hat{\mathbb{P}}_k, \mathbb{F}^{\hat{W}, \hat{\mu}_k})$ -compensator  $\lambda(da) dt$ , so that in particular it is a Poisson random measure. It can also be proved by standard arguments (see, e.g., [15], p. 2155, for detailed verifications in a similar framework)

that  $\hat{B}$  is a  $(\hat{\mathbb{P}}_k, \mathbb{F}^{\hat{W}, \hat{\mu}_k})$ -Wiener process and that  $\hat{B}$  and  $\hat{\mu}_k$  are independent under  $\hat{\mathbb{P}}_k$ . We have thus constructed a setting  $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}_k, \hat{V}, \hat{W}, \hat{\mu}_k, \hat{x}_0)$  for a randomized control problem as in Section 3.1. Since  $\hat{v}^k$  is a bounded, strictly positive and  $\mathcal{P}(\mathbb{F}^{\hat{W}, \hat{\mu}}) \otimes \mathcal{B}(A)$ -measurable random field it belongs to the class  $\hat{\mathcal{V}}^k$  of admissible controls for the randomized control problem and we now proceed to evaluating its gain  $J^{\mathcal{R}}(\hat{v}^k)$  and to comparing it with  $J(\alpha)$ . Our aim is to show that, as a consequence of the fact that  $\tilde{\rho}(\hat{I}^k, \hat{\alpha}) < 1/k$ , we have  $J^{\mathcal{R}}(\hat{v}^k) \rightarrow J(\alpha)$  as  $k \rightarrow \infty$ .

We introduce the Doléans exponential process  $\kappa^{\hat{v}^k}$  corresponding to  $\hat{v}^k$  [compare formula (3.5)]:

$$(4.22) \quad \kappa_t^{\hat{v}^k} = \exp\left(\int_0^t \int_A (1 - \hat{v}_s^k(a)) \lambda(da) ds\right) \prod_{\hat{S}_n^k \leq t} v_{\hat{S}_n^k}^k(\hat{\eta}_n^k), \quad t \in [0, T],$$

we define the probability  $d\hat{\mathbb{P}}_k^{\hat{v}^k} = \kappa_T^{\hat{v}^k} d\hat{\mathbb{P}}_k$  and we obtain the gain

$$J^{\mathcal{R}}(\hat{v}) = \hat{\mathbb{E}}^{\hat{v}^k} \left[ \int_0^T f_t(\hat{X}^k, \hat{I}^k) dt + g(\hat{X}^k) \right],$$

where  $\hat{X}^k$  is the solution to the equation

$$(4.23) \quad d\hat{X}_t^k = b_t(\hat{X}^k, \hat{I}^k) dt + \sigma_t(\hat{X}^k, \hat{I}^k) d\hat{B}_t, \quad \hat{X}_0^k = \hat{x}_0.$$

However, comparing (4.21) and (4.22) shows that  $\kappa_T^{\hat{v}^k} M_T^k \equiv 1$ , so that the Girsanov transformation  $\hat{\mathbb{P}}_k \mapsto \hat{\mathbb{P}}_k^{\hat{v}^k}$  is the inverse to the transformation  $\mathbb{Q} \mapsto \hat{\mathbb{P}}_k$  made above, and changes back the probability  $\hat{\mathbb{P}}_k$  into  $\mathbb{Q}$  considered above. Therefore, we have  $\hat{\mathbb{P}}_k^{\hat{v}^k} = \mathbb{Q}$  and also

$$(4.24) \quad J^{\mathcal{R}}(\hat{v}^k) = \mathbb{E}^{\mathbb{Q}} \left[ \int_0^T f_t(\hat{X}^k, \hat{I}^k) dt + g(\hat{X}^k) \right].$$

On the other hand, the gain  $J(\alpha)$  of the initial control problem with partial observation was defined in (2.6) in terms of the solution  $X^\alpha$  to the controlled equation (2.5). Denoting  $\hat{X}^\alpha$  the extension of  $X^\alpha$  to  $\hat{\Omega}$ , it is easy to verify that it is the solution to

$$(4.25) \quad d\hat{X}_t^\alpha = b_t(\hat{X}^\alpha, \hat{\alpha}) dt + \sigma_t(\hat{X}^\alpha, \hat{\alpha}) d\hat{B}_t, \quad \hat{X}_0^\alpha = \hat{x}_0,$$

and that

$$(4.26) \quad J(\alpha) = \mathbb{E}^{\mathbb{Q}} \left[ \int_0^T f_t(\hat{X}^\alpha, \hat{\alpha}) dt + g(\hat{X}^\alpha) \right].$$

Equations (4.25) and (4.23) are considered in the same probability space  $(\hat{\Omega}, \hat{\mathcal{F}}, \mathbb{Q})$ . In (4.23), we find a solution adapted to the filtration  $\mathbb{G}^k := \mathbb{F}^{\hat{x}_0, \hat{B}, \hat{\mu}_k}$  [defined as in (3.3)] and in (4.25) we find a solution adapted to the filtration  $\mathbb{G}^0 := \mathbb{F}^{\hat{x}_0, \hat{B}}$

generated by  $\hat{x}_0$  and  $\hat{B}$  (since  $\alpha$  was  $\mathbb{F}^W$ -progressive and so  $\hat{\alpha}$  is progressive with respect to  $\mathbb{F}^{\hat{W}} \subset \mathbb{F}^{\hat{x}_0, \hat{B}}$ ).

In order to conclude, we need the following stability lemma, where the continuity condition **(A1)**(iii) plays a role.

LEMMA 4.4. *Given a probability space  $(\hat{\Omega}, \hat{\mathcal{F}}, \mathbb{Q})$  with filtrations  $\mathbb{G}^k = (\mathcal{G}_t^k)_{t \geq 0}$  ( $k \geq 0$ ) consider the equations*

$$dY_t^k = b_t(Y^k, \gamma^k) dt + \sigma_t(Y^k, \gamma^k) d\beta_t, \quad \hat{Y}_0^k = y_0,$$

where  $\beta$  is a Wiener process with respect to each  $\mathbb{G}^k$ ,  $\mathbb{E}^{\mathbb{Q}}|y_0|^p < \infty$ ,  $y_0$  is  $\mathcal{G}_t^k$ -measurable and  $\gamma^k$  is  $\mathbb{G}^k$ -progressive for every  $k$ . If  $\tilde{\rho}(\gamma^k, \gamma^0) \rightarrow 0$  as  $k \rightarrow \infty$ , then

$$\mathbb{E}^{\mathbb{Q}} \sup_{t \in [0, T]} |Y_t^k - Y_t^0|^p \rightarrow 0,$$

$$\mathbb{E}^{\mathbb{Q}} \left[ \int_0^T f_t(Y^k, \gamma^k) dt + g(Y^k) \right] \rightarrow \mathbb{E}^{\mathbb{Q}} \left[ \int_0^T f_t(Y^0, \gamma^0) dt + g(Y^0) \right].$$

PROOF. This stability result for control problems was first proved in [25] in the standard diffusion case. The extension to the non-Markovian case presented in [15], Lemma 4.1 and Remark 4.1, also holds in our case with the same proof, using the continuity assumption **(A1)**(iii) and the Lipschitz and growth conditions (2.2)–(2.4). □

Applying the lemma to  $\beta = \hat{B}$ ,  $Y^k = \hat{X}^k$ ,  $\gamma^k = \hat{I}^k$  (for  $k \geq 1$ ) and  $Y^0 = \hat{X}^\alpha$ ,  $\gamma^0 = \hat{\alpha}$ , and recalling that  $\tilde{\rho}(\hat{I}^k, \hat{\alpha}) < 1/k \rightarrow 0$  we conclude by (4.24), (4.26) that  $J^{\mathcal{R}}(\hat{v}^k) \rightarrow J(\alpha)$  as  $k \rightarrow \infty$ . Therefore, relation (4.20) is satisfied for this choice of the sequence  $(\hat{v}^k)_k$  and for the corresponding settings  $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}_k, \hat{V}, \hat{W}, \hat{\mu}_k, \hat{x}_0)_k$ . This ends the proof of the inequality  $\mathcal{U}_0 \leq \mathcal{U}_0^{\mathcal{R}}$ .

**5. The randomized equation.** In this section, the assumptions **(A1)** and **(A2)** are assumed to hold. We will show how the randomized formulation of the control problem leads to a randomized equation in terms of a backward SDE. We choose a setting for the randomized control problem (3.8) as in Remark 3.2, that is, a product extension  $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}}, V, W, \mu, x_0)$  of the setting  $(\Omega, \mathcal{F}, \mathbb{P}, V, W, x_0)$  for the initial control problem (2.7). In view of Proposition 3.1, entirely analogous results hold true in any setting  $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}, \hat{V}, \hat{W}, \hat{\mu}, \hat{x}_0)$  for the randomized control problem as described in Section 3.1.

5.1. *The separated randomized control problem.* We first consider the (path-dependent) filtering of the randomized process  $X$  solution to (3.2), which consists in the process of conditional distributions  $\rho_t$  of  $X_{\cdot \wedge t}$  given  $\mathcal{F}_t^{W, \mu}$ . More precisely,

let  $\mathcal{P}(\mathbf{C}_n)$  be the space of probability measures on  $\mathbf{C}_n$  and let  $B_b(\mathbf{C}_n)$  denote the space of bounded Borel measurable real functions on  $\mathbf{C}_n$ . We define  $\rho = (\rho_t)_{0 \leq t \leq T}$  as an  $\mathbb{F}^{W,\mu}$ -optional process valued in  $\mathcal{P}(\mathbf{C}_n)$  satisfying, for every  $\varphi \in B_b(\mathbf{C}_n)$ , [we use the notation  $\rho_t(\varphi) = \int_{\mathbf{C}_n} \varphi(x) \rho_t(dx)$ ]

$$(5.1) \quad \rho_t(\varphi) = \bar{\mathbb{E}}[\varphi(X_{\cdot \wedge t}) \mid \mathcal{F}_t^{W,\mu}], \quad t \in [0, T], \bar{\mathbb{P}}\text{-a.s.}$$

The process  $t \mapsto \bar{\mathbb{E}}[\varphi(X_{\cdot \wedge t}) \mid \mathcal{F}_t^{W,\mu}]$  is understood as an optional projection. The existence of such a process  $\rho$  follows, for example, from Theorem 2.24 in [1]. While (5.1) is defined for bounded  $\varphi$ , since  $\rho_t$  is constructed as a  $\mathcal{P}(\mathbf{C}_n)$ -valued process, relation (5.1) holds for unbounded  $\varphi$  once the conditional expectation is well-defined, that is,  $\rho_t(|\varphi|) < \infty$  for all  $t \in [0, T]$ ,  $\bar{\mathbb{P}}$ -a.s. (see, e.g., Remark 2.27 in [1]).

We can now express the randomized gain functional in terms of the  $\mathbb{F}^{W,\mu}$ -optional processes  $\rho$  and  $I$ .

LEMMA 5.1. *For any  $v \in \mathcal{V}$ , we have*

$$J^{\mathcal{R}}(v) = \mathbb{E}^v \left[ \int_0^T \rho_t(f_t(\cdot, I)) dt + \rho_T(g) \right]$$

and, more generally,

$$(5.2) \quad \begin{aligned} &\mathbb{E}^v \left[ \int_t^T f_s(X, I) ds + g(X) \mid \mathcal{F}_t^{W,\mu} \right] \\ &= \mathbb{E}^v \left[ \int_t^T \rho_s(f_s(\cdot, I)) ds + \rho_T(g) \mid \mathcal{F}_t^{W,\mu} \right], \end{aligned}$$

for all  $0 \leq t \leq T$ .

PROOF. The result follows from the Bayes formula and the  $(\bar{\mathbb{P}}, \mathbb{F}^{W,\mu})$ -martingale property of the density process  $\kappa^v$ .  $\square$

The above Lemma 5.1 together with Theorem 3.1 proves that the randomized control problem, and thus the primal control problem under partial observation, can be written in a *separated* form involving  $\mathbb{F}^{W,\mu}$ -optional state processes:

$$(5.3) \quad v_0 = \sup_{v \in \mathcal{V}} \mathbb{E}^v \left[ \int_0^T \rho_t(f_t(\cdot, I)) dt + \rho_T(g) \right].$$

5.2. *BSDE representation.* The purpose of this paragraph is to show that the separated randomized control problem, described by the right-hand side of (5.3), admits a dual representation in terms of a constrained backward SDE, which then characterizes both the primal control problem and the randomized control problem (as well as the separated randomized control problem). We shall refer to it as the *randomized equation*.

On the space  $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$  equipped with the filtration  $\mathbb{F}^{W,\mu}$ , let us consider the following constrained BSDE on the time interval  $[0, T]$ :

$$(5.4) \quad \begin{cases} Y_t = \rho_T(g) + \int_t^T \rho_s(f_s(\cdot, I)) ds + K_T - K_t \\ \quad - \int_t^T Z_s dW_s - \int_t^T \int_A U_s(a) \mu(ds da), \\ U_t(a) \leq 0. \end{cases}$$

We look for a (minimal) solution to (5.4) in the sense of the following definition.

**DEFINITION 5.1.** A quadruple  $(Y_t, Z_t, U_t(a), K_t)$  ( $t \in [0, T], a \in A$ ) is called a solution to the BSDE (5.4) if:

1.  $Y \in \mathcal{S}^2(\mathbb{F}^{W,\mu})$ , the set of real-valued càdlàg  $\mathbb{F}^{W,\mu}$ -adapted processes satisfying  $\|Y\|_{\mathcal{S}^2}^2 := \bar{\mathbb{E}}[\sup_{0 \leq t \leq T} |Y_t|^2] < \infty$ ;
2.  $Z \in L^2_W(\mathbb{F}^{W,\mu})$ , the set of  $\mathbb{F}^{W,\mu}$ -predictable processes with values in  $\mathbb{R}^d$  satisfying  $\|Z\|_{L^2_W}^2 := \bar{\mathbb{E}}[\int_0^T |Z_t|^2 dt] < \infty$ ;
3.  $U \in L^2_{\tilde{\mu}}(\mathbb{F}^{W,\mu})$ , the set of real-valued  $\mathcal{P}(\mathbb{F}^{W,\mu}) \otimes \mathcal{B}(A)$ -measurable processes satisfying  $\|U\|_{L^2_{\tilde{\mu}}}^2 := \bar{\mathbb{E}}[\int_0^T \int_A |U_t(a)|^2 \lambda(da) dt] < \infty$ ;
4.  $K \in \mathcal{K}^2(\mathbb{F}^{W,\mu})$ , the subset of  $\mathcal{S}^2(\mathbb{F}^{W,\mu})$  consisting of  $\mathbb{F}^{W,\mu}$ -predictable non-decreasing process with  $K_0 = 0$ ;
5.  $\bar{\mathbb{P}}$ -a.s. the equality in (5.4) holds for every  $t \in [0, T]$  and the constraint  $U_t(a) \leq 0$  is understood to hold  $\bar{\mathbb{P}}(d\bar{\omega})\lambda(da) dt$ -almost everywhere.

A minimal solution  $(Y, Z, U, K)$  is a solution to (5.4) such that for any other solution  $(Y', Z', U', K')$ , we have  $\bar{\mathbb{P}}$ -a.s.,  $Y_t \leq Y'_t$  for all  $t \in [0, T]$ .

We now state the main result of this section.

**THEOREM 5.1.** *There exists a unique minimal solution  $(Y, Z, U, K) \in \mathcal{S}^2(\mathbb{F}^{W,\mu}) \times L^2_W(\mathbb{F}^{W,\mu}) \times L^2_{\tilde{\mu}}(\mathbb{F}^{W,\mu}) \times \mathcal{K}^2(\mathbb{F}^{W,\mu})$  to the randomized equation (5.4). Moreover, we have  $Y_0 = \sup_{v \in \mathcal{V}} J^{\mathcal{R}}(v)$ , and more generally,*

$$(5.5) \quad Y_t = \operatorname{ess\,sup}_{v \in \mathcal{V}} \mathbb{E}^v \left[ \int_t^T \rho_s(f_s(\cdot, I)) ds + \rho_T(g) \mid \mathcal{F}_t^{W,\mu} \right].$$

**REMARK 5.1.** Combining Theorems 3.1 and 5.1, we deduce the BSDE representation for the primal problem

$$(5.6) \quad Y_0 = \sup_{\alpha \in \mathcal{A}^W} J(\alpha).$$

We refer sometimes to  $Y_0 = \sup_{v \in \mathcal{V}} J^{\mathcal{R}}(v)$  as *duality relation*, since  $Y_0$  coincides with the infimum  $\inf\{Y'_0 : (Y', Z', U', K') \in \mathcal{S}^2(\mathbb{F}^{W,\mu}) \times L^2_W(\mathbb{F}^{W,\mu}) \times L^2_{\tilde{\mu}}(\mathbb{F}^{W,\mu}) \times \mathcal{K}^2(\mathbb{F}^{W,\mu}) \text{ solution to (5.4)}\}$ .



REMARK 5.2. In the Markovian setting studied in our companion paper [4], in addition to the probabilistic representation formula of Theorem 5.1, we are also able to find an optimal control when the value function is regular enough; see the verification Theorem 5.1 in [4].

PROOF OF THEOREM 5.1. Let us introduce for every  $n \in \mathbb{N}$  the following penalized BSDE on  $[0, T]$ :

$$(5.7) \quad \begin{aligned} Y_t^n &= \rho_T(g) + \int_t^T \rho_s(f_s(\cdot, I)) ds + K_T^n - K_t^n \\ &\quad - \int_t^T Z_s^n dW_s - \int_t^T \int_A U_s^n(a) \mu(ds da), \end{aligned}$$

where

$$K_t^n = n \int_0^t \int_A (U_s^n(a))^+ \lambda(da) ds.$$

Set  $\xi := \rho_T(g)$  and  $F_t := \rho_t(f_t(\cdot, I))$ . By (2.4) and (3.6), we see that

$$\mathbb{E}|\xi|^2 < \infty, \quad \mathbb{E}\left[\int_0^T |F_t|^2 dt\right] < \infty$$

[here we use the assumption that  $p \geq 2r$  in **(A1)(v)**]. Then, from Lemma 2.4 in [33], it follows that, for every  $n \in \mathbb{N}$ , there exists a unique solution  $(Y^n, Z^n, U^n) \in \mathcal{S}^2(\mathbb{F}^{W,\mu}) \times L^2_W(\mathbb{F}^{W,\mu}) \times L^2_{\tilde{\mu}}(\mathbb{F}^{W,\mu})$  to the above penalized BSDE.

Now, proceeding along the same lines as in the proof of Lemma 4.8 in [15], we obtain the formula

$$(5.8) \quad Y_t^n = \operatorname{ess\,sup}_{v \in \mathcal{V}_n} \mathbb{E}^v \left[ \int_t^T \rho_s(f_s(\cdot, I)) ds + \rho_T(g) \mid \mathcal{F}_t^{W,\mu} \right],$$

where  $\mathcal{V}_n = \{v \in \mathcal{V} : v \text{ takes values in } (0, n]\}$ . By (5.2), together with estimates (2.4) and (3.6), we deduce that

$$(5.9) \quad \sup_n Y_t^n < \infty \quad \text{for all } 0 \leq t \leq T.$$

Notice that equation (5.4) can be written as follows:

$$(5.10) \quad \begin{cases} Y_t = \rho_T(g) + \int_t^T \left( \rho_s(f_s(\cdot, I)) - \int_A U_s(a) \lambda(da) \right) ds + K_T - K_t \\ \quad - \int_t^T Z_s dW_s - \int_t^T \int_A U_s(a) \tilde{\mu}(ds da), \\ U_t(a) \leq 0. \end{cases}$$

Then we see that the above equation is a particular case of a backward stochastic differential equation studied in a general non-Markovian framework in [24]. In particular, existence and uniqueness of the minimal solution  $(Y, Z, U, K)$  to

equation (5.10) [or, equivalently, to equation (5.4)] follow from Theorem 2.1 in [24]. Indeed, Assumption **(H0)** in [24] is clearly satisfied. Concerning Assumption **(H1)**, this is only used in Lemma 2.2 of [24] to prove that the sequence  $(Y^n)_n$  satisfies (5.9), a property that in our setting follows directly from (2.4) and (3.6).

Finally, from Theorem 2.1 in [24] we also have that  $Y_t^n(\bar{\omega})$  converges increasingly to  $Y_t(\bar{\omega})$  as  $n \rightarrow \infty$ ,  $\mathbb{P}(d\bar{\omega})$ -a.s. Since  $\mathcal{V} = \bigcup_n \mathcal{V}_n$ , letting  $n \rightarrow \infty$  in (5.8) we obtain (5.5).  $\square$

We end this section proving the following generalization of formula (5.5).

**THEOREM 5.2** (Randomized dynamic programming principle). *For all  $0 \leq t \leq T$ , we have*

$$\begin{aligned}
 (5.11) \quad Y_t &= \operatorname{ess\,sup}_{\nu \in \mathcal{V}} \operatorname{ess\,sup}_{\tau \in \mathcal{T}} \mathbb{E}^\nu \left[ \int_t^\tau \rho_r(f_r(\cdot, I)) \, dr + Y_\tau \mid \mathcal{F}_t^{W, \mu} \right] \\
 &= \operatorname{ess\,sup}_{\nu \in \mathcal{V}} \operatorname{ess\,inf}_{\tau \in \mathcal{T}} \mathbb{E}^\nu \left[ \int_t^\tau \rho_r(f_r(\cdot, I)) \, dr + Y_\tau \mid \mathcal{F}_t^{W, \mu} \right],
 \end{aligned}$$

where  $\mathcal{T}$  denotes the class of  $[0, T]$ -valued  $\mathbb{F}^{W, \mu}$ -stopping times.

**PROOF.** For every  $n$ , proceeding along the same lines as in the proof of Lemma 4.8 in [15], we obtain

$$\begin{aligned}
 (5.12) \quad Y_t^n &= \mathbb{E}^\nu \left[ Y_\tau^n + \int_t^\tau \rho_r(f_r(\cdot, I)) \, dr \right. \\
 &\quad \left. + \int_t^\tau \int_A [n(U_r^n(a))^+ - v_r(a)U_r^n(a)] \lambda(da) \, dr \mid \mathcal{F}_t^{W, \mu} \right] \\
 &\geq \mathbb{E}^\nu \left[ Y_\tau^n + \int_t^\tau \rho_r(f_r(\cdot, I)) \, dr \mid \mathcal{F}_t^{W, \mu} \right] \quad \text{for all } \nu \in \mathcal{V}_n, \tau \in \mathcal{T}.
 \end{aligned}$$

Recalling that  $Y \geq Y^n$ , we find

$$Y_t \geq \mathbb{E}^\nu \left[ Y_\tau^n + \int_t^\tau \rho_r(f_r(\cdot, I)) \, dr \mid \mathcal{F}_t^{W, \mu} \right] \quad \text{for all } \nu \in \mathcal{V}_n, \tau \in \mathcal{T}.$$

Letting  $n \rightarrow \infty$ , and recalling that  $\mathcal{V}_n \subset \mathcal{V}$ , we end up with

$$Y_t \geq \mathbb{E}^\nu \left[ Y_\tau + \int_t^\tau \rho_r(f_r(\cdot, I)) \, dr \mid \mathcal{F}_t^{W, \mu} \right] \quad \text{for all } \nu \in \mathcal{V}, \tau \in \mathcal{T}.$$

The above inequality yields

$$\begin{aligned}
 Y_t &\geq \operatorname{ess\,sup}_{\nu \in \mathcal{V}} \operatorname{ess\,sup}_{\tau \in \mathcal{T}} \mathbb{E}^\nu \left[ \int_t^\tau \rho_r(f_r(\cdot, I)) \, dr + Y_\tau \mid \mathcal{F}_t^{W, \mu} \right], \\
 Y_t &\geq \operatorname{ess\,sup}_{\nu \in \mathcal{V}} \operatorname{ess\,inf}_{\tau \in \mathcal{T}} \mathbb{E}^\nu \left[ \int_t^\tau \rho_r(f_r(\cdot, I)) \, dr + Y_\tau \mid \mathcal{F}_t^{W, \mu} \right].
 \end{aligned}$$

It remains to prove the reverse inequalities. As in the proof of Lemma 4.8 in [15], for every  $n$  and  $\varepsilon \in (0, 1)$ , we define  $v_r^{\varepsilon,n}(a) = n1_{\{U_r^n(a) \geq 0\}} + \varepsilon 1_{\{-1 < U_r^n(a) < 0\}} - \varepsilon(U_r^n(a))^{-1}1_{\{U_r^n(a) \leq -1\}}$ . Then  $v^{\varepsilon,n} \in \mathcal{V}_n$  and

$$n(U_r^n(a))^+ - v_r^{\varepsilon,n}(a)U_r^n(a) \leq \varepsilon \quad \text{for all } r \in [0, T].$$

Therefore, from equality (5.12), we find

$$\begin{aligned} Y_t^n &= \mathbb{E}^{v^{\varepsilon,n}} \left[ Y_\tau^n + \int_t^\tau \rho_r(f_r(\cdot, I)) dr \right. \\ &\quad \left. + \int_t^\tau \int_A [n(U_r^n(a))^+ - v_r^{\varepsilon,n}(a)U_r^n(a)] \lambda(da) dr \mid \mathcal{F}_t^{W,\mu} \right] \\ &\leq \mathbb{E}^{v^{\varepsilon,n}} \left[ Y_\tau^n + \int_t^\tau \rho_r(f_r(\cdot, I)) dr \mid \mathcal{F}_t^{W,\mu} \right] + \varepsilon \lambda(A)T \quad \text{for all } \tau \in \mathcal{T}. \end{aligned}$$

Then we obtain the two following inequalities:

$$\begin{aligned} Y_t^n &\leq \operatorname{ess\,sup}_{\tau \in \mathcal{T}} \mathbb{E}^{v^{\varepsilon,n}} \left[ Y_\tau^n + \int_t^\tau \rho_r(f_r(\cdot, I)) dr \mid \mathcal{F}_t^{W,\mu} \right] + \varepsilon \lambda(A)T, \\ Y_t^n &\leq \operatorname{ess\,inf}_{\tau \in \mathcal{T}} \mathbb{E}^{v^{\varepsilon,n}} \left[ Y_\tau^n + \int_t^\tau \rho_r(f_r(\cdot, I)) dr \mid \mathcal{F}_t^{W,\mu} \right] + \varepsilon \lambda(A)T. \end{aligned}$$

As a consequence, we get (we continue the proof with “ess inf” over  $\tau \in \mathcal{T}$ , since the proof with “ess sup” can be done proceeding along the same lines)

$$Y_t^n \leq \operatorname{ess\,sup}_{v \in \mathcal{V}_n} \operatorname{ess\,inf}_{\tau \in \mathcal{T}} \mathbb{E}^v \left[ Y_\tau^n + \int_t^\tau \rho_r(f_r(\cdot, I)) dr \mid \mathcal{F}_t^{W,\mu} \right] + \varepsilon \lambda(A)T.$$

Using the arbitrariness of  $\varepsilon$ , and recalling that  $\mathcal{V}_n \subset \mathcal{V}$  and  $Y^n \leq Y$ , we obtain

$$Y_t^n \leq \operatorname{ess\,sup}_{v \in \mathcal{V}} \operatorname{ess\,inf}_{\tau \in \mathcal{T}} \mathbb{E}^v \left[ Y_\tau + \int_t^\tau \rho_r(f_r(\cdot, I)) dr \mid \mathcal{F}_t^{W,\mu} \right].$$

The claim follows letting  $n \rightarrow \infty$ .  $\square$

### REFERENCES

[1] BAIN, A. and CRISAN, D. (2009). *Fundamentals of Stochastic Filtering. Stochastic Modelling and Applied Probability* **60**. Springer, New York. MR2454694

[2] BANDINI, E. (2017). Optimal control of piecewise deterministic Markov processes: A BSDE representation of the value function. *ESAIM Control Optim. Calc. Var.* To appear. DOI:10.1051/cocv/2017006.

[3] BANDINI, E., COSSO, A., FUHRMAN, M. and PHAM, H. (2016). Randomization method and backward SDEs for optimal control of partially observed path-dependent stochastic systems. Preprint. Available at arXiv:1511.09274v2.

[4] BANDINI, E., COSSO, A., FUHRMAN, M. and PHAM, H. (2016). Randomized filtering and Bellman equation in Wasserstein space for partial observation control problem. Preprint. Available at arXiv:1609.02697.

- [5] BANDINI, E. and FUHRMAN, M. (2017). Constrained BSDEs representation of the value function in optimal control of pure jump Markov processes. *Stochastic Process. Appl.* **127** 1441–1474.
- [6] BENSOUSSAN, A. (1992). *Stochastic Control of Partially Observable Systems*. Cambridge Univ. Press, Cambridge.
- [7] BERTSEKAS, D. P. and SHREVE, S. E. (1978). *Stochastic Optimal Control: The Discrete Time Case. Mathematics in Science and Engineering* **139**. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York–London.
- [8] BOUCHARD, B. (2009). A stochastic target formulation for optimal switching problems in finite horizon. *Stochastics* **81** 171–197. [MR2571686](#)
- [9] CHOUKROUN, S. and COSSO, A. (2016). Backward SDE representation for stochastic control problems with nondominated controlled intensity. *Ann. Appl. Probab.* **26** 1208–1259. [MR3476636](#)
- [10] CHOUKROUN, S., COSSO, A. and PHAM, H. (2015). Reflected BSDEs with nonpositive jumps, and controller-and-stopper games. *Stochastic Process. Appl.* **125** 597–633.
- [11] COSSO, A., FUHRMAN, M. and PHAM, H. (2016). Long time asymptotics for fully nonlinear Bellman equations: A backward SDE approach. *Stochastic Process. Appl.* **126** 1932–1973.
- [12] ELIE, R. and KHARROUBI, I. (2014). Adding constraints to BSDEs with jumps: An alternative to multidimensional reflections. *ESAIM Probab. Stat.* **18** 233–250.
- [13] EL KAROUI, N., HUU NGUYEN, D. and JEANBLANC-PICQUÉ, M. (1988). Existence of an optimal Markovian filter for the control under partial observations. *SIAM J. Control Optim.* **26** 1025–1061.
- [14] FABBRI, G., GOZZI, F. and ŚWIĘCH, A. (2017). *Stochastic Optimal Control in Infinite Dimensions: Dynamic Programming and HJB Equations*. Springer, Berlin. With Chapter 6 by, Fuhrman, M. and Tessitore, G.
- [15] FUHRMAN, M. and PHAM, H. (2015). Randomized and backward SDE representation for optimal control of non-Markovian SDEs. *Ann. Appl. Probab.* **25** 2134–2167.
- [16] FUHRMAN, M., PHAM, H. and ZENI, F. (2016). Representation of non-Markovian optimal stopping problems by constrained BSDEs with a single jump. *Electron. Commun. Probab.* **21** Paper No. 3, 7.
- [17] FUHRMAN, M. and TESSITORE, G. (2004). Existence of optimal stochastic controls and global solutions of forward–backward stochastic differential equations. *SIAM J. Control Optim.* **43** 813–830. [MR2114377](#)
- [18] GOZZI, F. and ŚWIĘCH, A. (2000). Hamilton–Jacobi–Bellman equations for the optimal control of the Duncan–Mortensen–Zakai equation. *J. Funct. Anal.* **172** 466–510. [MR1753181](#)
- [19] JACOD, J. (1974/75). Multivariate point processes: Predictable projection, Radon–Nikodým derivatives, representation of martingales. *Z. Wahrsch. Verw. Gebiete* **31** 235–253. [MR0380978](#)
- [20] JACOD, J. (1979). *Calcul Stochastique et Problèmes de Martingales. Lecture Notes in Math.* **714**. Springer, Berlin.
- [21] KHARROUBI, I., LANGRENÉ, N. and PHAM, H. (2014). A numerical algorithm for fully nonlinear HJB equations: An approach by control randomization. *Monte Carlo Methods Appl.* **20** 145–165.
- [22] KHARROUBI, I., LANGRENÉ, N. and PHAM, H. (2015). Discrete time approximation of fully nonlinear HJB equations via BSDEs with nonpositive jumps. *Ann. Appl. Probab.* **25** 2301–2338.
- [23] KHARROUBI, I., MA, J., PHAM, H. and ZHANG, J. (2010). Backward SDEs with constrained jumps and quasi-variational inequalities. *Ann. Probab.* **38** 794–840.
- [24] KHARROUBI, I. and PHAM, H. (2015). Feynman–Kac representation for Hamilton–Jacobi–Bellman IPDE. *Ann. Probab.* **43** 1823–1865.

- [25] KRYLOV, N. V. (2009). *Controlled Diffusion Processes. Stochastic Modelling and Applied Probability* **14**. Springer, Berlin. [MR2723141](#)
- [26] LIONS, P.-L. (1989). Viscosity solutions of fully nonlinear second order equations and optimal stochastic control in infinite dimensions. II. Optimal control of Zakai's equation. In *Stochastic Partial Differential Equations and Applications. Lecture Notes in Math.* **1390** 147–170. Springer, Berlin.
- [27] PARDOUX, É. (1982). Equations of nonlinear filtering and application to stochastic control with partial observation. In *Nonlinear Filtering and Stochastic Control. Lect. Notes in Math.* **972** 208–248. Springer, Berlin.
- [28] PARDOUX, É. and PENG, S. (1992). Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic Partial Differential Equations and Their Applications (Charlotte, NC, 1991). Lect. Notes Control Inf. Sci.* **176** 200–217. Springer, Berlin. [MR1176785](#)
- [29] PENG, S. (2007).  $G$ -expectation,  $G$ -Brownian motion and related stochastic calculus of Itô type. In *Stochastic Analysis and Applications. Abel Symp.* **2** 541–567. Springer, Berlin. [MR2397805](#)
- [30] ROGERS, L. C. G. and WILLIAMS, D. (2000). *Diffusions, Markov Processes, and Martingales. Vol. 2. Itô Calculus. Cambridge Mathematical Library.* Cambridge Univ. Press, Cambridge. [MR1780932](#)
- [31] SONER, H. M., TOUZI, N. and ZHANG, J. (2012). Wellposedness of second order backward SDEs. *Probab. Theory Related Fields* **153** 149–190.
- [32] TANG, S. (1998). The maximum principle for partially observed optimal control of stochastic differential equations. *SIAM J. Control Optim.* **36** 1596–1617. [MR1626880](#)
- [33] TANG, S. and LI, X. (1994). Necessary conditions for optimal control of stochastic systems with random jumps. *SIAM J. Control Optim.* **32** 1447–1475.
- [34] ZABCZYK, J. (1996). *Chance and Decision: Stochastic Control in Discrete Time.* Scuola Normale Superiore, Pisa. [MR1678432](#)

E. BANDINI  
 DIPARTIMENTO DI ECONOMIA E FINANZA  
 LIBERA UNIVERSITÀ DEGLI  
 STUDI SOCIALI “GUIDO CARLI”  
 VIALE ROMANIA 32  
 00197 ROMA  
 ITALY  
 E-MAIL: [ebandini@luiss.it](mailto:ebandini@luiss.it)

M. FUHRMAN  
 DIPARTIMENTO DI MATEMATICA  
 UNIVERSITÀ DI MILANO  
 VIA SALDINI 50  
 20133 MILANO  
 ITALY  
 E-MAIL: [marco.fuhrman@unimi.it](mailto:marco.fuhrman@unimi.it)

A. COSSO  
 DIPARTIMENTO DI MATEMATICA  
 POLITECNICO DI MILANO  
 VIA BONARDI 9  
 20133 MILANO  
 ITALY  
 E-MAIL: [andrea.cosso@polimi.it](mailto:andrea.cosso@polimi.it)

H. PHAM  
 LABORATOIRE DE PROBABILITÉS  
 ET MODÈLES ALÉATOIRES  
 CNRS, UMR 7599  
 UNIVERSITÉ PARIS DIDEROT  
 75013 PARIS  
 FRANCE  
 E-MAIL: [pham@math.univ-paris-diderot.fr](mailto:pham@math.univ-paris-diderot.fr)