MARTINGALE APPROACH TO CONTROL FOR GENERAL JUMP PROCESSES

HERNÁNDEZ-HERNÁNDEZ, M. E.[1], JACKA, S. D.[2], AND MIJATOVIĆ, A.[3]

ABSTRACT. We provide verification theorems (at different levels of generality) for infinite horizon stochastic control problems in continuous time for semimartingales. The control framework is given as an abstract "martingale formulation", which encompasses a broad range of standard control problems. Under appropriate conditions we show that the set of admissible controls gives rise to a certain class of controlled special semimartingales. Our results generalise both the standard controlled Itô - and Lévy -diffusion settings as we allow ourselves to locally control not only the drift and diffusion coefficients, but also the jump intensity measure of the jumps. As an illustration, we present a few examples with explicit solutions.

1. INTRODUCTION

As far as we know, to date the most general setting for controlled Markov processes in continuous time is that in [20]. In this paper we provide sufficient conditions of optimality for the control problem of minimising the cost functional

$$J(\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}^{\mathbf{x}}) := \mathbb{E}^{\alpha}_{\mathbf{x}} \left[\int_0^\infty e^{-\gamma_t^{\boldsymbol{\alpha}}} f\left(\mathbf{X}_t^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}_t^{\mathbf{x}}\right) \mathrm{d}t \right], \qquad (1.1)$$

over all *admissible* state-control processes $\left(\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}^{\mathbf{x}}\right) \in \mathcal{A}_{\mathbf{x}}^p$, where $\mathbf{X}_0^{\boldsymbol{\alpha}^{\mathbf{x}}} = \mathbf{x}$. We assume that each control process $\boldsymbol{\alpha}_\cdot^{\mathbf{x}} = (\boldsymbol{\sigma}_\cdot, \nu_\cdot, \boldsymbol{\mu}_\cdot)$ determines the volatility, the jump intensity and the drift of the system $\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}$. We stress that, unlike most of the literature including [20], the action space allows us to choose the jump measure ν in a position-dependent way and from an arbitrary class. The local dynamics of the controlled process $\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}$ are described by nonlocal operators of the form

$$(L^{\mathbf{a}}g)(\cdot) := (\mathbf{u} + \boldsymbol{\mu})^T \nabla g(\cdot) + \frac{1}{2} \operatorname{Tr}(\boldsymbol{\sigma}^T \mathrm{H}g\, \boldsymbol{\sigma})(\cdot) + \int_{\mathbb{R}_0^n} \left(g(\cdot + \mathbf{y}) - g(\cdot) - \mathbf{y}^T \nabla g(\cdot)\right) \nu(\mathrm{d}\mathbf{y}), \quad (1.2)$$

where each $\mathbf{a} = (\boldsymbol{\sigma}, \nu, \boldsymbol{\mu})$ is an element in the (properly defined) set of available actions A. Here, f is the given running cost function and $\gamma_t^{\boldsymbol{\alpha}}$ denotes a discounting process depending on the admissible pair $\left(\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}^{\mathbf{x}}\right)$.

The contribution of this paper to the subject of stochastic control is twofold: firstly, we set up a general (and abstract) martingale formulation for multidimensional controlled semimartingales whose differential characteristics are controlled continuously in time. Secondly, we explore different characterisations of the value function at different levels of generality, including the standard verification theorems which are based on the corresponding HJB type equation. We also show the interplay between the different assumptions and the relationship with the classical stochastic differential equation (SDE) setting (see, for example, [22], [19] and references therein).

In particular, our control formulation encompasses the controlled Itô diffusion case and, further, generalises the controlled jump diffusion case presented in [20].

The basic structure of a stochastic control problem involves the description of the system dynamics and the corresponding class of admissible controls. Most of the literature dealing with verification theorems is set up via either a *strong* or a *weak* SDE formulation with a *Markovian* structure. That is, the evolution of each controlled system is governed by an SDE whose relevant coefficients are functions of the type $h(t, \mathbf{x}, u(t, \mathbf{x}))$, where u is the control policy. In such a setting the definition of the "smallest" class of admissible controls relies then on well-posedness results (existence and uniqueness of solutions) of the corresponding SDE's. Along these lines, the study of continuous time stochastic control of Markov processes, and more generally of Ito-diffusions, has been widely researched. For a quick overview of this formulation, see the surveys [3] and [21]. For a more detailed study we refer to [8], [7], [33] and references therein. Control settings allowing discontinuous processes include those of stochastic control for Markovian jump systems (also known as Markov decision processes) as given, for example, in [18], [31], [23]; and, more generally, controlled jump diffusions (also known as Lévy diffusions) as in [20].

When compared to the controlled SDE approaches, either for (continuous) Itô diffusions or for (discontinuous) Lévy diffusions, the amount of research outside these settings, such as the *control martingale approach*, is less. Some works in this direction are [13], [14], [15], [12], [9]. In this paper we also consider a control setting outside the SDE formulation. The main features of our model are the following:

i) Our control problem is based on a *martingale approach* similar to the one given in [9] (see also [4], [15]), wherein the underlying local dynamics, i.e. each admissible pair $(\mathbf{X}^{\alpha^{\mathbf{x}}}, \alpha^{\mathbf{x}})$, is characterised as a solution to a *control martingale problem*. This approach is also similar to the one introduced by Stroock and Varadhan for SDE's. This formulation allows very general underlying dynamics for the controlled system where neither is the dynamics of the system assumed to be Markovian nor is the class of admissible policies restricted to be *Markovian or state-feedback* policies. We assume a multidimensional state space with no state constraints. The general setting presented here encompasses typical stochastic control problems for Itô difusions.

ii) The action set A is an open abstract space for which each action $\mathbf{a} \in A$ is a triplet which determines (locally) the diffusion coefficient, the jump intensity and the drift of the underlying process. We are thus outside the standard assumption of taken $A \subset \mathbb{R}^k$, $k \in \mathbb{N}$. The local dynamics of the controlled system $\mathbf{X}^{\alpha^{\mathbf{x}}}$ is described by means of the family of Lévy operators $\{L^{\mathbf{a}} : \mathbf{a} \in A\}$ as defined in (1.2). The use of these operators allows us to study a general class of controlled processes without any *a priori* restriction to the Markovian class.

Our main results are the following:

i) *Controlled semimartingale dynamics.* We prove that our controlled processes are special semimartingales whose *differential semimartingale characteristics* are related to the associated control policy $\alpha^{\mathbf{x}}$ (Corollary 4.4). The generality of our setting imposes stronger integrability conditions for the control $\alpha^{\mathbf{x}}$ so as to guarantee the finiteness of the pth moments of $\mathbf{X}^{\alpha^{\mathbf{x}}}$ (Proposition 4.8), which are crucial when dealing with value functions of polynomial growth. In the standard SDE approach, where polynomial growth is also a natural assumption, the corresponding integrability conditions are usually granted by the Itô conditions. This is discussed in Section 6.1.

ii) *Dynamic Programming Principle (DPP).* In order to use the dynamic programming approach to study our control problem, our starting point is to prove that the DPP (also known as Bellman's principle of optimality) holds true in our abstract martingale setting (Lemma 5.5 and Lemma 5.6). Such a principle is well-known for continuous controlled Markov processes. However, under more general frameworks, its validity is either assumed

true because of *its clear intuitive meaning* or the reader is referred to references which, in many cases, deal with slightly different models.

ii) *Verification theorems.* Under different assumptions on the initial control data (running cost function) and considering different classes of admissible policies (including the Markovian class), we provide different results to characterise the value function including the standard *verification theorems* (Theorem 5.13). These results provide sufficient conditions to characterise the optimal payoff function (or value function) as a solution to the corresponding Hamilton-Jacobi-Bellman (HJB) equation and, at the same time, they allow us to determine optimal feedback controls via the pointwise optimisation of the HJB equation. The proofs of these theorems rely on the probabilistic counterpart of *the dynamic programming principle.* Essentially, one needs to show that the so-called *Bellman process* associated with a control process (see definition in (B.10)) is a submartingale for any admissible policy, and it is a martingale whenever the policy is optimal. In particular, we cover in detail the case where the running cost function f has polynomial growth.

iii) *Examples.* As applications of our results, we provide a few examples with explicit solutions. In one of them the optimal controlled process is a Brownian motion that is jumped to the origin whenever it leaves a certain region. We also give the explicit solution for a control problem with a quadratic running cost function. For this case we determine the value function and exhibit an optimal policy whose associated optimal controlled process is an Ornstein-Uhnlenbeck type process. There is a close connection between linear-quadratic (LQ) optimal control problems and the quadratic case presented here. Indeed, although the controlled system associated with each admissible control is not assumed to be linear, we still obtain a linear, optimally controlled process. Furthermore, we get an optimal control whose drift component is linear in the state variable and whose diffusion component and jump intensity are related to the solution to an algebraic Riccati equation (see (8.11)-(8.12) in Section 4).

Needless to say, we are aware of the limitations arising when considering verification results in the context of *classical* (regular enough) solutions. It is well-known that (even for standard controlled diffusions) the regularity of the value function cannot be guaranteed in general. Nevertheless, even though our verification results are based on the assumption of a smooth solution to the HJB equation, our control setting gives a promising starting point to handing very general stochastic control problems in continuous time and with abstract control sets. The generalisation of our results to the context of *viscosity solution* is left as part of our future research. A brief discussion of the finite horizon case and other possible extensions are given in Section 7.

The rest of the paper is organised as follows. Section 2 introduces some standard notation. Section 3 describes the control setting of interest. In Section 4 we study the underlying dynamics of our controlled process. Here we show the semimartingale structure of our formulation. Then, in Section 5 we define the cost structure of the problem and state the main results of the paper: the dynamic programming principle and its (sub-)martingale formulation (Lemma 5.6), as well as the characterisation results of the value function (Lemma 5.10, Theorem 5.13 and Theorem 6.2). In Section 6 we consider different classes of admissible controls such as the Markovian case and its connection with the standard SDE setting. The finite horizon case and other possible generalisations are briefly discussed in Section 7. We provide some applications in Section 8. Finally, for the sake of clarity, the proofs of all our results are collected in the Appendices.

2. Notation

Let $\mathbb{R}^n$ be the n-dimensional Euclidean space where each point $\mathbf{x}$ is expressed by a column vector $\mathbf{x} = (x_1, \ldots, x_n)^T$. As usual, the superscript "T" denotes the transpose of a vector (or a matrix). The set of nonnegative real numbers is denoted by $\mathbb{R}_+$. For any two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

the inner product and the Eucliden norm are denoted by $\mathbf{x} \cdot \mathbf{y} := \sum_{i=1}^{n} x_i y_i$ and $|\mathbf{x}| := \sqrt{\mathbf{x} \cdot \mathbf{x}}$, respectively.

Notation $M(\mathbb{R}^n)$, $B(\mathbb{R}^n)$ and $C(\mathbb{R}^n)$ denote the spaces of real-valued measurable functions, bounded measurable functions and continuous functions on $\mathbb{R}^n$, respectively. These spaces are endowed with the sup-norm $\|f\| = \sup_{\mathbf{x}} |f(\mathbf{x})|$. We denote by $C^k(\mathbb{R}^n)$, $k \in \mathbb{N}$, the space of real-valued k-times continuously differentiable functions defined on $\mathbb{R}^n$. An additional subscript c will be used to denote the corresponding space of functions with compact support and by $C_c^\infty(\mathbb{R}^n) := \bigcap_{k=1}^{\infty} C_c^k(\mathbb{R}^n)$ we denote the space of infinitely often continuously differentiable functions on $\mathbb{R}^n$ with compact support.

Given a metric space (E, d), $\mathcal{B}(E)$ denotes the Borel σ-algebra compatible with the metric d. Given the measure space $(E, \mathcal{B}(E), \rho)$, $L^q(E, \mathcal{B}(E), \rho)$ (in brief $L^q(\rho)$ if there is no risk of confusion), $1 \le q < \infty$, stands for the Banach space of all equivalence classes of mappings $f : E \to \mathbb{R}^n$ which agree a.e. with respect to ρ and for which $\|f\|_q < \infty$, where the q-norm $\| \cdot \|_q$ is given by $\|f\|_q := \left(\int_E |f(x)|^q \, \mathrm{d}\rho \right)^{1/q}$.

Let $\mathrm{M}_{n \times n}(\mathbb{R})$ be the set of real-valued $n \times n$-matrices endowed with the Frobenius norm $\|\mathbf{B}\| := \mathrm{Tr}(\mathbf{B}\mathbf{B}^T)^{1/2}$, $\mathbf{B} = (B_{ij})_{1 \le i,j \le n} \in \mathrm{M}_{n \times n}(\mathbb{R})$, where $\mathrm{Tr}(\mathbf{C}) := \sum_{i=1}^{n} C_{ii}$ denotes the trace of the matrix $\mathbf{C} \in \mathrm{M}_{n \times n}(\mathbb{R})$. The $n \times n$ identity matrix will be denoted by $\mathbf{I}$.

Let $\mathbb{R}_0^n := \mathbb{R}^n \backslash \{0\}$ and fix a $p \ge 2$. Define

$$\mathcal{M}_p := \left\{ \text{ measures } \nu \text{ on } (\mathbb{R}_0^n, \mathcal{B}(\mathbb{R}_0^n)) \text{ such that } \int_{\mathbb{R}_0^n} |\mathbf{y}|^2 \vee |\mathbf{y}|^p \nu(\mathrm{d}\mathbf{y}) < +\infty \right\}. \tag{2.1}$$

Since $1 \wedge |\mathbf{y}|^2 \le |\mathbf{y}|^2$ for all $\mathbf{y} \in \mathbb{R}^n$, it follows that each $\nu \in \mathcal{M}_p$ is a Lévy measure, i.e.

$$\int_{\mathbb{R}_0^n} (1 \wedge |\mathbf{y}|^2) \nu(\mathrm{d}\mathbf{y}) < +\infty, \tag{2.2}$$

and, further, ν satisfies $\int_{\mathbb{R}_0^n} |\mathbf{y}|^2 \nu(\mathrm{d}\mathbf{y}) < +\infty$. Here, we used the notation $a \wedge b := \min\{a, b\}$ and $a \vee b := \max\{a, b\}$, for any $a, b \in \mathbb{R}$. The space $\mathcal{M}_p$ is endowed with a suitable weak convergence topology and its Borel σ-algebra is denoted by $\mathcal{B}(\mathcal{M}_p)$.

For any $\boldsymbol{\sigma} \in \mathrm{M}_{n \times n}(\mathbb{R})$, the matrix $\boldsymbol{\sigma}^T \boldsymbol{\sigma}$ is symmetric positive semidefinite. By Theorem 8.1 in [28, Chapter 8, p. 37] and Theorem 7.10 in [28, Chapter 7, p. 35], given $\boldsymbol{\sigma} \in \mathrm{M}_{n \times n}(\mathbb{R})$, $\boldsymbol{\mu} \in \mathbb{R}^n$ and $\nu \in \mathcal{M}_p$, there exists a unique in law $\mathbb{R}^n$-valued process $\mathbf{X}^{\mathbf{x}} := (\mathbf{X}_t^{\mathbf{x}})_{t \in \mathbb{R}_+}$, started at $\mathbf{x} \in \mathbb{R}^n$, having stationary independent increments and satisfying

$$\mathbb{E}_{\mathbf{x}} \left[e^{i\mathbf{u} \cdot (\mathbf{X}_t^{\mathbf{x}} - \mathbf{x})} \right] = e^{t\Psi(\mathbf{u})}, \quad \text{for all } \mathbf{u} \in \mathbb{R}^n, \ t \in \mathbb{R}_+,$$

where

$$\Psi(\mathbf{u}) := -\frac{1}{2} \mathbf{u} \cdot \boldsymbol{\sigma}^T \boldsymbol{\sigma} \mathbf{u} + i\mathbf{u} \cdot \boldsymbol{\mu} + \int_{\mathbb{R}_0^n} \left(e^{i\mathbf{u} \cdot \mathbf{y}} - 1 - i\mathbf{u} \cdot \mathbf{y} \right) \nu(\mathrm{d}\mathbf{y}), \quad \mathbf{u} \in \mathbb{R}^n. \tag{2.3}$$

The infinitesimal generator of the n-dimensional process $\mathbf{X}^{\mathbf{x}}$ is the operator L, with domain $\mathrm{Dom}(L)$, defined for any $g \in C_c^2(\mathbb{R}^n)$ by

$$(Lg)(\cdot) := \boldsymbol{\mu}^T \nabla g(\cdot) + \frac{1}{2} \mathrm{Tr}(\boldsymbol{\sigma}^T \mathrm{H}g \, \boldsymbol{\sigma})(\cdot) + \int_{\mathbb{R}_0^n} \left(g(\cdot + \mathbf{y}) - g(\cdot) - \mathbf{y}^T \nabla g(\cdot) \right) \nu(\mathrm{d}\mathbf{y}), \tag{2.4}$$

see [28, Theorem 31.5, p. 208]. Here, ∇g and $\mathrm{H}g$ denote the gradient and the Hessian in the variable $\mathbf{x}$ with components $\partial_i g := \frac{\partial}{\partial x_i} g$, $1 \le i \le n$, and $\partial_{ij}^2 g := \frac{\partial^2}{\partial x_i \partial x_j} g$, $1 \le i, j \le n$, respectively.

3. STOCHASTIC CONTROL SETTING

We are interested in studying a stochastic control problem in a continuous setting: continuous time and continuous action and state space. We shall consider a *martingale control formulation* to describe the dynamics of each controlled system, or rather its law, as determined by the

martingale property of a certain class of processes. For each controlled system, the local dynamics are described by a given family of operators defined on a certain class of test functions.

Notation. Given a metric space (E, d), $\mathcal{P}(E)$ denotes the set of probability measures on $(E, \mathcal{B}(E))$. The set of càdlàg functions (right-continuous with left-limits) on $\mathbb{R}_+$ with values in E is denoted by $D(\mathbb{R}_+; E)$ and is endowed with the Skorohod topology [6, Chapter 3, Section 5]. Given a stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, all equalities and inequalities between random variables are understood to hold $\mathbb{P}-$almost surely, unless stated otherwise.

3.1. **Control setting.** A *controlled system* with state space $\mathbb{R}^n$ and distribution $\xi \in \mathcal{P}(\mathbb{R}^n)$ at an initial time $r \in \mathbb{R}_+$ is described by the following elements.

(i) *Action set.* It is given by an open subset $A \subseteq M_{n\times n}(\mathbb{R}) \times \mathcal{M}_p \times \mathbb{R}^n$ and denotes the set of possible actions available at each instant time and it is endowed with the corresponding product Borel σ−algebra denoted by $\mathcal{B}(A)$ and derived from the Borel σ-algebra on $M_{n\times n}(\mathbb{R}) \times \mathcal{M}_p \times \mathbb{R}^n$.

(ii) *Instantaneous underlying dynamics.* The local dynamics of the system are specified by a family $\{(L^{\mathbf{a}}, D^{\mathbf{a}}) : \mathbf{a} \in A\}$ of operators $L^{\mathbf{a}} : D^{\mathbf{a}} \subset M(\mathbb{R}^n) \to M(\mathbb{R}^n)$, such that for each $\mathbf{a} = (\boldsymbol{\sigma}, \nu, \boldsymbol{\mu}) \in A$ and for a fixed vector $\mathbf{u} \in \mathbb{R}^n$, the operator $L^{\mathbf{a}}$ with domain $D^{\mathbf{a}}$ is defined, for each $g \in C_c^2(\mathbb{R}^n) \subset D^{\mathbf{a}}$, by

$$(L^{\mathbf{a}}g)(\cdot) := (\mathbf{u} + \boldsymbol{\mu})^T \nabla g(\cdot) + \frac{1}{2}\mathrm{Tr}(\boldsymbol{\sigma}^T \mathrm{H}g\,\boldsymbol{\sigma})(\cdot) + \int_{\mathbb{R}_0^n} \left(g(\cdot + \mathbf{y}) - g(\cdot) - \mathbf{y}^T \nabla g(\cdot)\right)\nu(\mathrm{d}\mathbf{y}). \quad (3.1)$$

(iii) *Control policies.* A control policy is an A-valued process $\boldsymbol{\alpha} = (\alpha_t)_{t\geq 0}$ determined by its probability law $\mathbf{P}_{\boldsymbol{\alpha}}$, i.e. $\mathbf{P}_{\boldsymbol{\alpha}}$ is a probability measure on $(D(\mathbb{R}_+; A))$. The set of all such control processes is denoted by $\mathcal{U}$.

(iv) *Admissible control policies.* The set of admissible policies $\mathcal{A}_{r,\xi}^p \subseteq \mathcal{U}$ with initial condition $(r, \xi) \in \mathbb{R}_+ \times \mathcal{P}(\mathbb{R}^n)$ is defined as the class of control processes $\boldsymbol{\alpha}^{r,\xi} = \left(\alpha_t^{r,\xi}\right)_{t\geq r} \in \mathcal{U}$ satisfying the following:

(**H1**) There exists a filtered, complete probability space[1] $(\Omega^{\alpha}, \mathcal{F}^{\alpha}, \mathbb{F}^{\alpha} = (\mathcal{F}_t^{\alpha}), \mathbb{P}^{\alpha})$ which supports both an $\mathbb{F}^{\alpha}$-adapted càdlàg copy of $\boldsymbol{\alpha}^{r,\xi}$ (denoted again by $\boldsymbol{\alpha}^{r,\xi}$) and an $\mathbb{F}^{\alpha}$-adapted $\mathbb{R}^n$-valued càdlàg process $\mathbf{X}^{\boldsymbol{\alpha}^{r,\xi}} = (\mathbf{X}_t^{\boldsymbol{\alpha}^{r,\xi}})_{t\geq r}$ with $\mathbb{P}^{\alpha} \circ \left(\mathbf{X}_s^{\boldsymbol{\alpha}^{r,\xi}}\right)^{-1} = \xi$ for all $0 \leq s \leq r$, such that the pair $(\mathbf{X}^{\boldsymbol{\alpha}^{r,\xi}}, \boldsymbol{\alpha}^{r,\xi})$ is unique in law and, further, for each $h \in C_c^2(\mathbb{R}^n)$, $\int_r^t \left|\left(L^{\alpha_s^{r,\xi}}h\left(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{r,\xi}}\right)\right)\right|\mathrm{d}s < \infty$, for $t > r$ $\mathbb{P}^{\alpha} - a.s.$, and the process $M^{h,\boldsymbol{\alpha}^{r,\xi}} = \left(M_t^{h,\boldsymbol{\alpha}^{r,\xi}}\right)_{t\geq r}$ defined by

$$M_t^{h,\boldsymbol{\alpha}^{r,\xi}} := h\left(\mathbf{X}_t^{\boldsymbol{\alpha}^{r,\xi}}\right) - \int_r^t \left(L^{\alpha_s^{r,\xi}}h\right)\left(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{r,\xi}}\right)\mathrm{d}s, \quad t \geq r, \quad (3.2)$$

is an $(\mathbb{F}^{\alpha}, \mathbb{P}^{\alpha})$-local martingale (i.e. an $\mathbb{F}^{\alpha}$-adapted local martingale under $\mathbb{P}^{\alpha}$).

(**H2**) If $\alpha_s^{r,\xi} = (\boldsymbol{\sigma}_s, \nu_s, \boldsymbol{\mu}_s)$, $s \geq r$, then for any $t \geq r$

$$\int_r^t Q_s^{p,\boldsymbol{\alpha}^{r,\xi}}\,\mathrm{d}s < \infty, \quad \text{where } Q_s^{p,\boldsymbol{\alpha}^{r,\xi}} := |\boldsymbol{\mu}_s| + \|\boldsymbol{\sigma}_s\|^2 + \int_{\mathbb{R}_0^n} |\mathbf{y}|^2 \vee |\mathbf{y}|^p \nu_s(\mathrm{d}\mathbf{y}), \quad \mathbb{P}^{\alpha} - a.s.. \quad (3.3)$$

Hereafter, any pair $\left(\mathbf{X}^{\boldsymbol{\alpha}^{r,\xi}}, \boldsymbol{\alpha}^{r,\xi}\right)$ with $\boldsymbol{\alpha}^{r,\xi} \in \mathcal{A}_{r,\xi}^p$ will be referred to as an *admissible pair*.

---

[1]As usual, we will always assume that any probability space satisfies the *usual conditions* [25, Chapter I]: $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ is a filtration of sub-σ−algebras of $\mathcal{F}$ such that $\mathcal{F}_0$ contains all the $\mathbb{P}$−null sets in $\mathcal{F}$, and $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ is right continuous, i.e. $\mathcal{F}_t = \bigcap_{s>t}\mathcal{F}_s =: \mathcal{F}_{t+}$.

Notation. For any $\mathbf{x} \in \mathbb{R}^n$, the initial conditions $(r, \delta_{\mathbf{x}})$ will be denoted by $(r, \mathbf{x})$, and thus $\mathcal{A}^p_{r, \delta_{\mathbf{x}}} \equiv \mathcal{A}^p_{r, \mathbf{x}}$ and $(\mathbf{X}^{\alpha^{r, \delta_{\mathbf{x}}}}, \alpha^{r, \delta_{\mathbf{x}}}) \equiv (\mathbf{X}^{\alpha^{r, \mathbf{x}}}, \alpha^{r, \mathbf{x}})$, whereas if $r = 0$, then $\mathcal{A}^p_{0, \delta_{\mathbf{x}}} \equiv \mathcal{A}^p_{\mathbf{x}}$ and $(\mathbf{X}^{\alpha^{0, \delta_{\mathbf{x}}}}, \alpha^{0, \delta_{\mathbf{x}}}) \equiv (\mathbf{X}^{\alpha^{\mathbf{x}}}, \alpha^{\mathbf{x}})$.

Definition 3.1 (**Control Martingale Problem**). *Let $\xi \in \mathcal{P}(\mathbb{R}^n)$ and $r \in \mathbb{R}_+$. The $\mathbb{R}^n \times A$-valued admissible pair $(\mathbf{X}^{\alpha^{r, \xi}}, \alpha^{r, \xi})$ or, more precisely, the sextuplet $(\Omega^\alpha, \mathcal{F}^\alpha, \mathbb{F}^\alpha, \mathbb{P}^\alpha, \mathbf{X}^{\alpha^{r, \xi}}, \alpha^{r, \xi})$, with $\mathbb{P}^\alpha \circ (\mathbf{X}^{\alpha^{r, \xi}}_r)^{-1} = \xi$, is called a solution to the* control martingale problem *for $(\{L^a : a \in A\}, C^2_c(\mathbb{R}^n), (r, \xi))$, whenever (**H1**) holds. Here $\{L^a : a \in A\}$ is the family of operators defined in (3.1) and ξ is the distribution of $\mathbf{X}^{\alpha^{r, \xi}}_r$.*

Remark 3.2. The set of admissible policies $\mathcal{A}^p_{r, \xi}$ can then be thought of as the set of laws of the pair $(\mathbf{X}^{\alpha^{r, \xi}}, \alpha^{r, \xi})$ viewed as a random element with values in $D(\mathbb{R}_+; A) \times D(\mathbb{R}_+; \mathbb{R}^n)$.

Convention. In order to extend the martingale problem (3.2) to $\mathbb{R}_+$ and given the assumption $\mathbb{P}^\alpha \circ (\mathbf{X}^{\alpha^{r, \xi}}_s)^{-1} = \xi$ for all $0 \leq s \leq r$, we set $\int_r^t \left| \left(L^{\alpha^{r, \xi}_s} h \right) (\mathbf{X}^{\alpha^{r, \xi}}_{s-}) \right| \mathrm{d}s = 0$ for $t \leq r$.

Remarks 3.3.

i) Since the stochastic basis $(\Omega^\alpha, \mathcal{F}^\alpha, \mathbb{F}^\alpha, \mathbb{P}^\alpha)$ in (**H1**) fulfils the usual conditions, by [26, Theorem 2.9, Chapter II, p. 65] any martingale defined on it admits a càdlàg version which will be the one we always work with.

ii) By (3.2) in (**H1**), the process $h(\mathbf{X}^{\alpha^{r, \xi}})$ is a special semimartingale for each $h \in C^2_c(\mathbb{R}^n)$. Observe that no further restrictions, such as assuming the Markov property, are imposed neither on the process $\mathbf{X}^{\alpha^{r, \xi}}$ nor on $\alpha^{r, \xi}$.

iii) It is feasible to consider a larger set of test functions in (3.2), for instance the space $C^2(\mathbb{R}^n)$. However, enlarging this set reduces the size of the set of admissible policies $\mathcal{A}^p_{r, \xi}$.

iv) Using the standard definition of (Lévy) generating triplet with respect to the truncation function $h(\mathbf{y}) = \mathbf{y} 1_{\{|\mathbf{y}| \leq 1\}}$ (see [28, Definition 8.2, p. 38]), one can see that the operator $L^{\mathbf{a}}$ in (3.1) is the infinitesimal generator of a Lévy process with triplet $(\sigma^T \sigma, \gamma, \nu)$ w.r.t h, where $\gamma := (\mathbf{u} + \boldsymbol{\mu})^T + \int_{|\mathbf{y}| > 1} |\mathbf{y}|^2 \nu(\mathrm{d}\mathbf{y})$ (for details, see also [28, Chapter 2, Section 7 and Theorem 31.5, p. 208]).

3.2. **Concatenation property and examples of admissible policies.** Since any solution to a martingale problem can only determine the law of its solution process and due to the càdlàg requirement for each admissible pair $(\mathbf{X}^{\alpha^{r, \xi}}, \alpha^{r, \xi})$, without loss of generality we may (and we will) consider solutions to the control martingale problem in the corresponding canonical space. More precisely, the sextuplet $(\Omega^\alpha, \mathcal{F}^\alpha, \mathbb{F}^\alpha, \mathbb{P}^\alpha, \mathbf{X}^{\alpha^{r, \xi}}, \alpha^{r, \xi})$ takes the form: $\Omega^\alpha := \Omega^{\alpha, 1} \times \Omega^{\alpha, 2}$ where $\Omega^{\alpha, 1} := D(\mathbb{R}_+; \mathbb{R}^n)$ and $\Omega^{\alpha, 2} := D(\mathbb{R}_+; A)$ are the space of càdlàg functions on $\mathbb{R}_+$ with values in $\mathbb{R}^n$ and A, respectively, both of them endowed with the corresponding Skorohod topology; $\mathbb{F}^\alpha := \mathcal{F}^{\alpha, 1} \otimes \mathcal{F}^{\alpha, 2}$, where $\mathcal{F}^{\alpha, i} := \mathcal{B}(\Omega^{\alpha, i})$, $i = 1, 2$. The solution process $(\mathbf{X}^{\alpha^{r, \xi}}_{\cdot}, \alpha^{r, \xi}_{\cdot})$ corresponds to the coordinate pair $(\mathbf{X}^{\alpha^{r, \xi}}_t(\omega), \alpha^{r, \xi}_t(\omega)) := (\omega^1(t), \omega^2(t))$, for each $\omega = (\omega^1, \omega^2) \in \Omega^\alpha$, $t \geq 0$. The filtration $\mathbb{F}^\alpha = (\mathcal{F}^\alpha_t)$ is defined by $\mathcal{F}^{\alpha, 1}_t \otimes \mathcal{F}^{\alpha, 2}_t$, where $\mathcal{F}^{\alpha, 1}_t := \sigma(\mathbf{X}^{\alpha^{r, \xi}}_s; s \leq t)$ and $\mathcal{F}^{\alpha, 2}_t := \sigma(\alpha^{r, \xi}_s; s \leq t)$ are the natural filtration generated by the coordinate processes $\mathbf{X}^{\alpha^{r, \xi}}$ and $\alpha^{r, \xi}$, respectively. Finally, the measure $\mathbb{P}^\alpha$ is a probability measure on $(\Omega^\alpha, \mathbb{F}^\alpha)$.

In the following result we prove a concatenation property for the admissible policies. This property is fundamental for the validity of the dynamic programming principle as will be shown in the next section.

Lemma 3.4. *Let $(r, \xi) \in \mathbb{R}_+ \times \mathcal{P}(\mathbb{R}^n)$ and let $\boldsymbol{\alpha}^{r,\xi} \in \mathcal{A}^p_{r,\xi}$. Denote by η^α_t the law of the r.v. $\mathbf{X}^{\boldsymbol{\alpha}^{r,\xi}}_t$, i.e. $\eta^\alpha_t := \mathbb{P}^\alpha \circ (\mathbf{X}^{\boldsymbol{\alpha}^{r,\xi}}_t)^{-1} \in \mathcal{P}(\mathbb{R}^n)$. Then, given any admissible policy $\boldsymbol{\beta}^\alpha_t \in \mathcal{A}^p_{t,\eta^\alpha_t}$, the concatenation of $\boldsymbol{\alpha}^{r,\xi}$ and $\boldsymbol{\beta}^\alpha_t$ at time $t \geq r$, denoted by $\boldsymbol{\alpha}^{r,\xi} \oplus_t \boldsymbol{\beta}^\alpha_t$ and defined by*

$$\boldsymbol{\alpha}^{r,\xi} \oplus_t \boldsymbol{\beta}^\alpha_t := \boldsymbol{\alpha}^{r,\xi}(s) 1_{[r,t)}(s) + \boldsymbol{\beta}^\alpha_t(s) 1_{[t,\infty)}(s). \tag{3.4}$$

belongs to the set of admissible controls $\mathcal{A}^p_{r,\xi}$.

Remark 3.5. All the proofs of our results are given in the Appendices. See A.1 for the proof of Lemma 3.4.

Examples of admissible policies.

i) Since each measure $\nu \in \mathcal{M}_p$, $p \geq 2$, has finite pth moments outside the unitary ball, for each *constant policy* the corresponding controlled process $\mathbf{X}^{\boldsymbol{\alpha}^\mathbf{x}}$ is a Lévy martingale with drift component and, further, it also has finite pth moments (see [2, Theorem 3.3.3, p. 163; p.133], [28, Theorem 31.5, p. 208; p. 39]). In this case, it is also known that there exists a unique solution to the martingale problem for each initial condition $X_0 = x$, and further, such a solution is the unique strong solution to the corresponding SDE.

ii) Since constant policies are admissible, Lemma 3.4 implies that (by pasting appropriately) the set of step controls (or piecewise constant controls) $\mathcal{A}^{step}_\mathbf{x}$ also belongs to $\mathcal{A}^p_\mathbf{x}$. We say that a control $\boldsymbol{\alpha}^\mathbf{x}$ belongs to $\mathcal{A}^{step}_\mathbf{x}$ if there exists $N \in \mathbb{N}$ and $\alpha_k \in A$, $k = 1, \ldots, N$, such that $\boldsymbol{\alpha}^\mathbf{x}_t = \alpha_k$ on $[t_{k-1}, t_k)$ for all $k = 1, \ldots, N$, where $t_0 = 0$ and $t_N = \infty$.

iii) The set $\mathcal{A}^p_\mathbf{x}$ also contains Markov policies of the form $\boldsymbol{\alpha}^\mathbf{x}_t := (\sigma(\mathbf{X}^\mathbf{x}_t), \nu(\mathbf{X}^\mathbf{x}_t, \cdot), \mu(\mathbf{X}^\mathbf{x}_t))$, $\mathbf{x} \in \mathbb{R}^n$, with continuous functions $\mu : \mathbb{R}^n \to \mathbb{R}^n$, $\sigma : \mathbb{R}^n \to \mathrm{M}_{n \times n}(\mathbb{R})$ and $\nu : \mathbb{R}^n \to \mathcal{M}_p$ satisfying

$$\sup_\mathbf{x} \left(|\mu(\mathbf{x})| + \|\sigma(\mathbf{x})\|^2 + \int_{\mathbb{R}^n_0} |\mathbf{y}|^2 \vee |\mathbf{y}|^p \nu(\mathbf{x}, \mathrm{d}\mathbf{y}) \right) < \infty.$$

The associated controlled process $\mathbf{X}^{\boldsymbol{\alpha}^\mathbf{x}}$ is a *Lévy -type process* $\mathbf{X}^\mathbf{x}$ *with bounded coefficients* (see [2, Chapter 3], [30]) and whose infinitesimal generator L^* is defined, on functions $g \in C^\infty_c(\mathbb{R}^n)$, by

$$(L^* g)(\mathbf{x}) := (\mathbf{u} + \mu(\mathbf{x}))^T \nabla g(\mathbf{x}) + \frac{1}{2} \mathrm{Tr}(\sigma^T(\mathbf{x}) \, \mathrm{H}g(\mathbf{x}) \, \sigma(\mathbf{x})) + \int_{\mathbb{R}^n_0} (g(\mathbf{x} + \mathbf{y}) - g(\mathbf{x}) - \mathbf{y}^T \nabla g(\mathbf{x}) \nu(\mathbf{x}, \mathrm{d}\mathbf{y}).$$
$$\tag{3.5}$$

4. First results: underlying control dynamics

We shall prove that, for each policy $\boldsymbol{\alpha}^\mathbf{x} \in \mathcal{A}^p_\mathbf{x}$ the corresponding controlled process is an $\mathbb{R}^n$-valued (special) semimartingale whose characteristics depend on $\boldsymbol{\alpha}^\mathbf{x}$.

Recall that, given a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} := (\mathcal{F}_t), \mathbb{P})$, an $\mathbb{F}$-adapted càdlàg process X is said to be a classical $(\mathbb{F}, \mathbb{P})$-*semimartingale* if it admits a (not necessarily unique) decomposition $X = X_0 + M + A$ $\mathbb{P}$-a.s., where X_0 is finite-valued and $\mathcal{F}_0$−measurable, M is an $(\mathbb{F}, \mathbb{P})$-local martingale with $M_0 = 0 = A_0$, and A is an $\mathbb{F}$-adapted càdlàg processes with paths of (locally) finite variation $\mathbb{P}$-a.s. If additionally A is predictable, then X is called a *special semimartingale* [10, Definition 4.21, I.4c, p. 43]. For *special semimartingales* its decomposition is unique (up to indistinguishability) and is known as the *canonical decomposition of X* ([10, Definition 4.22, I.4c, p. 43]). An adapted stochastic process is said to be *predictable* (resp. *optional*) if it is measurable with respect to the *predictable* σ-algebra $\mathbb{P}$ (resp. *optional* σ-algebra $\mathcal{O}$), i.e. the σ-algebra on $\Omega \times \mathbb{R}_+$ generated by all the left-continuous (resp. càdlàg) adapted processes Z considered as mappings $(\omega, t) \mapsto Z_t(\omega)$ on $\Omega \times \mathbb{R}_+$. A function $W : \Omega \times \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$ is called *predictable* (resp. *optional*) if it is measurable with respect to the σ-algebra of *predictable* sets (resp. *optional* sets) in $\Omega \times \mathbb{R}_+ \times \mathbb{R}^n$, given by $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^n)$ (resp. $\mathcal{O} \otimes \mathcal{B}(\mathbb{R}^n_t)$).

Another explicit representation for semimartingales can be given in terms of the so-called *semimartingale characteristics relative to a truncation function h*. Such a decomposition is known as the *canonical representation relative to h*. Let us briefly recall these concepts (see [10, Chapter II] for a detailed study). Let $X = (X^i)_{1 \leq i \leq n}$ be an n-dimensional semimartingale. Fix a truncation function $h : \mathbb{R}^n \to \mathbb{R}^n$ (i.e., a bounded measurable function with compact support such that $h(x) = x$ in a neighbourhood of the origin [10, Definition 2.3]). Define

$$
\begin{aligned}
\tilde{X}(h). &:= \sum_{s \leq \cdot} [\Delta X_s + h(\Delta X_s)] \\
X(h). &:= X. - \tilde{X}(h).,
\end{aligned}
$$

then $\tilde{X}(h)$ is an n-dimensional finite variation process and $X(h)$ is a process with uniformly bounded jumps (and thus is a special semimartingale) and differs from X by a finite variation process.

Now, recall also that, for a given state space S, a *random measure* μ on $\mathbb{R}_+ \times S$ is a family $\{\mu(\omega) : \omega \in \Omega\}$ of measures $\mu(\omega)$ on $(\mathbb{R}_+ \times S, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(S))$ satisfying $\mu(\omega; \{0\} \times S) = 0$ for each $\omega \in \Omega$. For any random measure μ and any optional function W on $\Omega \times \mathbb{R}_+ \times S$, we denote by $W * \mu$ the integral process

$$
\int_0^\cdot \int_S W(\omega; t, y) \mu(\omega; \mathrm{d}t, \mathrm{d}y). \tag{4.1}
$$

Given an adapted càdlàg $\mathbb{R}^n$-valued process X, the measure associated to its jumps is defined as the integer-valued *random measure* η^X on $\mathbb{R}^+ \times \mathbb{R}^n$ given by

$$
\eta^X(\omega; \mathrm{d}t, \mathrm{d}y) := \sum_s \mathbf{1}_{\{\Delta X_s(\omega) \neq 0\}} \delta_{(s, \Delta X_s(\omega))}(\mathrm{d}t, \mathrm{d}y),
$$

see [10, II.1b, Definition 1.13, Proposition 1.16, pp. 68-69]. Thus, for each path ω, $\eta^X(\omega; [0, t] \times D)$ gives the number of jumps whose sizes fall in the measurable set $D \subset \mathbb{R}^n$, during the time interval $[0, t]$. Since X is càdlàg , the random measure $\eta^X(\omega; \mathrm{d}t, \mathrm{d}y)$ takes only finite values for any Borel subset D of $\mathbb{R}^n$ bounded away from zero. Moreover, by [10, Theorem 1.8, Chapter II.1a, p. 66], there exists an unique (up to indistinguishability) *predictable* random measure η on $\mathbb{R}^+ \times \mathbb{R}^n$ for which, in particular, the integral process $W * (\eta^X - \eta)$ given by

$$
W * (\eta^X - \eta)_t(\omega) := \int_0^t \int_{\mathbb{R}^n} W(\omega, s, x)(\eta^X - \eta)(\omega; \mathrm{d}s, \mathrm{d}x))
$$

is a local martingale for each predictable function W on $\Omega \times \mathbb{R}_+ \times \mathbb{R}^n$ for which $|W| * (\eta^X - \eta)_t(\omega)$ is finite. The random measure $\eta(\omega; \mathrm{d}s, \mathrm{d}y)$ is called the *predictable compensator*, or *dual predictable projection* of the jump measure η^X.

Definition 4.1 (Semimartingale characteristics). *The characteristics of the semimartingale X w.r.t $\mathbb{P}$ and relative to h, are given by a triplet $(\mathbf{B}^h, \mathbf{C}, \eta)$, where* (i) *$\mathbf{B}^h = (B^i)_{i \leq n}$ is the $\mathbb{R}^n$-valued predictable process of finite variation in the canonical representation of the special semimartingale $X(h)$;* (ii) *$\mathbf{C} = (C^{ij})_{1 \leq i, j \leq n}$, $C^{ij} := \langle X^{i,c}, X^{j,c} \rangle$, is the $M_{n \times n}(\mathbb{R})$-valued continuous process with entries $< X^{i,c}, X^{j,c} >$, $1 \leq i, j \leq n$, where X^c denotes the continuous martingale part of X (with components $X^{i,c}$); and* (iii) *$\eta = \eta(\omega; \mathrm{d}t, \mathrm{d}y)$ is a random measure on $\mathbb{R}_+ \times \mathbb{R}_0^n$ and is the predictable compensator of the integer-valued random measure $\eta^X = \eta^X(\omega; \mathrm{d}t, \mathrm{d}x)$ associated with the jumps of X.*

Remark 4.2. Notice that the characteristics are unique up to indistinguishability. Furthermore, as notation indicates, $\mathbf{B}^h$ depends on the chosen truncation function h, whereas $\mathbf{C}$ and η are independent of h.

Remark 4.3. To simplify notation we shall fix the initial state $(0, \delta_{\mathbf{x}}) \in \mathbb{R}^n \times \mathcal{P}(\mathbb{R}^n)$ and work with the class of controls $\mathcal{A}_{\mathbf{x}}^p$. Our results can easily be extended to the general case $\mathcal{A}_{r, \xi}^p$ with $(r, \xi) \in \mathbb{R}_+ \times \mathcal{P}(\mathbb{R}^n)$.

We now state the relationship between each policy $\boldsymbol{\alpha}^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}^p$ and the semimartingale characteristics of its corresponding process $\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}$.

Proposition 4.4. *Let $h : \mathbb{R}^n \to \mathbb{R}^n$ be the function $h(\mathbf{y}) := \mathbf{y}\mathbf{1}_{\{|\mathbf{y}|\leq 1\}}$, $\mathbf{y} \in \mathbb{R}^n$. For each admissible policy $\boldsymbol{\alpha}^{\mathbf{x}} = (\boldsymbol{\sigma}, \nu, \boldsymbol{\mu}) \in \mathcal{A}_{\mathbf{x}}^p$, $p \geq 2$, the process $\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}$ is an $(\mathbb{F}^{\alpha}, \mathbb{P}^{\alpha})$-semimartingale with the predictable semimartingale characteristics (relative to h) $(\mathbf{B}^h, \mathbf{C}, \eta)$ given by*

i) $\mathbf{B}^h$ *is the process* $\mathbf{B}^h(t, \omega) = \int_0^t (\mathbf{u} + \boldsymbol{\mu}_s(\omega)) \, \mathrm{d}s + \int_0^t \mathrm{d}s \int_{\mathbb{R}_0^n} (\mathbf{y} - h(\mathbf{y})) \nu_s(\omega; \mathrm{d}\mathbf{y})$, *for* $t \geq 0$.

ii) $\mathbf{C} = (C_{ij}(\cdot))_{1 \leq i,j \leq n}$, *where* $C_{ij}(t, \omega) = \int_0^t \mathrm{d}s (\boldsymbol{\sigma}_s^T \boldsymbol{\sigma}_s)(\omega)_{ij}$, *for* $t \geq 0$.

iii) η *is the random measure on* $\mathbb{R}_+ \times \mathbb{R}_0^n$ *given by* $\eta(\omega; \mathrm{d}s, \mathrm{d}\mathbf{y}) := \mathrm{d}s \otimes \nu_s(\omega; \mathrm{d}\mathbf{y})$, $\omega \in \Omega^{\alpha}$.

Remark 4.5. $\mathbf{B}^h$ and $\mathbf{C}$ are the local drift coefficient and the local covariance matrix, respectively. We also recall that the predictable characteristics do not characterise the law of the process, however they provide useful information related to its jumps as we will see in Proposition 4.8.

Remarks 4.6.

i) The process $\mathbf{C}$ takes values on the set of all positive semidefinite symmetric $n \times n$-matrices. By definition, $\mathbf{C}$ is the unique (up to null sets) adapted continuous process, starting at $\mathbf{C}_0 = 0$, with paths $t \mapsto C_{ij}(t)$ having finite variation over compact intervals and such that $X_t^{i,c} X_t^{j,c} - C_{ij}(t)$ is a local martingale (see [27, IV.26]).

ii) By Proposition 4.4, for each admissible $\boldsymbol{\alpha}^{\mathbf{x}}$, the process $\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}$ is an *Itô semimartingale*, in the sense that its characteristics are absolutely continuous with respect to the Lebesgue measure [1, Definition 1.16, p. 45]. Hence, the control process $\boldsymbol{\alpha}^{\mathbf{x}}$ determines the *differential semimartingale characteristics* of $\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}$.

Remark 4.7. If X is a Lévy process with generating triplet (b, c, F), then its characteristics in the semimartingale sense are the not random functions:

$$\mathbf{B}_t(\omega) := bt, \qquad \mathbf{C}_t(\omega) := ct, \qquad \eta(\omega, \mathrm{d}t, \mathrm{d}x) := \mathrm{d}t \otimes F(\mathrm{d}x). \tag{4.2}$$

Proposition 4.4 identifies $\mathrm{d}s \otimes \nu_s(\omega, \mathrm{d}\mathbf{y})$ as the compensator of the random measure associated to the jumps of $\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}$, hereafter denoted by $\eta^{X, \boldsymbol{\alpha}^{\mathbf{x}}}$. This plays an important role in obtaining estimates of the running maximum of $|\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}|^q$ for $q > 2$, which are fundamental to deal with value functions of super-quadratic order. We now state conditions on the control process $\boldsymbol{\alpha}^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}^p$ which guarantee the finite expectation of the running maximum of $|\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}|^q$, for $q \geq 2$.

Notation. $\mathbb{E}_{\mathbf{x}}^{\alpha}$ *will denote the expectation with respect to the measure induced by the process* $\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}$ *started at* $\mathbf{x}$.

Proposition 4.8. *For each admissible pair* $(\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}^{\mathbf{x}})$, $\boldsymbol{\alpha}^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}^p$, $p \geq 2$, *the following holds.*

i) *The semimartingale* $\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}$ *admits the* canonical representation

$$\mathbf{X}_t^{\boldsymbol{\alpha}^{\mathbf{x}}} = \mathbf{x} + \mathbf{X}_t^{\boldsymbol{\alpha}^{\mathbf{x}}, c} + \int_0^t (\mathbf{u} + \boldsymbol{\mu}_s) \, \mathrm{d}s + \int_0^t \int_{\mathbb{R}_0^n} \mathbf{y} \left(\eta^{X, \boldsymbol{\alpha}^{\mathbf{x}}} - \eta \right) (\mathrm{d}s, \mathrm{d}\mathbf{y}), \tag{4.3}$$

where $\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}, c}$ *is the continuous martingale part of* $\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}$, $\eta^{X, \boldsymbol{\alpha}^{\mathbf{x}}}$ *is the jump measure associated with* $\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}$ *and* η *is its predictable compensator.*

ii) *If for each $t \geq 0$ the process $\int_0^t \|\boldsymbol{\sigma}_s\|^2 \, \mathrm{d}s \in L^{q/2}(\mathbb{P}^{\boldsymbol{\alpha}^{\mathbf{x}}})$ for $q \in [2, p]$, then there exists a constant $C > 0$ such that*

$$\mathbb{E}_{\mathbf{x}}^{\alpha} \left(\sup_{0 \leq s \leq t} |\mathbf{X}_s^{\boldsymbol{\alpha}^{\mathbf{x}}, c}|^q \right) \leq C \, \mathbb{E}_{\mathbf{x}}^{\alpha} \left(\int_0^t \|\boldsymbol{\sigma}_s\|^2 \, \mathrm{d}s \right)^{q/2} < \infty, \quad t \geq 0. \tag{4.4}$$

9

iii) *Let $\mathbf{X}^{\alpha^{\mathbf{x}},d}$ be the discontinuous martingale part of the process $\mathbf{X}^{\alpha^{\mathbf{x}}}$ and let $q \in [2,p]$. Suppose that for each $t \geq 0$*

$$\mathbf{G}_t^{\alpha^{\mathbf{x}}} := \int_0^t \mathrm{d}s \int_{\mathbb{R}_0^n} |\mathbf{y}|^2 \nu_s(\mathrm{d}\mathbf{y}) \in L^{q/2}(\mathbb{P}^{\alpha^{\mathbf{x}}}) \quad and \quad \mathbf{H}_t^{\alpha^{\mathbf{x}}} := \int_0^t \mathrm{d}s \int_{|y|>1} |\mathbf{y}|^q \nu_s(\mathrm{d}\mathbf{y}) \in L^1(\mathbb{P}^{\alpha^{\mathbf{x}}})$$

(4.5)

Then, there exists a constant $C_1 > 0$ such that

$$\mathbb{E}_{\mathbf{x}}^{\alpha}\left(\sup_{0 \leq s \leq t} |\mathbf{X}_s^{\alpha^{\mathbf{x}},d}|^q\right) \leq C_1 \, \mathbb{E}_{\mathbf{x}}^{\alpha}\left[\left(\mathbf{G}_t^{\alpha^{\mathbf{x}}}\right)^{q/2}\right] + C_1 \, \mathbb{E}_{\mathbf{x}}^{\alpha}(\mathbf{H}_t^{\alpha^{\mathbf{x}}}) < +\infty, \quad t \geq 0.$$

(4.6)

Remarks 4.9.

i) In general, the *canonical representation* of a (not necessarily special) semimartingale depends on function h associated with its characteristics (see [10, II.2c, Theorem 2.34, p.84]). However, for the special semimartingale $\mathbf{X}^{\alpha^{\mathbf{x}}}$ its representation (4.3) is independent of h. This is because of the finite first moments of the measures ν outside B_1, which allows one to compensate the *big* jumps. In fact, the representation (4.3) coincides with the *canonical decomposition* of $\mathbf{X}^{\alpha^{\mathbf{x}}}$ (by [10, II.2c, Corollary 2.38, p. 85]). Indeed, $\mathbf{X}^{\alpha^{\mathbf{x}}}$ is an $(\mathbb{F}^\alpha, \mathbb{P}^\alpha)$-*special* $\mathbb{R}^n$-valued semimartingale with compensator

$$\mathbf{A}_\cdot^{\alpha^{\mathbf{x}}} := \int_0^\cdot (\mathbf{u} + \boldsymbol{\mu}_s)\, \mathrm{d}s.$$

Also observe that, under constant policies we recover the corresponding Lévy -Itô decomposition for Lévy processes.

ii) Proposition 4.8 establishes a relationship between the existence of qth-moments of $\mathbf{X}^{\alpha^{\mathbf{x}}}$ and the finiteness of $\int_{|y|>1} |y|^q \nu(\mathrm{d}y) < +\infty$ for each measure ν. In particular, by considering constant policies we recover the well-known results for the existence of moments of Lévy processes [28, Theorem 25.3]: a Lévy process with generating triplet (b, c, F) has finite qth-moments whenever $\int_{|y|>1} |y|^q F(\mathrm{d}y) < +\infty$. That is, the finiteness of the moments depends on the tail behaviour of F (the big jumps).

iii) Since $\mathbf{X}^{\alpha^{r,\xi}}$ is a semimartingale, the Meyer-Itô formula ensures that, for any $f \in C^2(\mathbb{R}^n)$, $f(\mathbf{X}^{\alpha^{r,\xi}})$ is also a semimartingale whose decomposition can be given explicitly [25, Chapter II, Theorem 33, p. 81]. In fact, convex functions are the most general functions that take semimartingales into semimartingales [25, Chapter IV, Theorem 67, p. 215].

Remark 4.10. We can also prove that the *predictable quadratic variation* of $\mathbf{X}^{\alpha^{\mathbf{x}}} = (X^1, \ldots, X^n)^T$ is given by

$$\langle X^i, X^j \rangle_t = \int_0^t \mathrm{d}s \left((\boldsymbol{\sigma}_s^T \boldsymbol{\sigma}_s)_{ij} + \int_{\mathbb{R}_0^n} y_i y_j \nu_s(\mathrm{d}\mathbf{y}) \right), \quad i, j \in \{1, \ldots, n\};$$

(4.7)

and, thus,

$$\mathrm{Tr}\langle \mathbf{X}^{\alpha^{\mathbf{x}}}, \mathbf{X}^{\alpha^{\mathbf{x}}} \rangle_t = \int_0^t \mathrm{d}s \left(\|\boldsymbol{\sigma}_t\|^2 + \int_{\mathbb{R}_0^n} |\mathbf{y}|^2 \nu_s(\mathrm{d}\mathbf{y}) \right).$$

5. Control problem and Verification theorems

Let us now describe the cost structure of the control problem we are interested in.

5.1. **Infinite horizon case.** Consider the infinite horizon stochastic control problem with *running cost* (or *instantaneous payoff*) given by a measurable function $f : \mathbb{R}^n \times A \to \mathbb{R}^+$, and *payoff function* J defined, for each admissible pair $(\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}^{\mathbf{x}})$, by

$$J(\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}^{\mathbf{x}}) := \mathbb{E}^{\alpha}_{\mathbf{x}} \left[\int_0^\infty e^{-\gamma_t^{\boldsymbol{\alpha}^{\mathbf{x}}}} f(\mathbf{X}_t^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}_t^{\mathbf{x}}) \, \mathrm{d}t \right], \tag{5.1}$$

where the discount process $\gamma^{\boldsymbol{\alpha}^{\mathbf{x}}} = (\gamma_t^{\boldsymbol{\alpha}^{\mathbf{x}}})$ is given by $\gamma_t^{\boldsymbol{\alpha}^{\mathbf{x}}} := \int_0^t q(\mathbf{X}_s^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}_s^{\mathbf{x}}) \, \mathrm{d}s$, for some given bounded, measurable and nonnegative function $q : \mathbb{R}^n \times A \to \mathbb{R}_+$.

The aim is to minimize (5.1) over all admissible policies $\boldsymbol{\alpha}^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}^p$. In order to solve this problem, the *dynamic programming approach* focuses on studying the associated family of optimisation problems indexed by the initial data $\mathbf{x}$. The solution of the control problem consists then in finding both

(a) the *value function* (or *optimal payoff function*) $V : \mathbb{R}^n \to \mathbb{R}^+$, given by

$$V(\mathbf{x}) := \inf_{\{(\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}^{\mathbf{x}}) : \boldsymbol{\alpha}^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}^p\}} J(\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}^{\mathbf{x}}), \quad \mathbf{x} \in \mathbb{R}^n, \tag{5.2}$$

(b) an *optimal policy* for each initial state $\mathbf{x} \in \mathbb{R}^n$ (whenever it exists); that is a family $\{\hat{\boldsymbol{\alpha}}^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}^p : \mathbf{x} \in \mathbb{R}^n\}$ such that the corresponding admissible pairs $(\mathbf{X}^{\hat{\boldsymbol{\alpha}}^{\mathbf{x}}}, \hat{\boldsymbol{\alpha}}^{\mathbf{x}})$ satisfy

$$V(\mathbf{x}) = J(\mathbf{X}^{\hat{\boldsymbol{\alpha}}^{\mathbf{x}}}, \hat{\boldsymbol{\alpha}}^{\mathbf{x}}), \quad \text{for each } \mathbf{x} \in \mathbb{R}^n. \tag{5.3}$$

Notation. For convenience we will write $J^\alpha(r, \mathbf{x}) \equiv J(\mathbf{X}^{\boldsymbol{\alpha}^{r,\mathbf{x}}}, \boldsymbol{\alpha}^{r,\mathbf{x}})$ and $J^\alpha(\mathbf{x}) = J^\alpha(0, \mathbf{x})$. We will also write $(\mathbf{X}^{\boldsymbol{\alpha}}, \boldsymbol{\alpha})$ instead of $(\mathbf{X}^{\boldsymbol{\alpha}^{r,\mathbf{x}}}, \boldsymbol{\alpha}^{r,\mathbf{x}})$ whenever these processes are inside the operator $\mathbb{E}^\alpha_{r,\mathbf{x}}$.

Remark 5.1. For many applications the cost functional (5.1) is considered a suitable model to analyse the long-time behaviour of a controlled system. The corresponding finite horizon case will be treated in Section 7.1.

It is well-known that in the previous form the infinite horizon formulation is time-invariant: it does not vary over time as long as the initial state is the same. This is stated in the following lemma.

Lemma 5.2. *Define*

$$J^\alpha(r, \mathbf{x}) := \mathbb{E}^\alpha_{r,\mathbf{x}} \left[\int_r^\infty e^{-\int_t^s q(\mathbf{X}_l^{\boldsymbol{\alpha}}, \boldsymbol{\alpha}_l) \, \mathrm{d}l} f(\mathbf{X}_s^{\boldsymbol{\alpha}}, \boldsymbol{\alpha}_s) \, \mathrm{d}s \right], \tag{5.4}$$

$$v(r, \mathbf{x}) := \inf_{\boldsymbol{\alpha}^{r,\mathbf{x}} \in \mathcal{A}_{r,\mathbf{x}}^p} J^\alpha(r, \mathbf{x}). \tag{5.5}$$

Then $v(r, \mathbf{x}) = v(0, \mathbf{x}) = V(\mathbf{x})$, for all $(r, \mathbf{x}) \in \mathbb{R}_+ \times \mathbb{R}^n$.

We now prove the validity of the *dynamic programming principle* (DPP) for our control setting. To do this, we will need the following definitions.

Definition 5.3 (ϵ-optimal controls)**.** *Let $\epsilon > 0$ and $\mathbf{x} \in \mathbb{R}^n$. An admissible pair $(\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}^{\mathbf{x}})$ with $\boldsymbol{\alpha}^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}^p$ is said to be an ϵ-optimal control if*

$$J^\alpha(\mathbf{x}) < V(\mathbf{x}) + \epsilon. \tag{5.6}$$

An admissible pair $(\mathbf{X}^{\boldsymbol{\alpha}^{\xi}}, \boldsymbol{\alpha}^{\xi})$ with $\boldsymbol{\alpha}^{\xi} \in \mathcal{A}_{\xi}^p$ and $\xi \in \mathcal{P}(\mathbb{R}^n)$ is said to be an ϵ-optimal control if

$$\int_{\mathbb{R}^n} J(\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}^{\mathbf{x}}) \xi(\mathrm{d}\mathbf{x}) \leq \int_{\mathbb{R}^n} V(\mathbf{x}) \xi(\mathrm{d}\mathbf{x}) + \epsilon. \tag{5.7}$$

Remark 5.4. Notice that, by definition of the value function V, the existence of ϵ-optimal controls is always guaranteed.

11

Lemma 5.5. *(DPP) Let V be a continuous function solving (5.2)-(5.3). Then, for each $\mathbf{x} \in \mathbb{R}^n$, the function V satisfies*

$$V(\mathbf{x}) = \inf_{\boldsymbol{\alpha}^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}^p} \mathbb{E}_{\mathbf{x}}^{\alpha} \left[\int_0^t e^{-\gamma_s^{\boldsymbol{\alpha}}} f\left(\mathbf{X}_s^{\boldsymbol{\alpha}}, \boldsymbol{\alpha}_s\right) \mathrm{d}s + V\left(\mathbf{X}_t^{\boldsymbol{\alpha}}\right) e^{-\gamma_t^{\boldsymbol{\alpha}}} \right]. \tag{5.8}$$

We want to establish conditions which help us to determine if a given function $\phi : \mathbb{R}^n \to \mathbb{R}_+$ is the value function of the control problem (5.1)-(5.3). The following lemma provides a *necessary* condition satisfied by the optimal cost function V. This result is a consequence of the DPP.

Lemma 5.6. *Let $V : \mathbb{R}^n \to \mathbb{R}^+$ be the value function of the control problem (5.1)-(5.3). For each admissible policy $\boldsymbol{\alpha}^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}^p$, $p \geq 2$, such that $J^{\alpha}(\mathbf{x}) < +\infty$, the process $S^{V,\boldsymbol{\alpha}^{\mathbf{x}}}$ defined by*

$$S_t^{V,\boldsymbol{\alpha}^{\mathbf{x}}} := \int_0^t e^{-\gamma_s^{\boldsymbol{\alpha}^{\mathbf{x}}}} f(\mathbf{X}_s^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}_s^{\mathbf{x}}) \, \mathrm{d}s + e^{-\gamma_t^{\boldsymbol{\alpha}^{\mathbf{x}}}} V(\mathbf{X}_t^{\boldsymbol{\alpha}^{\mathbf{x}}}), \quad t \geq 0 \tag{5.9}$$

is a positive $\mathbb{P}^{\alpha}$-submartingale. Furthermore, if $\boldsymbol{\alpha}^{\mathbf{x}}$ is optimal, then $S^{V,\boldsymbol{\alpha}^{\mathbf{x}}}$ is a (true) $\mathbb{P}^{\alpha}$-martingale.

Remarks 5.7.

i) The process $S^{V,\boldsymbol{\alpha}^{\mathbf{x}}}$ is called *the Bellman process*. At each time t, it can be thought of as the minimum expected total cost, given the evolution of the process up to time t under the policy $\boldsymbol{\alpha}^{\mathbf{x}}$ and then changing to an optimal control afterwards. The submartingale property tells us then that by using an arbitrary control $\boldsymbol{\alpha}^{\mathbf{x}}$ for a longer time, the expected cost keeps on increasing and such an increase is zero whenever the policy is optimal.

ii) Lemma 5.6 provides a first necessary (but not sufficient) condition to characterise the value function. As in the standard diffusion case, some sufficient conditions can be given in terms of the so-called *transversality condition* (see (5.10) below).

iii) Notice that policies with infinite payoff were not excluded from the definition of admissible policies. However, when solving the optimisation problem we will only focus on policies with finite payoff, as otherwise it is clear that such a policy cannot be an optimal one.

Let us give some further assumptions for a given candidate value function $\phi \in C^2(\mathbb{R}^n)$:

(**SC**) For any admissible policy $\boldsymbol{\alpha}^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}^p$ with $J^{\alpha}(\mathbf{x}) < +\infty$, $S_t^{\phi,\boldsymbol{\alpha}^{\mathbf{x}}}$ is a positive submartingale.

(**MC**) There exists a family of admissible pairs $\{ (\mathbf{X}^{\hat{\boldsymbol{\alpha}}^{\mathbf{x}}}, \hat{\boldsymbol{\alpha}}^{\mathbf{x}}) : \mathbf{x} \in \mathbb{R}^n \}$ such that, for each $\hat{\boldsymbol{\alpha}}^{\mathbf{x}}$, $S_t^{\phi,\hat{\boldsymbol{\alpha}}^{\mathbf{x}}}$ is a martingale.

(**TC**) Each admissible policy $\boldsymbol{\alpha}^{\mathbf{x}}$ is such that either $J^{\alpha}(\mathbf{x}) = \infty$ or

$$\liminf_{t \to \infty} \mathbb{E}_{\mathbf{x}}^{\alpha} \left[e^{-\gamma_t^{\boldsymbol{\alpha}}} \phi\left(\mathbf{X}_t^{\boldsymbol{\alpha}}\right) \right] = 0, \tag{5.10}$$

(**nC**) For each $\mathbf{x} \in \mathbb{R}^n$, there exists a sequence $\{\boldsymbol{\alpha}_n^{\mathbf{x}}\}_{n \in \mathbb{N}} \subset \mathcal{A}_{\mathbf{x}}^p$ such that $J^{\alpha_n}(\mathbf{x}) < \phi(\mathbf{x}) + \frac{1}{n}$.

Remarks 5.8.

i) The submartingale and martingale conditions (**SC**)-(**MC**) are the core feature in the standard *martingale approach* for stochastic control problems (see [4]). By Lemma (5.6), these conditions are only necessary conditions for optimality.

ii) The *transversality condition* (5.10) implicitly prescribes some kind of growth condition on V. Essentially, it ensures that the value function $V(\mathbf{x})$ does not growth too rapidly for large $|\mathbf{x}|$. This condition plays a similar role than a terminal condition does in the finite horizon case.

Remark 5.9. If our control problem were a maximisation problem then the condition "*positive submartingale*" in (**SC**) would be replaced by "*positive supermartingale*", whereas the inequalities (7.10)-(7.11) would be reversed.

Lemma 5.10 (**Verification Result 1**). *Let $p \geq 2$ and let $\phi \in C^2(\mathbb{R}^n)$ be a nonnegative function satisfying (**SC**) and (**TC**). Then the following holds.*

  i) *For any admissible policy $\boldsymbol{\alpha}^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}^p$, $\phi(\mathbf{x}) \leq J^\alpha(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^n$.*

 ii) *If condition (**MC**) holds, then $\phi(\mathbf{x}) = J^{\hat{\alpha}}(\mathbf{x})$ and, thus, ϕ is the value function of the control problem (5.1)-(5.3) and $\{\hat{\boldsymbol{\alpha}}^{\mathbf{x}} : \mathbf{x} \in \mathbb{R}^n\}$ is a family of optimal admissible policies.*

iii) *If condition (**nC**) holds, then ϕ is the value function and, for each n, $\{\boldsymbol{\alpha}_n^{\mathbf{x}} : \mathbf{x} \in \mathbb{R}^n\}$ is a family of ϵ_n-optimal admissible policies with $\epsilon_n := \frac{1}{n}$.*

Remarks 5.11.

  i) Under conditions (**SC**) and (**TC**), Lemma 5.10 implies that the candidate value function ϕ is a lower bound for the value function V. However, to apply this lemma one needs to verify both conditions *for each policy $\boldsymbol{\alpha}^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}^p$.*

 ii) The transversality condition (**TC**) is not much of a problem in the maximisation case with nonnegative running cost as such a condition becomes

$$\limsup_{t \to \infty} e^{-qt} \mathbb{E}\left[\phi\left(X_t^x\right)\right] \geq 0.$$

iii) Condition (**SC**) is usually dealt with via the corresponding **HJB** equation related to the so-called verifications theorems.

Verification theorems provide sufficient conditions for optimality when the value function is smooth enough. Such conditions are given in terms of the solution to an integro-differential equation and require the concept of *Markov policies.*

Definition 5.12 (**Markov controls**). *Let $(\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}^{\mathbf{x}})$ be an admissible pair with $\boldsymbol{\alpha}^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}^p$, $p \geq 2$. The process $\boldsymbol{\alpha}^{\mathbf{x}} = (\sigma, \nu, \mu)$ is called a* Markov control *if there exists an $\mathbb{F}^\alpha$-adapted càdlàg process $X^{\mathbf{x}}$ starting at $\mathbf{x}$ and there exist measurable mappings $\bar{\mu} : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$, $\bar{\sigma} : \mathbb{R}_+ \times \mathbb{R}^n \to M_{n \times n}(\mathbb{R})$ and $\bar{\nu} : \mathbb{R}_+ \times \mathbb{R}^n \to \mathcal{M}_p$ such that $\mathbb{P}^\alpha - a.s.$ $\mu_s = \bar{\mu}(s, X_s^{\mathbf{x}})$, $\sigma_s = \bar{\sigma}(s, X_s^{\mathbf{x}})$ and $\nu_s = \bar{\nu}(s, X_s^{\mathbf{x}})$, for all $s \geq 0$, and, furthermore, if $\bar{\alpha} := (\bar{\sigma}, \bar{\nu}, \bar{\mu})$ then $(\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}_{\cdot}, \boldsymbol{\alpha}^{\mathbf{x}}_{\cdot}) = (X^{\mathbf{x}}_{\cdot}, \bar{\alpha}^{\mathbf{x}}(\cdot, X^{\mathbf{x}}_{\cdot}))$ have the same law. If all the mappings in $\bar{\alpha}$ are independent of time, then the corresponding control is said to be a* stationary Markov control.

To state the verification theorem of our control problem, we introduce the following conditions:

(**HJB**) ϕ satisfies the integro-differential equation

$$\inf_{\mathbf{a} \in A}\{L^{\mathbf{a}}\phi(\mathbf{x}) - q(\mathbf{x}, \mathbf{a})\phi(\mathbf{x}) + f(\mathbf{x}, \mathbf{a})\} = 0, \quad \text{for all } \mathbf{x} \in \mathbb{R}^n. \tag{5.11}$$

(**UI**) For each $\mathbf{x} \in \mathbb{R}^n$ and for each admissible policy $\boldsymbol{\alpha}^{\mathbf{x}}$, the family $\{e^{-\gamma_{t \wedge \tau}^{\boldsymbol{\alpha}^{\mathbf{x}}}}\phi\left(\mathbf{X}_{\tau \wedge t}^{\boldsymbol{\alpha}^{\mathbf{x}}}\right)\}_{\tau \in \mathcal{T}}$ is uniformly integrable (UI) for each $t \geq 0$. Here $\mathcal{T}$ denotes the family of all stopping times.

(**OC**) There exists an admissible policy $\hat{\boldsymbol{\alpha}}^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}^p$ with corresponding controlled process $\mathbf{X}^{\hat{\alpha}^{\mathbf{x}}}$, defined on a filtered probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{F}} := (\hat{\mathcal{F}}_t), \hat{\mathbb{P}})$, such that

$$L^{\hat{\alpha}_s^{\mathbf{x}}}\phi(\mathbf{X}_s^{\hat{\alpha}^{\mathbf{x}}}) - q(\mathbf{X}_s^{\hat{\alpha}^{\mathbf{x}}}, \hat{\alpha}_s^{\mathbf{x}})\phi(\mathbf{X}_s^{\hat{\alpha}^{\mathbf{x}}}) + f(\mathbf{X}_s^{\hat{\alpha}^{\mathbf{x}}}, \hat{\alpha}_s^{\mathbf{x}}) = 0, \quad \hat{\mathbb{P}} - a.s. \quad \text{for all } s \geq 0, \tag{5.12}$$

A few comments about the previous assumptions:

(i) Equations of the type (7.6) are usually referred to as *Hamilton-Jacobi-Bellman (HJB) equations* (or the Dynamic Programming Equation) and they are the infinitesimal version of the dynamic programming principle. By assuming sufficient regularity conditions for the value function, this equation can be derived formally by a standard limiting procedure.

ii) Assumption (**UI**) is important for the case when the Bellman process $S^{\phi,\boldsymbol{\alpha}^{\mathbf{x}}}$ is only a *local* submartingale. This condition implies that the nonnegative local submartingale $S^{\phi,\boldsymbol{\alpha}^{\mathbf{x}}}$ is of class (DL) and thus it is a true submartingale.

iii) Assumption (**OC**) allows one to identify and construct an optimal (stationary) Markov control policy via a pointwise minimisation of the associated HJB equation.

The main result of this paper is the following verification theorem. This theorem characterises the value function V as a solution to the integro-differential equation (7.6) and it also identifies optimal Markov controls.

Theorem 5.13 (Verification Result 2). *Let $p \geq 2$ and suppose that $\phi \in C^2(\mathbb{R}^n)$ is a nonnegative function satisfying $|\phi(\mathbf{x})| \leq C(|\mathbf{x}|^q + 1)$ for some $C > 0$ and $q \in [2, p]$. Under conditions (**HJB**), (**UI**) and (**TC**), the following holds.*

i) *For any admissible policy $\boldsymbol{\alpha}^{\mathbf{x}} \in \mathcal{A}^p_{\mathbf{x}}$, $\phi(\mathbf{x}) \leq J^\alpha(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^n$.*

ii) *If, additionally, condition (**OC**) holds, then $\phi(\mathbf{x}) = J^{\hat{\alpha}}(\mathbf{x})$, the family $\{\hat{\boldsymbol{\alpha}}^{\mathbf{x}} : \mathbf{x} \in \mathbb{R}^n\}$ is a family of optimal (stationary) Markov policies, and $\phi(\mathbf{x})$ is the value function of the control problem* (5.1)-(5.3).

Remark 5.14.

i) The proof of Theorem 5.13 is given in the appendix B.5 and follows a localised version of the standard *martingale approach* for stochastic control problems. Therein we first show that the Bellman process $S^{\phi,\boldsymbol{\alpha}^{\mathbf{x}}}$ is a *local* submartingale for any arbitrary admissible policy and, further, it is a *local* martingale whenever the policy is optimal. Then, thanks to condition (**UI**) we conclude the proof by standard localising arguments.

ii) Since $\phi \in C^2$, we are seeking smooth solutions to the HJB equation. This regularity also justifies the use of the Itô-Meyer formula in the corresponding proof. However, by Theorem 71 in [25, Chapter IV, p. 221], the smoothness condition can be relaxed, at least for the one-dimensional case, by considering $\phi \in C^1$ with an absolutely continuous derivative f'.

iii) For a maximisation problem the *infimum* in (7.6) should be replaced by a *supremum*, so that statement i) in Theorem 5.13 would imply that ϕ is an upper bound for the value function, that is $\phi(\mathbf{x}) \geq J^\alpha(\mathbf{x})$, for each admissible policy.

Remark 5.15. It is worth recalling the importance of (**TC**) at prescribing growth conditions for the candidate value function ϕ. In the diffusion setting, a standard example is the following (see [7, Example 3.1, p. 130]): consider the equation $\frac{1}{2}\phi''(x) - q\phi(x) + f(x) = 0$, whose general solution ϕ is given by

$$\phi(x) = x^2 + 1 + c_1 \exp(\sqrt{2}x) + c_2 \exp(-\sqrt{2}x). \tag{5.13}$$

It can be proved that the corresponding transversality condition $\lim_{t\to\infty} e^{-qt} \mathbb{E}_x(\phi(B^x_t)) = 0$ is satisfied only when $c_1 = 0 = c_2$. Here B^x is a standard Brownian motion started at x. On the other hand, the other solutions given by (5.13) grow exponentially as $x \to \infty$ or $x \to -\infty$.

Remark 5.16. For any fixed $a \in A$, let $Y^{x,a}$ be an $\mathbb{R}^n$-valued Lévy process started at x whose infinitesimal generator coincides with the operator L^a on $C^2_c(\mathbb{R}^n)$. If $\phi \in C^2(\mathbb{R}^n)$ has polynomial growth of degree $p \geq 2$ and, further, ϕ solves the integro-differential equation

$$L^a\phi(x) - q(x,a)\phi(x) + f(x,a) = 0, \quad x \in \mathbb{R}^n, \tag{5.14}$$

$$\lim_{t\to\infty} \mathbb{E}\left[e^{-\int_0^t q(X^x_r,a)\,\mathrm{d}r} f(X^x_t,a)\right] = 0, \quad x \in \mathbb{R}^n, \tag{5.15}$$

then, by Theorem 5.13, such a solution admits the probabilistic representation:

$$\phi(\mathbf{x}) = \mathbb{E}\left[\int_0^\infty e^{-\int_0^t -q(X^x_r,a))\,\mathrm{d}r} f(X^x_t,a)\,\mathrm{d}t\right].$$

In general, it can be difficult to verify the validity of conditions (**UI**) and (**TC**). However, if the function ϕ is bounded, both conditions follow straightforwardly. In particular, if the running cost function f is bounded, then the value function is also bounded and so is any candidate value function ϕ. We thus obtain the following result for the bounded case.

Corollary 5.17. (**Bounded case**) *Suppose that $f : \mathbb{R}^n \times A \to \mathbb{R}_+$ is a measurable bounded function. Let $\phi \in C^2(\mathbb{R}^n)$ be a bounded nonnegative function satisfying condition (**HJB**), then the conclusions of Theorem 5.13 are valid.*

Thanks to the finiteness of the pth moment of each measure ν outside the ball B_1, the growth condition of ϕ in Theorem 5.13 ensures that the non-local term in $L^{\mathbf{a}}\phi$ (and so equation (7.6)) is well-defined. We are now interested in providing conditions on the running cost function f that guarantee that the corresponding value function is of polynomial growth.

Lemma 5.18. *Consider the control problem (5.1)-(5.3) with admissible policies $\mathcal{A}_{\mathbf{x}}^p$ for a fixed $p \geq 2$. If there exist a positive constant $c > 0$, an action control $\mathbf{a}_0 \in A$ and $q \in [2, p]$ such that the running cost $f : \mathbb{R}^n \times A \to \mathbb{R}_+$ satisfies $|f(\mathbf{x}, \mathbf{a}_0)| \leq c(1 + |\mathbf{x}|^q)$ for all $\mathbf{x} \in \mathbb{R}^n$, then there exists a positive constant $C > 0$ such that the value function $V : \mathbb{R}^n \to \mathbb{R}^n$ satisfies $V(\mathbf{x}) \leq C(1 + |\mathbf{x}|^q)$, for all $\mathbf{x} \in \mathbb{R}^n$.*

Remarks 5.19.

i) The key fact in the proof of this lemma (see section B.6) is to show that for the constant policy $\boldsymbol{\alpha}^{\mathbf{x}} = \mathbf{a}_0$, with associated controlled process $\mathbf{X}^{\mathbf{a}^{\mathbf{x}}}$, the following holds

$$\int_0^\infty e^{-\gamma_s^{\mathbf{a}^{\mathbf{x}}}} \mathbb{E}_{\mathbf{x}} |\mathbf{X}_t^{\mathbf{a}^{\mathbf{x}}}|^q \, \mathrm{d}t \leq C(1 + |\mathbf{x}|^q).$$

Observe that the latter inequality is valid as the process $\mathbf{X}^{\mathbf{a}^{\mathbf{x}}}$ is a *Lévy martingale with drift* whose qth-moments estimates are well-known and satisfy the inequality above.

ii) In the Itô diffusion setting another standard condition for the running cost function f is the (uniform) growth condition $|f(\mathbf{x}, \mathbf{a})| \leq C(1 + |\mathbf{x}^q|)$. In our framework, such a case is covered by Lemma 5.18 which also implies that the value function V has polynomial growth of degree at most q.

iii) It is not difficult to see that to allow a value function with exponential growth one needs, in principle, an exponential moment for the measures ν outside the unitary ball B_1. Appropriate assumptions on the control $\boldsymbol{\alpha}^{\mathbf{x}}$ are required to guarantee the finite expectation of the exponential moments of $\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}$. The latter will then ensure the validity of the corresponding transversality condition.

We now state conditions on the running cost function f that imply (**TC**).

Lemma 5.20. *Consider the control problem (5.1)-(5.3) with admissible policies $\mathcal{A}_{\mathbf{x}}^p$ for a fixed $p \geq 2$. Let $f : \mathbb{R}^n \times A \to \mathbb{R}_+$ be the running cost function. If there exists a positive constant $c > 0$ such that $|f(\mathbf{x}, \mathbf{a})| \geq c(1 + |\mathbf{x}|^p)$ for all $\mathbf{a} \in A$, then the transversality condition (5.10) holds for any nonnegative function $\phi \in C^2(\mathbb{R}^n)$ with polynomial growth of degree p.*

As a direct consequence of Lemma 5.18 and Lemma 5.20, we can now give conditions on the running cost f under which the value function is of polynomial growth and, further, satisfies the transversality condition.

Corollary 5.21. *Consider the control problem (5.1)-(5.3) with admissible policies $\mathcal{A}_{\mathbf{x}}^p$ for a fixed $p \geq 2$. Suppose that there exist $C > 0$ and $\mathbf{a}_0 \in A$ such that the running cost $f : \mathbb{R}^n \times A \to \mathbb{R}_+$ satisfies $|f(\mathbf{x}, \mathbf{a}_0)| \leq C(1 + |\mathbf{x}|^p)$ for all $\mathbf{x} \in \mathbb{R}^n$ and, further, there exists $c > 0$ such that $|f(\mathbf{x}, \mathbf{a})| \geq c(1 + |\mathbf{x}|^p)$ for all $\mathbf{a} \in A$. Then the value function V is of polynomial growth of degree p and satisfies (**TC**).*

Remark 5.22. In particular, this corollary is valid when f is of the same order than a polynomial function of degree $p \geq 2$. Recall that the parameter p is related to the moments of each ν outside of the ball B_1 (see (2.1)).

6. Different classes of admissible controls

Apart from the case where the running cost function is bounded, we have not explored yet conditions that guarantee the validity of (**UI**). This assumption, as was pointed out before, is not easy to verify in practical applications. To fill in this gap, we provide the following three classes of admissible controls for which condition (**UI**) is no longer needed in the corresponding verification theorems.

Case 1. Integrability conditions on the control processes $\boldsymbol{\alpha}^{\mathbf{x}}$.

Definition 6.1. *An admissible policy $\boldsymbol{\alpha}^{\mathbf{x}} = (\boldsymbol{\sigma}, \nu, \boldsymbol{\mu}) \in \mathcal{A}^p_{\mathbf{x}}$ is said to belong to the class $\tilde{\mathcal{A}}^p_{\mathbf{x}}(q)$, $q \in [2, p]$, whenever it satisfies the additional condition:*

*(**H3**) For all $t \in \mathbb{R}_+$,*

$$\int_0^t |\boldsymbol{\mu}_s| + \|\boldsymbol{\sigma}_s\|^2 \, \mathrm{d}s \in L^{q/2}(\mathbb{P}^{\boldsymbol{\alpha}^{\mathbf{x}}}), \quad \mathbf{G}_t^{\boldsymbol{\alpha}^{\mathbf{x}}} \in L^{q/2}(\mathbb{P}^{\boldsymbol{\alpha}^{\mathbf{x}}}), \quad and \quad \mathbf{H}_t^{\boldsymbol{\alpha}^{\mathbf{x}}} \in L^1(\mathbb{P}^{\boldsymbol{\alpha}^{\mathbf{x}}}), \qquad (6.1)$$

where $\mathbf{G}_t^{\boldsymbol{\alpha}^{\mathbf{x}}}$ and $\mathbf{H}_t^{\boldsymbol{\alpha}^{\mathbf{x}}}$ are as defined in (4.5).

Theorem 6.2 (Verification Result 3). *Let $p \geq 2$ and $q \in [2, p]$. If $\phi \in C^2(\mathbb{R}^n)$ is a nonnegative function bounded by a polynomial function of order q for which both conditions (**TC**) and (**HJB**) hold, then the conclusions of* Theorem 5.13 *are also valid for the corresponding minimisation problem over the set $\tilde{\mathcal{A}}_{\mathbf{x}}(q)$.*

Remark 6.3. Condition (**H3**) seems to be quite restrictive, however, we will see that due to the generality of our set-up is fairly natural to impose integrability conditions of this type. We will also see that in the standard SDE framework the corresponding integrability assumptions are implied by the very-well known Itô conditions.

Remark 6.4. Thanks to the assumptions on the growth of the candidate function ϕ, the proof of Theorem 6.2 follows easily by observing that (6.1) guarantees the finiteness of the expectation of the running maximum of $\left|\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}\right|^q$.

Case 2. Growth conditions on the control processes $\boldsymbol{\alpha}^{\mathbf{x}}$.

Definition 6.5. *An admissible policy $\boldsymbol{\alpha}^{\mathbf{x}} = (\boldsymbol{\sigma}, \nu, \boldsymbol{\mu}) \in \mathcal{A}^p_{\mathbf{x}}$ is said to belong to the class $\hat{\mathcal{A}}^p_{\mathbf{x}}$, $p \geq 2$, if the following conditions hold:*

*(**H4**) There exist measurable functions*

$$\hat{\mu}^i : \mathbb{R}^n \times \Omega \to \mathbb{R}^n, \quad 1 \leq i \leq n, \qquad \hat{\sigma}_{ij} : \mathbb{R}^n \times \Omega \to M_{n \times n}(\mathbb{R}), \quad 1 \leq i, j \leq n,$$
$$\hat{\nu} : \mathbb{R}^n \times \Omega \to \mathcal{M}_p,$$

such that, for each $\hat{g} \in \{\hat{\sigma}, \hat{\nu}, \hat{\mu}\}$, $\hat{g}\left(\mathbf{X}_t^{\boldsymbol{\alpha}^{\mathbf{x}}}, \cdot\right)$ is $\mathcal{F}_t^{\alpha}$-adapted and, further,

$$\boldsymbol{\alpha}_t^{\mathbf{x}} = (\hat{\sigma}(\mathbf{X}_t^{\boldsymbol{\alpha}^{\mathbf{x}}}(\omega), \omega), \hat{\nu}(\mathbf{X}_t^{\boldsymbol{\alpha}^{\mathbf{x}}}(\omega), \omega), \hat{\mu}(\mathbf{X}_t^{\boldsymbol{\alpha}^{\mathbf{x}}}(\omega), \omega)), \quad \mathbb{P}^{\alpha} - a.s..$$

*(**GC**) There exist a deterministic positive constant K and a real-valued process $\kappa = (\kappa_t)_{t \geq 0}$ such that for all $t \geq 0$ and $\mathbf{x} \in \mathbb{R}^n$*

$$|\hat{\mu}(\mathbf{x}, \omega)|^p + \|\hat{\sigma}(\mathbf{x}, \omega)\|^p + \int_{\mathbb{R}_0^n} |\mathbf{z}|^2 \vee |\mathbf{z}|^p \hat{\nu}(\mathbf{x}, \omega, \mathrm{d}\mathbf{z}) \leq |\kappa_t(\omega)|^p + K|\mathbf{x}|^p, \; \mathbb{P}^{\alpha} - a.s. \qquad (6.2)$$

with $\int_0^t |\kappa_s|^p \, \mathrm{d}s \in L^{p/2}(\mathbb{P}^{\alpha})$.

Remark 6.6. The previous definition of admissible controls can be thought of as a generalisation of the Itô SDE setting as presented in [22]. In this reference, under an additional Lipschitz condition, the process κ_t can be given explicitly in terms of the associated drift and diffusion (stochastic) coefficients, see [22, Section 1.3, p. 23]).

Theorem 6.7 (Verification Result 4). *Let $p \geq 2$ and $q \in [2, p]$. If $\phi \in C^2(\mathbb{R}^n)$ is a non-negative function bounded by a polynomial function of order q, for which both conditions (**TC**) and (**HJB**) hold, then the conclusions of Theorem 5.13 are also valid for the corresponding minimisation problem over the set of admissible controls $\hat{\mathcal{A}}_{\mathbf{x}}^p$.*

Remarks 6.8. The key fact in the proof of Theorem 6.7 is to show that, for each admissible policy $\boldsymbol{\alpha}^{\mathbf{x}} \in \hat{\mathcal{A}}_{\mathbf{x}}^p$, $\mathbb{E}_{\mathbf{x}}^\alpha [\sup_{0 < s \leq t} |\mathbf{X}_s^\alpha|^p] < +\infty$, for each $t \geq 0$. The latter follows directly from the estimates given in Proposition 4.8. Notice that, apart from the growth condition (6.3), the assumption $\int_0^t |\kappa_s|^p \, \mathrm{d}s \in L^{p/2}(\mathbb{P}^\alpha)$ is crucial. The rest of the proof follows the same arguments used in the proof of Theorem 6.2, so that we omit the details.

Case 3. Markov conditions on the control processes $\boldsymbol{\alpha}^{\mathbf{x}}$.

As a particular case of the class of admissible controls $\hat{\mathcal{A}}_{\mathbf{x}}^{\mathbf{x}}$ given in Definition 6.5, we can now restrict our attention to the class of stationary Markov controls, i.e. when $\boldsymbol{\alpha}^{\mathbf{x}}$ is of the form

$$\boldsymbol{\alpha}_s^{\mathbf{x}}(\omega) = (\hat{\sigma}(\mathbf{X}_s^{\boldsymbol{\alpha}^{\mathbf{x}}}(\omega)), \hat{\nu}(\mathbf{X}_s^{\boldsymbol{\alpha}^{\mathbf{x}}}(\omega)), \hat{\mu}(\mathbf{X}_s^{\boldsymbol{\alpha}^{\mathbf{x}}}(\omega))),$$

for some (deterministic) measurable functions $\hat{\sigma}$, $\hat{\nu}$ and $\hat{\mu}$ (recall Definition 5.12).

Definition 6.9. *An admissible policy $\boldsymbol{\alpha}^{\mathbf{x}} = (\boldsymbol{\sigma}, \nu, \boldsymbol{\mu}) \in \mathcal{A}_{\mathbf{x}}^p$ is said to belong to the class $\mathcal{A}_{\mathbf{x}}^{M,p}$, $p \geq 2$, if $\boldsymbol{\alpha}^{\mathbf{x}}$ is a stationary Markov control process and the following condition is satisfied:*

(GM) *If $\boldsymbol{\alpha}^{\mathbf{x}} = (\hat{\sigma}(\mathbf{X}_s^{\boldsymbol{\alpha}^{\mathbf{x}}}(\omega)), \hat{\nu}(\mathbf{X}_s^{\boldsymbol{\alpha}^{\mathbf{x}}}(\omega)), \hat{\mu}(\mathbf{X}_s^{\boldsymbol{\alpha}^{\mathbf{x}}}(\omega)))$, then there exists a deterministic positive constant K such that for all $t \geq 0$ and $\mathbf{x} \in \mathbb{R}^n$, the functions $\hat{\sigma}$, $\hat{\nu}$ and $\hat{\mu}$ satisfy*

$$|\hat{\mu}(\mathbf{x})|^p + \|\hat{\sigma}(\mathbf{x})\|^p + \int_{\mathbb{R}_0^n} |\mathbf{z}|^2 \vee |\mathbf{z}|^p \hat{\nu}(\mathbf{x}, \mathrm{d}\mathbf{z}) \quad \leq \quad K(1 + |\mathbf{x}|^p). \tag{6.3}$$

Theorem 6.10 (Verification Result 5). *Let $p \geq 2$ and $q \in [2, p]$. If $\phi \in C^2(\mathbb{R}^n)$ is a non-negative function bounded by a polynomial function of order q, for which both conditions (**TC**) and (**HJB**) hold, then the conclusions of Theorem 5.13 are also valid for the corresponding minimisation problem over the set of admissible (stationary) Markov controls $\hat{\mathcal{A}}_{\mathbf{x}}^{M,p}$.*

Remark 6.11. Since this result is just a particular case of Theorem 6.7, we omit its proof.

6.1. **About the standard SDE settings.** Most of the literature dealing with optimal control of Lévy -Itô diffusions share two important characteristics in the definition of admissible policies:

a) The optimisation is usually done over the smaller class of Markov controls u. Hence, each process X^u is defined as a solution to a controlled SDE with a Markovian structure:

$$X_t^u = x + \int_0^t b(X_r^u, u_r) \, \mathrm{d}r + \int_0^t s(X_r^u, u_r) \, \mathrm{d}W_r +$$
$$+ \int_0^t \int_{|\gamma| \leq 1} \gamma(X_{r-}^u, z, u_r) \tilde{N}(\mathrm{d}r, \mathrm{d}z) + \int_0^t \int_{|\gamma| > 1} \gamma(X_{r-}^u, z, u_{r-}) N(\mathrm{d}r, \mathrm{d}z), \quad (6.4)$$

i.e., the drift and diffusion coefficients b and s, as well as the function γ (which determines the size of the jumps), are assumed to be functions of both the space and the control variable. Here $\tilde{N}(\mathrm{d}s, \mathrm{d}z) := N(\mathrm{d}s, \mathrm{d}z) - \mathrm{d}s\nu(\mathrm{d}z)$ is a compensated Poisson random measure (independent of the Brownian motion W), whose mean measure ν is fixed and satisfies that $\int_{\mathbb{R}_0^n} 1 \wedge \gamma^2(\cdot, z, \cdot)\nu(\mathrm{d}z) < \infty$.

b) The drift and diffusion coefficients as well as γ satisfy appropriate Itô -type conditions: Lipschitz and linear growth conditions in the space variable and uniformly on the control variable.

The importance of such conditions is that they guarantee both (i) the existence and uniqueness of a strong solution to the corresponding SDE, and (ii) the square integrability of the associated control process (see [8, Chapter 4], [22, Section 1.3, p.22] for the (continuous) Itô diffusion case, or [20, Theorem 1.19, p. 10] for the (jump) Lévy diffusion case).

Remarks 6.12.

i) Although in our setting condition (**H3**) seems to be quite strong when compared to the Itô or Lévy diffusion cases, this assumption is crucial to guarantee the finiteness of the qth moments of each process $\mathbf{X}^{\alpha^{\mathbf{x}}}$. Nevertheless, by Theorem 6.7, the qth moments are also finite whenever the growth condition (6.3) holds.

ii) By imposing an additional Lipschitz condition we can also guarantee the existence of a strong solution to the corresponding SDE. We then see that our framework encompasses the controlled Itô SDE case.

iii) Using Definition 6.5, we can recover the Itô diffusion case by setting $\nu \equiv 0$ and by proceeding as follows: given an admissible control $\alpha_t^x(\omega) := (\bar{\sigma}, \bar{\mu})(\cdot, \omega)$, where $\bar{\sigma} : \mathbb{R}^n \times \Omega \to M_{n \times n}(\mathbb{R})$ and $\bar{\mu} : \mathbb{R}^n \times \Omega \to \mathbb{R}^n$ satisfy (**H4**), define $s(\cdot, \alpha_t(\omega)) := \bar{\sigma}(\cdot, \omega)$ and $b(\cdot, \alpha_t(\omega)) := \bar{\mu}(\cdot, \omega)$ as the corresponding diffusion and drift coefficients.

iv) In the Lévy diffusion (6.4), the jump intensity measure ν is fixed and deterministic. Hence, the control affects only the jump sizes determined by the function $\gamma(x, z, u)$. Since our martingale approach allows us to control the jump intensity measure, our setting is more general than the jump case in [20].

v) We can reformulate the controlled SDE (6.4) in terms of our martingale approach as follows. Define the action control set Γ' as a subset of $M_{n \times n}(\mathbb{R}) \times \mathbb{R}^n \times F(\mathbb{R}^n; \mathbb{R}^n)$, where $F(\mathbb{R}^n; \mathbb{R}^n)$ is the set of measurable functions on $\mathbb{R}^n$ with values on $\mathbb{R}^n$. Fixed a jump intensity measure ν and replace the operator $L^{\mathbf{a}}$ in (3.1) by the operator

$$(\mathcal{G}^{\mathbf{a}}h)(\cdot) := \boldsymbol{\mu}^T \nabla h(\cdot) + \frac{1}{2} \operatorname{Tr}(\boldsymbol{\sigma}^T \operatorname{H}h \, \boldsymbol{\sigma})(\cdot) + \int_{\mathbb{R}_0^n} \left(h(\cdot + \mathbf{y}) - h(\cdot) - \mathbf{y}^T \nabla h(\cdot) \right) \mathrm{d}\nu \circ \theta^{-1}(\mathbf{y}), \quad (6.5)$$

where each action $\mathbf{a} = (\boldsymbol{\sigma}, \boldsymbol{\mu}, \theta) \in \Gamma'$. The set of admissible controls can now be defined (with the appropriate changes) as was done in Section 3.1.

7. Finite horizon case and other extensions

7.1. Finite Horizon Case. We will sketch briefly the formulation for the case when the planning horizon is a finite (deterministic) time interval $[0, T] \subset \mathbb{R}_+$. We take the control setting of Section 3.1 except for the fact that all processes are defined on the time interval $[0, T]$. Keeping this restriction in mind, we shall use the same notation for the class of admissible controls (recall also Remark 4.3). Let $p \geq 2$ be fixed throughout this section. For each $\mathbf{z} \in \mathbb{R}^n$, define the cost functional

$$\mathcal{J} : (\mathbf{X}^{\alpha^{\mathbf{z}}}, \boldsymbol{\alpha}^{\mathbf{z}}) \mapsto \mathbb{E}_{\mathbf{z}}\left[\int_0^T e^{-\gamma_t^{\alpha^{\mathbf{z}}}} f(t, \mathbf{X}_t^{\alpha^{\mathbf{z}}}, \boldsymbol{\alpha}_t^{\mathbf{z}}) \, \mathrm{d}t + g(\mathbf{X}_T^{\alpha^{\mathbf{z}}}) \right], \quad (7.1)$$

where $f : [0, T] \times \mathbb{R}^n \times A \to \mathbb{R}^+$ and $h : \mathbb{R}^n \to \mathbb{R}^+$ are the *running cost* and the *terminal cost* functions, respectively. As before, the discount process is given by $\gamma_\cdot^{\alpha^{\mathbf{z}}} := \int_0^\cdot q(s, \mathbf{X}_s^{\alpha^{\mathbf{z}}}, \boldsymbol{\alpha}_s^{\mathbf{z}}) \, \mathrm{d}s$ with $q : [0, T] \times \mathbb{R}^n \times A \to \mathbb{R}_+$ being a bounded measurable function.

The stochastic control problem on the interval $[0, T]$ consists in solving the optimisation problem:

$$\inf_{(\mathbf{X}^{\alpha^{\mathbf{z}}}, \boldsymbol{\alpha}^{\mathbf{z}}) : \boldsymbol{\alpha}^{\mathbf{z}} \in \mathcal{A}_{\mathbf{z}}^p} \mathcal{J}(\mathbf{X}^{\alpha^{\mathbf{z}}}, \boldsymbol{\alpha}^{\mathbf{z}}) \quad (7.2)$$

for some initial state $\mathbf{z} \in \mathbb{R}^n$.

The solution to (7.1)-(7.2) consists then in finding the optimal value of $\mathcal{J}$ and an optimal policy (whenever it exists). A policy $\hat{\boldsymbol{\alpha}}^{\mathbf{z}} \in \mathcal{A}_{\mathbf{z}}^p$ is called *optimal* if the corresponding admissible pair $(\mathbf{X}^{\hat{\alpha}^{\mathbf{z}}}, \hat{\boldsymbol{\alpha}}^{\mathbf{z}})$ satisfies

$$J(\mathbf{X}^{\hat{\alpha}^{\mathbf{z}}}, \hat{\boldsymbol{\alpha}}^{\mathbf{z}}) = \inf_{\boldsymbol{\alpha}^{\mathbf{z}} \in \mathcal{A}_{\mathbf{z}}^p} J(\mathbf{X}^{\alpha^{\mathbf{z}}}, \boldsymbol{\alpha}^{\mathbf{z}}). \quad (7.3)$$

Unlike the infinite horizon, the cost functional (and thus the value function) in the finite horizon case does depend on both the initial state and the initial time of the system. Hence, in order to solve (7.1)-(7.3) via the *dynamic programming approach*, we consider the associated family of control problems indexed by the initial time-state points.

Given $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^n$, for each admissible pair $(\mathbf{X}^{\boldsymbol{\alpha}^{t,\mathbf{x}}}, \boldsymbol{\alpha}^{t,\mathbf{x}})$, we define

$$\mathcal{J}^\alpha(t, \mathbf{x}) := \mathbb{E}^\alpha_{t,\mathbf{x}} \left[\int_t^T e^{-\gamma_s^\alpha} f(s, \mathbf{X}_s^\alpha, \boldsymbol{\alpha}_s) \, \mathrm{d}s + h(\mathbf{X}_T^\alpha) \right], \tag{7.4}$$

where $\mathbb{E}^\alpha_{t,\mathbf{x}}$ stands for the mathematical expectation conditional to $\mathbf{X}_t^{\boldsymbol{\alpha}^{t,\mathbf{x}}} = \mathbf{x}$. As before, we have omitted the superscripts $t, \mathbf{x}$ when appearing inside the operator $\mathbb{E}^\alpha_{t,\mathbf{x}}$. We have that $\mathcal{J}^\alpha(t, \mathbf{x})$ is the expected cost of using the control policy $\boldsymbol{\alpha}^{t,\mathbf{x}}$ over the time interval $[t, T]$ given the initial time-state point $(t, \mathbf{x})$. If $t = 0$, we write $\mathcal{J}^\alpha(\mathbf{x}) \equiv \mathcal{J}^\alpha(t, \mathbf{x})$.

The optimal cost function $V : [0, T] \times \mathbb{R}^n \to \mathbb{R}^+$ is then defined by

$$V(t, \mathbf{x}) := \inf_{\left\{ (\mathbf{X}^{\boldsymbol{\alpha}^{t,\mathbf{x}}}, \boldsymbol{\alpha}^{t,\mathbf{x}}) : \boldsymbol{\alpha}^{t,\mathbf{x}} \in \mathcal{A}^p_{t,\mathbf{x}} \right\}} V^{\boldsymbol{\alpha}}(t, \mathbf{x}). \tag{7.5}$$

Hence, $V(t, \mathbf{x})$ gives the minimum *cost-to-go*, starting at time t from state $\mathbf{x}$.

Following similar arguments than those used in Section 5, we can obtain the finite horizon counterpart of our previous results. We thus omit the repetition and only present the following verification theorem.

Theorem 7.1. *Let $f : \mathbb{R}_+ \times \mathbb{R}^n \times A \to \mathbb{R}_+$ and $h : \mathbb{R}^n \to \mathbb{R}_+$ be measurable functions. Let $\phi \in C([0, T] \times \mathbb{R}^n) \cap C^{1,2}([0, T) \times \mathbb{R}^n)$ be a nonnegative function with polynomial growth in $\mathbf{x}$ of degree p (uniformly in t), which solves*

$$\partial_t \phi(t, \mathbf{x}) + \inf_{\mathbf{a} \in A} \{ L^{\mathbf{a}} \phi(t, \mathbf{x}) - q(t, \mathbf{x}, \mathbf{a}) \phi(t, \mathbf{x}) + f(t, \mathbf{x}, \mathbf{a}) \} = 0, \quad \textit{for all } (t, \mathbf{x}) \in [0, T) \times \mathbb{R}^n, \tag{7.6}$$

with the boundary condition $\phi(T, \cdot) = h(\cdot)$. Suppose that, for each $\boldsymbol{\alpha}^{t,\mathbf{x}} \in \mathcal{A}^p_{t,\mathbf{x}}$, the family $\left\{ \phi \left(\tau \wedge T, \mathbf{X}^{\boldsymbol{\alpha}^{t,\mathbf{x}}}_{\tau \wedge T} \right) \right\}_{\tau \in \mathcal{T}}$ is uniformly integrable.[2] Then the following holds.

(a) *For any admissible policy $\boldsymbol{\alpha}^{t,\mathbf{x}} \in \mathcal{A}^p_{t,\mathbf{x}}$, $\phi(t, \mathbf{x}) \leq \mathcal{J}^\alpha(t, \mathbf{x})$, $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^n$.*

(b) *If there exists an admissible policy $\hat{\boldsymbol{\alpha}}^{t,\mathbf{x}} \in \mathcal{A}^p_{t,\mathbf{x}}$ with corresponding controlled process $\mathbf{X}^{\hat{\boldsymbol{\alpha}}^{t,\mathbf{x}}}$, defined on a filtered probability space $(\hat{\Omega}, \hat{\mathcal{F}}, (\hat{\mathcal{F}}_t), \hat{\mathbb{P}})$, such that $\hat{\mathbb{P}}$-a.s. for all $s \geq 0$*

$$\partial_s \phi(s, \mathbf{X}_s^{\hat{\boldsymbol{\alpha}}^{t,\mathbf{x}}}) + L^{\hat{\boldsymbol{\alpha}}_s^{t,\mathbf{x}}} \phi(s, \mathbf{X}_s^{\hat{\boldsymbol{\alpha}}^{t,\mathbf{x}}}) - q(s, \mathbf{X}_s^{\hat{\boldsymbol{\alpha}}^{t,\mathbf{x}}}, \hat{\boldsymbol{\alpha}}_s^{t,\mathbf{x}}) \phi(s, \mathbf{X}_s^{\hat{\boldsymbol{\alpha}}^{t,\mathbf{x}}}) + f(s, \mathbf{X}_s^{\hat{\boldsymbol{\alpha}}^{t,\mathbf{x}}}, \hat{\boldsymbol{\alpha}}_s^{t,\mathbf{x}}) = 0, \tag{7.7}$$

with the boundary condition $\phi(T, \cdot) = h(\cdot)$. Then, $\phi(t, \mathbf{x}) = J(\mathbf{X}^{\hat{\boldsymbol{\alpha}}^{t,\mathbf{x}}}, \hat{\boldsymbol{\alpha}}^{t,\mathbf{x}})$, the family $\{ \hat{\boldsymbol{\alpha}}^{t,\mathbf{x}} : (t, \mathbf{x}) \in [0, T] \times \mathbb{R}^n \}$ is a family of optimal policies, and $\phi(t, \mathbf{x})$ is the value function of the control problem (7.1)-(7.2).

Remark 7.2. The proof follows the same arguments used in the proof of Theorem 5.13, so the details are omitted.

7.2. **Possible extensions.** Let us now comment on three possible generalisations.

7.2.1. *Local dynamics and action set A.* . We can generalise our framework by, for example, replacing $\mathcal{M}_p$ by the set

$$\mathcal{M}'_p := \left\{ \text{measures } \nu \text{ on } \mathbb{R}^n \text{ such that } \int_{\mathbb{R}^n} \left(1 \wedge |\mathbf{y}|^2 \right) \vee |\mathbf{y}|^p \nu(\mathrm{d}\mathbf{y}) < \infty \right\}, \quad p \geq 1. \tag{7.8}$$

---

[2]As before, $\mathcal{T}$ denotes the family of all stopping times.

Hence, to define the local dynamics of a control process, we can replace (3.1) by the more general Lévy operator

$$(L^{\mathbf{a}}g)(\cdot) := (\mathbf{u} + \boldsymbol{\mu})^T \nabla g(\cdot) + \frac{1}{2} \operatorname{Tr}(\boldsymbol{\sigma}^T \operatorname{H}g \boldsymbol{\sigma})(\cdot) + \int_{\mathbb{R}_0^n} \left(g(\cdot + \mathbf{y}) - g(\cdot) - \mathbf{y}^T \nabla g(\cdot) 1_{\{|\mathbf{y}| \geq 1\}} \right) \nu(\mathrm{d}\mathbf{y}).$$
(7.9)

7.2.2. *Running cost function f.* Notice that we have assumed that $f \geq 0$. This condition can be relaxed, for instance, by considering that f is bounded below, say $f \geq m$. In such a case, the previous results can be applied to the modified running cost $\tilde{f} := f - m$. Indeed, under the following assumptions:

(**TCa**) Each admissible policy $\boldsymbol{\alpha}^{\mathbf{x}}$ is such that either $J^{\alpha}(\mathbf{x}) = \infty$ or

$$\liminf_{t \to \infty} \mathbb{E}_{\mathbf{x}}^{\alpha} \left[e^{-\gamma_t^{\boldsymbol{\alpha}}} \phi(\mathbf{X}_t^{\boldsymbol{\alpha}}) \right] \leq 0,$$
(7.10)

(**TCb**) Each admissible policy $\boldsymbol{\alpha}^{\mathbf{x}}$ is such that

$$\limsup_{t \to \infty} \mathbb{E}_{\mathbf{x}}^{\alpha} \left[e^{-\gamma_t^{\boldsymbol{\alpha}}} \phi(\mathbf{X}_t^{\boldsymbol{\alpha}}) \right] \geq 0.$$
(7.11)

Lemma 5.10 becomes:

Lemma 7.3 (Verification Result 1'). *Let $p \geq 2$ and the running cost function f be bounded below. Let $\phi \in C^2(\mathbb{R}^n)$ be a function bounded below and satisfying (**SC**). Then the following holds.*

i) *If for every admissible policy $\boldsymbol{\alpha}^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}^p$ conditions (**SC**) and (**TCa**) hold, then $\phi(\mathbf{x}) \leq J^{\alpha}(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^n$.*

ii) *If (**MC**) and (**TCb**) hold, then $\phi(\mathbf{x}) \geq J^{\hat{\alpha}}(\mathbf{x})$. The equality holds when, for every admissible policy with finite payoff,*

$$\lim_{t \to \infty} \mathbb{E}_{\mathbf{x}}^{\alpha} \left[e^{-\gamma_t^{\boldsymbol{\alpha}}} \phi(\mathbf{X}_t^{\boldsymbol{\alpha}}) \right] = 0.$$
(7.12)

In the latter case, $\{\hat{\boldsymbol{\alpha}}^{\mathbf{x}} : \mathbf{x} \in \mathbb{R}^n\}$ is a family of admissible optimal policies and ϕ equals the value function.

In a similar way, we can extend all our results for a lower bounded running cost. For more general functions f, additional constraints are required to ensure that the payoff $J^{\alpha}(\mathbf{x})$ is well-defined.

7.2.3. *Weak solutions to HJB equation.* The assumption of having a C^2 solution to the HJB equation guarantees that such a solution is regular enough for the integro-differential equation to make sense. However, since the existence of C^2 solutions is difficult to guarantee, a natural approach to deal with this issue is to introduce the concept of *viscosity solutions*, as has been done in other settings. We leave the study of this issue to future research.

8. Applications

8.1. **Example 1.** Let $f : \mathbb{R} \to \mathbb{R}_+$ be a symmetric convex function with polynomial growth of degree $p \geq 2$. Define

$$\mathcal{M}_{\leq 1} := \{\nu \in \mathcal{M}_p \text{ such that } \nu(\mathbb{R}) \leq 1\}.$$
(8.1)

Let A'' be the action control set given by

$$A'' := \left\{ a = (\sigma, \nu, \mu) \,\middle|\, \sigma = 1; \nu \in \mathcal{M}_{\leq 1}, \mu = \int y\nu(\mathrm{d}y) \right\},$$
(8.2)

and let $\mathcal{A}_x$ be the set of A''-valued admissible control processes $\alpha^x = (\alpha_s^x)_{s \geq 0}$ (as defined in Section (3.1) with $p \geq 2$ and $u = 0$ for the operator L^a).

20

We seek the optimal function

$$V : x \mapsto \inf_{\alpha^x \in \mathcal{A}_x} \mathbb{E}_x \left[\int_0^\infty e^{-qt} f(X_t^{\alpha^x}) \, dt \right]. \tag{8.3}$$

Let B be a standard Brownian motion started at zero and define

$$\psi : x \mapsto \mathbb{E} \left[\int_0^\infty e^{-(q+1)} f(x + B_t) \, dt \right]. \tag{8.4}$$

Lemma 8.1. *Suppose that*

$$\beta := \sup_{\nu \in \mathcal{M}_{\leq 1}} \int_{|y|>1} |y|^p \nu(\mathrm{d}y) < \infty. \tag{8.5}$$

Then the following holds:

  i) ψ *is symmetric and convex and its global minimum is attained at zero.*

  ii) The optimal payoff is given by

$$V = \psi + c, \quad \text{where} \quad c = \frac{\psi(0)}{q}, \tag{8.6}$$

  and $\{ \hat{\alpha}^x = (1, -X_\cdot^x, \delta_{-X^x}), : x \in \mathbb{R} \}$ *is a family of admissible optimal policies, where* X^x *is the Markov process, started at* x, *whose infinitesimal generator defined on* $C_c^2(\mathbb{R}^n)$ *is given by*

$$Gh(\cdot) := \frac{1}{2} h''(\cdot) + \int (h(\cdot + y) - h(\cdot)) \hat{\nu}(\cdot; \mathrm{d}y), \tag{8.7}$$

  with $\hat{\nu}(x; \mathrm{d}y) := \delta_{-x}(\mathrm{d}y)$, *the unit point mass at* $-x$. *Thus, the optimal control is to always jump* X *to zero at maximal (i.e. unit) rate.*

Remark 8.2. In this example, conditions (**TC**) and (**UI**) are easily verified thanks to the definition of the action control set and condition (8.5).

8.2. **Example 2.** Let us consider the same control setting as in Example 1, but with a slightly different running cost function. Let $f : \mathbb{R} \to \mathbb{R}_+$ be a polynomial of degree $p \geq 2$. Suppose that f is a symmetric, C^2, convex function increasing on $\mathbb{R}_+$. Given $\kappa > 0$, we seek

$$V : x \mapsto \inf_{\alpha^x \in \mathcal{A}_x} \mathbb{E}_x \left[\int_0^\infty e^{-qt} \left[f(X_t^{\alpha^x}) + \kappa \nu_t(\mathbb{R}) \right] \mathrm{d}t \right]. \tag{8.8}$$

The HJB equation for the control problem (8.8) is now given by

$$\inf_{a \in [0,1]} \left\{ \frac{1}{2} g'' - qg + f + \kappa a + \int (g(x+y) - g(x)) \nu(\mathrm{d}y) \right\} = 0. \tag{8.9}$$

Theorem 8.3. *Given* $b \geq 0$, *define* $\phi_b : \mathbb{R} \to \mathbb{R}_+$ *by*

$$\phi_b(x) = \mathbb{E}_x \left[\int_0^\infty e^{-qt} \left(f(B_t^{b,x}) + \kappa 1_{|B_t^{b,x}| \geq b} \right) \mathrm{d}t \right], \tag{8.10}$$

with $B^{b,x}$ *being a controlled BM, started at* x, *which is jumped to the origin at rate 1 whenever* $|B^{b,x}| \geq b$ *and is otherwise uncontrolled. Then, the value function* V *defined in (8.8) is given by* $\phi \equiv \phi_{\hat{b}}$, *where* $\hat{b}$ *solves* $\phi_{\hat{b}}(\hat{b}) - \phi_{\hat{b}}(0) = \kappa$, *with corresponding optimal control.*

Remark 8.4. Unlike Example 1 wherein condition (8.5) is key to guarantee the transversality condition, in this second example such a condition is a consequence of Corollary 5.21 and the polynomial form of the running cost f.

8.3. **Example 3 (Quadratic Control).** We now consider the case when the running cost function $f(\mathbf{x}, (\boldsymbol{\sigma}, \nu, \boldsymbol{\mu}))$ is a quadratic form as a function of $\mathbf{x}$ and $\boldsymbol{\mu}$. We will see that the associated payoff function turns out to be a quadratic form as well and we obtain an explicit solution to the stochastic problem (5.1)-(5.3).

Let Λ and Θ be positive definite symmetric matrices in $\mathrm{M}_{n \times n}(\mathbb{R})$. Consider the running cost function $f : \mathbb{R}^n \times A \to \mathbb{R}^+$ defined as the quadratic form $f(\mathbf{x}, \mathbf{a}) := \mathbf{x}^T \Lambda \mathbf{x} + \boldsymbol{\mu}^T \Theta \boldsymbol{\mu}$ for each $\mathbf{a} = (\boldsymbol{\sigma}, \nu, \boldsymbol{\mu}) \in A$. Let $\mathbf{B}$ be a symmetric positive definite matrix solving the algebraic Riccati equation

$$\mathbf{B}^T \Theta^{-1} \mathbf{B} + q\mathbf{B} - \Lambda = 0. \tag{8.11}$$

Let D be an open subset of $M_{n \times n}(\mathbb{R}) \times \mathcal{M}_2$ and set $\Gamma = D \times \mathbb{R}^n$ as the action set A. Define

$$\hat{\delta} := \inf_{(\boldsymbol{\sigma}, \nu) \in D} \left(\mathrm{Tr}(\boldsymbol{\sigma}^T \mathbf{B} \, \boldsymbol{\sigma}) + \int_{\mathbb{R}_0^n} \mathbf{y}^T \mathbf{B} \, \mathbf{y} \, \nu(\mathrm{d}\mathbf{y}) \right), \tag{8.12}$$

and suppose that the infimum in (8.12) is attained at $(\hat{\boldsymbol{\sigma}}, \hat{\nu}) \in D$.

Theorem 8.5. *Consider the control problem (5.1)-(5.3) over the class of admissible controls $\hat{\mathcal{A}}_{\mathbf{x}}^2$ (see Definition 6.1) and with the quadratic running cost f given above. Define $\mathbf{Q} := \Theta^{-1} \mathbf{B}$ and $\mathbf{v} := -\Theta^{-1} \mathbf{P} \mathbf{u}$, where $\mathbf{P} := \mathbf{B} \Lambda^{-1} \mathbf{B}$, and $\mathbf{u} \in \mathbb{R}^n$ is the drift term in (3.1). Let $\hat{\mu} : \mathbb{R}^n \to \mathbb{R}^n$ be defined by $\hat{\mu}(\mathbf{x}) := -\mathbf{Q}\mathbf{x} + \mathbf{v}$. Then the following holds:*

(i) The family $\{\hat{\boldsymbol{\alpha}}^{\mathbf{x}} : \mathbf{x} \in \mathbb{R}^n\}$ defined by

$$\hat{\boldsymbol{\alpha}}_t^{\mathbf{x}} := \left(\hat{\boldsymbol{\sigma}}, \hat{\nu}, \hat{\mu}(\hat{\mathbf{X}}_t^{\mathbf{x}}) \right), \tag{8.13}$$

is a family of optimal policies, where the associated controlled process $\mathbf{X}_t^{\hat{\boldsymbol{\alpha}}^{\mathbf{x}}} \equiv \hat{\mathbf{X}}_t^{\mathbf{x}}$ is the $\mathbb{R}^n$-valued process with infinitesimal generator $\hat{L}$ defined on functions $f \in C_c^2(\mathbb{R}^n)$ by

$$(\hat{L}g)(\cdot) := (\mathbf{u} + \hat{\mu}(\cdot))^T \nabla g(\cdot) + \frac{1}{2} \mathrm{Tr}(\hat{\boldsymbol{\sigma}}^T \mathrm{H}g(\cdot) \, \hat{\boldsymbol{\sigma}}) + \int_{\mathbb{R}_0^n} (g(\cdot + \mathbf{y}) - g(\cdot) - \mathbf{y}^T \nabla g(\cdot)) \hat{\nu}(\mathrm{d}\mathbf{y}), \tag{8.14}$$

(ii) The value function V is given by $V(\mathbf{x}) = \mathbf{x}^T \mathbf{B} \mathbf{x} + \mathbf{x}^T \mathbf{c} + d$ for all $\mathbf{x} \in \mathbb{R}^n$, where $\mathbf{c} \in \mathbb{R}^n$ and $d \in \mathbb{R}$ are given by

$$\mathbf{c} := 2\mathbf{P}^T \mathbf{u}, \qquad d := \frac{1}{q} \left(2\mathbf{u}^T \mathbf{P}^T \mathbf{u} + \hat{\delta} - \mathbf{u}^T \mathbf{P} \Theta^{-1} \mathbf{P}^T \mathbf{u} \right). \tag{8.15}$$

Remark 8.6. The proof, given in Section C.3, follows again a verification approach: we first show that $\hat{\boldsymbol{\alpha}}^{\mathbf{x}}$ as defined in (8.13) is an admissible policy for each $\mathbf{x} \in \mathbb{R}^n$ (see Lemma C.2 in Section A). We then prove that $\phi(\mathbf{x}) := \mathbf{x}^T \mathbf{B} \mathbf{x} + \mathbf{x}^T \mathbf{c} + d$ satisfies the assumptions of Theorem 5.13. Here we verify that the pointwise minimisation of the corresponding HJB equation yields the algebraic matrix equation (8.11).

Remarks 8.7.

i) Notice that the optimal family of policies defined in (8.13) is a linear function of the state $\mathbf{x}$. This family depends on the solvability of the algebraic matrix Riccati equation (8.11). Although the dynamics of the controlled system are not linear, this example can be thought of as a generalisation of the standard linear quadratic regulator (LQR) problem, see, for example, the finite horizon case in [8, Chapter VI, Section 5, p.165] .

ii) Various criteria to guarantee the existence and uniqueness of a positive definite solution to the Riccati equation (8.11) are very well-known in the literature (see, for example, [11,17,32], and references therein). Furthermore, it is also known that such a solution can be expressed in terms of the eigenvectors of the $2n \times 2n$-matrix

$$\left\{ \begin{array}{cc} -\frac{q}{2}\mathbf{I} & \Theta^{-1} \\ -\boldsymbol{\Lambda} & \frac{q}{2}\mathbf{I} \end{array} \right\},$$

see [24, Theorem 1], [16, Theorem 1].

Particular case. If the weight cost matrices for the control problem (5.1)-(5.3) are the diagonal matrices $\Lambda := \lambda \mathbf{I}$ and $\Theta := \theta \mathbf{I}$, with $\lambda \geq 0$ and $\theta > 0$, then the coefficients of the corresponding value function V take the explicit values

$$\mathbf{B} := \frac{\theta}{2}(p-q)\mathbf{I}, \qquad \mathbf{c} := \frac{8\lambda}{\theta(p+q)^2}\mathbf{u}, \qquad d := \frac{\|\mathbf{u}\|^2}{q\theta(p+q)^2}(8\lambda - (p-q)^2) + \frac{\theta\tilde{\delta}(p-q)}{2q},$$

where $p := \sqrt{q^2 + 4\lambda/\theta}$.

APPENDIX A. PROOFS OF RESULTS IN SECTIONS 3 AND 4

A.1. **Proof of Lemma 3.4.**

Proof. Define $\mathbf{u}^{\xi,\oplus_t} := \boldsymbol{\alpha}^{r,\xi} \oplus_t \boldsymbol{\beta}_t^{\alpha}$. Consider the admissible pairs $(\mathbf{X}^{\boldsymbol{\alpha}^{r,\xi}}, \boldsymbol{\alpha}^{r,\xi})$ and $(\mathbf{Y}^{\boldsymbol{\beta}_t^{\alpha}}, \boldsymbol{\beta}_t^{\alpha})$ defined on the probability spaces $(\Omega^{\alpha}, \mathcal{F}^{\alpha}, (\mathcal{F}_t^{\alpha}), \mathbb{P}_{\xi})$ and $(\Omega^{\beta}, \mathcal{F}^{\beta}, (\mathcal{F}_t^{\beta}), \mathbb{P}_{\eta_t^{\alpha}}^{\beta})$, respectively. Notice that in the notation $\mathbb{P}_{\xi}^{\alpha}$ and $\mathbb{P}_{\eta_t^{\alpha}}^{\beta}$ we have made explicit the initial distributions ξ and η_t^{α} of the corresponding control processes $\mathbf{X}^{\boldsymbol{\alpha}^{r,\xi}}$ and $\mathbf{Y}^{\boldsymbol{\beta}_t^{\alpha}}$, respectively. Define a new filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ by setting $\Omega := \Omega^{\alpha} \times \Omega^{\beta}$ endowed with the product σ-algebra $\mathcal{F} := \mathcal{F}^{\alpha} \otimes \mathcal{F}^{\beta}$ generated by the measurable rectangles. Define the probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ as the probability measure on $(\Omega^{\alpha} \times \Omega^{\beta}, \mathcal{F}^{\alpha} \otimes \mathcal{F}^{\beta})$ given by

$$\mathbb{P}(B) := \int_{\Omega^{\alpha}} \mathbb{P}_{\eta_t^{\alpha}}^{\beta}(B_{\omega_1}) \, d\mathbb{P}_{\xi}^{\alpha}(\omega_1), \qquad B \in \mathcal{F}^{\alpha} \otimes \mathcal{F}^{\beta}, \tag{A.1}$$

where, for each $\omega_1 \in \Omega^{\alpha}$, $B_{\omega_1} := \{\omega_2 \in \Omega^{\beta} : (\omega_1, \omega_2) \in B\}$ denotes the ω_1-section of B. Note that $B_{\omega_1} \in \mathcal{F}^{\beta}$. We can now define the process $\mathbf{Z}^{\mathbf{u}^{\xi,\oplus_t}}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ as follows

$$\mathbf{Z}_s^{\mathbf{u}^{\xi,\oplus_t}}(\omega_1, \omega_2) = \omega_1(s)1_{[r,t)}(s) + \omega_2(s)1_{[t,\infty)}(s), \quad s \in \mathbb{R}_+, \, \omega_1 \in \Omega^{\alpha}, \omega_2 \in \Omega^{\beta}. \tag{A.2}$$

By construction $(\mathbf{Z}^{\mathbf{u}^{\xi,\oplus_t}}, \mathbf{u}^{\xi,\oplus_t})$ is an $\mathbb{R}^n \times A$-valued, $(\mathcal{F}_s)$-adapted càdlàg process which agrees in law with $(\mathbf{X}^{\boldsymbol{\alpha}^{r,\xi}}, \boldsymbol{\alpha}^{r,\xi})$ on $[r,t)$ and with $(\mathbf{Y}^{\boldsymbol{\beta}_t^{\alpha}}, \boldsymbol{\beta}_t^{\alpha})$ on $[t,\infty)$.

Note now that the integrability of $\int_r^s \left| L^{\mathbf{u}_l^{\xi,\oplus_t}} h\left(\mathbf{Z}_l^{\mathbf{u}^{\xi,\oplus_t}}\right) \right| dl < +\infty$, for $s > r$ and $h \in C_c^2(\mathbb{R}^n)$, as well as the validity of condition (3.3), follow from the validity of such conditions for the admissible pairs $(\mathbf{X}^{\boldsymbol{\alpha}^{r,\xi}}, \boldsymbol{\alpha}^{r,\xi})$ and $(\mathbf{Y}^{\boldsymbol{\beta}_t^{\alpha}}, \boldsymbol{\beta}_t^{\alpha})$ associated with $\mathbf{u}^{\xi,\oplus_t}$ and $\mathbf{Z}^{\mathbf{u}^{\xi,\oplus_t}}$, respectively. It thus remains to prove that

$$M_t^{h,\mathbf{u}^{\xi,\oplus_t}} := h(\mathbf{Z}_t^{\mathbf{u}^{\xi,\oplus_t}}) - \int_r^t \left(L^{\mathbf{u}_s^{\xi,\oplus_t}} h\right)\left(\mathbf{Z}_{s-}^{\mathbf{u}^{\xi,\oplus_t}}\right) ds, \quad t \geq r, \tag{A.3}$$

is an $(\mathcal{F}_t)$-local martingale under $\mathbb{P}$. Let us then prove that $\mathbb{E}\left(M_s^{h,\mathbf{u}^{\xi,\oplus_t}} \middle| \mathcal{F}_k\right) = M_k^{h,\mathbf{u}^{\xi,\oplus_t}}$ for each $r \leq k \leq s$. For this, we shall use that $M_{\cdot}^{h,\boldsymbol{\alpha}^{r,\xi}}$ and $M_{\cdot}^{h,\boldsymbol{\beta}_t^{\alpha}}$ are $(\mathcal{F}_s^{\alpha})$- and $(\mathcal{F}_s^{\beta})$-local martingales, respectively.

Case 1. If $k \leq s \leq t$, then $\mathbb{E}\left(M_s^{h,\mathbf{u}^{\xi,\oplus_t}} \middle| \mathcal{F}_k\right) = \mathbb{E}\left(M_s^{h,\boldsymbol{\alpha}^{r,\xi}} \middle| \mathcal{F}_k^{\alpha}\right) = M_k^{h,\boldsymbol{\alpha}^{r,\xi}}$. Similarly, if $t \leq k \leq s$, then $\mathbb{E}\left(M_s^{h,\mathbf{u}^{\xi,\oplus_t}} \middle| \mathcal{F}_k\right) = \mathbb{E}\left(M_s^{h,\boldsymbol{\beta}_t^{\alpha}} \middle| \mathcal{F}_k^{\beta}\right) = M_k^{h,\boldsymbol{\beta}_t^{\alpha}}$, as required.

Case 2. If $k \leq t \leq s$, then, by definition of $(\mathbf{Z}^{\mathbf{u}^{\xi,\oplus_t}}, \mathbf{u}^{\xi,\oplus_t})$ and by the law of iterated conditional expectation, we obtain that

$$M_s^{h,\mathbf{u}^{\xi,\oplus_t}} = M_s^{h,\boldsymbol{\beta}_t^{\alpha}} - \int_r^t L^{\boldsymbol{\alpha}_l^{r,\xi}} h\left(\mathbf{X}_l^{\boldsymbol{\alpha}^{r,\xi}}\right) dl \tag{A.4}$$

$$\mathbb{E}\left(M_s^{h,\boldsymbol{\beta}_t^{\alpha}} \middle| \mathcal{F}_k\right) = \mathbb{E}\left[\mathbb{E}\left(M_s^{h,\boldsymbol{\beta}_t^{\alpha}} \middle| \mathcal{F}_t^{\beta}\right) \middle| \mathcal{F}_k\right] = \mathbb{E}\left[M_t^{h,\boldsymbol{\beta}_t^{\alpha}} \middle| \mathcal{F}_k\right] = \mathbb{E}\left[h\left(\mathbf{X}_t^{\boldsymbol{\alpha}^{r,\xi}}\right) \middle| \mathcal{F}_k\right]. \tag{A.5}$$

Therefore,

$$
\begin{aligned}
\mathbb{E}\left(M_s^{h,\mathbf{u}^{\xi,\oplus_t}}\Big|\mathcal{F}_k\right) &= \mathbb{E}\left(M_s^{h,\boldsymbol{\beta}_t^{\alpha}} - \int_r^t L^{\boldsymbol{\alpha}_l^{r,\xi}} h\left(\mathbf{X}_l^{\boldsymbol{\alpha}^{r,\xi}}\right)\mathrm{d}l\Big|\mathcal{F}_k\right) \\
&= \mathbb{E}\left(h\left(\mathbf{X}_t^{\boldsymbol{\alpha}^{r,\xi}}\right) - \int_r^t L^{\boldsymbol{\alpha}_l^{r,\xi}} h\left(\mathbf{X}_l^{\boldsymbol{\alpha}^{r,\xi}}\right)\mathrm{d}l\Big|\mathcal{F}_k\right) \\
&= \mathbb{E}\left(M_t^{h,\boldsymbol{\alpha}^{r,\xi}}\Big|\mathcal{F}_k^{\alpha}\right) = M_k^{h,\boldsymbol{\alpha}^{r,\xi}}, \quad\quad\quad\quad\quad\quad (A.6)
\end{aligned}
$$

as desired. ∎

A.2. **Preliminary results.** Let us introduce some additional notation and give some preliminary technical results.

Given a function $f \in C^2(\mathbb{R}^n)$, define a sequence $\{f_K\}_K \subset C_c^2(\mathbb{R}^n)$ as follows. For each $K > 1$, $K \in \mathbb{N}$, set

$$f_K := f\,\zeta_K, \quad \text{where} \quad \zeta_K \in C_c^2(\mathbb{R}^n), \ \mathbf{1}_{B(0,2K)} \le \zeta_K \le \mathbf{1}_{B(0,3K)}. \quad\quad (A.7)$$

Note that $f_K \to f$ pointwise as $K \to \infty$.

Given an admissible pair $(\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}^{\mathbf{x}})$ with $\boldsymbol{\alpha}^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}^p$, $p \ge 2$, we define, for each $m < K$, $m \in \mathbb{N}$, the stopping times $\tau_m := T_m \wedge S_m$, where

$$T_m := \inf\left\{r \in \mathbb{R}_+ : \left|\mathbf{X}_r^{\boldsymbol{\alpha}^{\mathbf{x}}}\right| > m\right\} \wedge m, \quad\quad\quad\quad\quad\quad (A.8)$$

and

$$S_m := \inf\left\{r : \int_0^r Q_s^{p,\boldsymbol{\alpha}^{\mathbf{x}}}\,\mathrm{d}s > m\right\}, \quad m \ge 1, \quad\quad\quad\quad (A.9)$$

with $Q_s^{p,\boldsymbol{\alpha}^{\mathbf{x}}}$ as given in (3.3) and the usual convention $\inf \varnothing = \infty$. Note that $\tau_m \to \infty$ $\mathbb{P}^{\alpha} - a.s.$ as $m \to \infty$ (thanks to the càdlàg property of $\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}$ and the continuity of $r \mapsto \int_0^r Q_s^{p,\boldsymbol{\alpha}^{\mathbf{x}}}\,\mathrm{d}s$).

For each $\mathbf{a} = (\boldsymbol{\sigma}, \nu, \boldsymbol{\mu}) \in A$, we will rewrite

$$\left(L^{\mathbf{a}} f\right)(\cdot) = \left(A^{(\boldsymbol{\mu},\boldsymbol{\sigma})} f\right)(\cdot) + (G^{\nu} f)(\cdot). \quad\quad\quad\quad (A.10)$$

where

$$\left(A^{(\boldsymbol{\mu},\boldsymbol{\sigma})} f\right)(\cdot) := (\mathbf{u} + \boldsymbol{\mu})^T \nabla f(\cdot) + \frac{1}{2}\mathrm{Tr}(\boldsymbol{\sigma}^T \mathrm{H}f\, \boldsymbol{\sigma})(\cdot) \quad\quad\quad (A.11)$$

$$(G^{\nu} f)(\cdot) := \int_{\mathbb{R}_0^n} (f(\cdot + \mathbf{y}) - f(\cdot) - \mathbf{y} \cdot \nabla f(\cdot))\,\nu(\mathbf{y}). \quad\quad\quad (A.12)$$

Lemma A.1. *Let $f \in C^2(\mathbb{R}^n)$ be a function satisfying $|f(\mathbf{x})| \le C(|\mathbf{x}|^q + 1)$ for some $q \in [1,p]$, $p \ge 2$. Take $m \in \mathbb{N}$ and let $\{f_K\}_{K>m}$ be a sequence of functions approximating f defined via (A.7). Then, for any admissible pair $(\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}^{\mathbf{x}})$, $\boldsymbol{\alpha}^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}^p$, there exists a positive constant $C(m,f)$ (independent of K), such that, for each $K > m$,*

$$\left|(L^{\boldsymbol{\alpha}_s^{\mathbf{x}}} f_K)(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{\mathbf{x}}})\right| \le C(m,f) Q_s^{p,\boldsymbol{\alpha}^{\mathbf{x}}}, \quad s \le \tau_m. \quad\quad\quad (A.13)$$

In particular, the $\mathbb{P}^{\alpha} - a.s.$ convergence

$$\int_0^{t \wedge \tau_m} (L^{\boldsymbol{\alpha}_s^{\mathbf{x}}} f_K)(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{\mathbf{x}}})\,\mathrm{d}s \ \to \ \int_0^{t \wedge \tau_m} (L^{\boldsymbol{\alpha}_s^{\mathbf{x}}} f)(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{\mathbf{x}}})\,\mathrm{d}s, \quad\quad\quad (A.14)$$

holds for all $t \in \mathbb{R}_+$ as $K \to \infty$.

Remark A.2. Recall that we will always omit the superscript $\mathbf{x}$ in $\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}$ and $\boldsymbol{\alpha}^{\mathbf{x}}$ whenever they appear inside the operator $\mathbb{E}_{\mathbf{x}}^{\alpha}$.

Proof. Take $\boldsymbol{\alpha}^{\mathbf{x}} = (\boldsymbol{\sigma}, \nu, \boldsymbol{\mu}) \in \mathcal{A}_{\mathbf{x}}^p$, f and $\{f_K\}_K$ as in the statement. Since $f_K \in C_c^2(\mathbb{R}^n)$, the continuity of f_K and the fact that $f_K = f$ on $[-2K, 2K]$ yields

$$\left|\left(A^{(\boldsymbol{\mu}_s,\boldsymbol{\sigma}_s)} f_K\right)(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{\mathbf{x}}})\right| = \left|\left(A^{(\boldsymbol{\mu}_s,\boldsymbol{\sigma}_s)} f\right)(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{\mathbf{x}}})\right| \le c_0(m,f)\left(|\boldsymbol{\mu}_s| + \|\boldsymbol{\sigma}_s\|^2\right), \quad\quad (A.15)$$

where $c_0(m,f) := \max\{ (\mathbf{u} + 1) \sup_{|\mathbf{z}| \le m} |\nabla f(\mathbf{z})|, \frac{1}{2}\sup_{|\mathbf{z}| \le m} \|\mathrm{H}f(\mathbf{z})\| \}$.

As for the non-local part, observe that for each $\mathbf{x} \in \mathbb{R}^n$ the integral term $\left(G^{\nu_s} f_K\right)(\mathbf{x})$ can be split into two regions: $E := \{\mathbf{y} \in \mathbb{R}_0^n : |\mathbf{y}| \le 1\}$ and $E^c := \{\mathbf{y} \in \mathbb{R}_0^n : |\mathbf{y}| > 1\}$. For $|\mathbf{x}| \le m$, by Taylor's theorem, there exists $\theta \in (0,1)$ such that

$$|f_K(\mathbf{x}+\mathbf{y}) - f_K(\mathbf{x}) - \mathbf{y} \cdot \nabla f_K(\mathbf{x})| = \frac{1}{2} \Big| \sum_{i,j=1}^n \partial_{ij}^2 f_K(\mathbf{x}+\theta\mathbf{y}) y_i y_j \Big| \le c_1(m,f) |\mathbf{y}|^2, \quad \mathbf{y} \in E, \quad (\text{A.16})$$

where $c_1(m,f) := \frac{1}{4} |\mathbf{y}|^2 \sum_{i,j=1}^n \sup_{|\mathbf{z}| \le m+1} \left| \partial_{ij}^2 f(\mathbf{z}) \right|$. Note the use of inequality $2 y_i y_j \le y_i^2 + y_j^2 \le |\mathbf{y}|^2$, as well as the fact that c_1 does not depend on K as (by construction) $f_K = f$ on $[-2K, 2K] \subset [-m-1, m+1]$.

On the other hand, again using that $|f_K| \le |f|$ and f has polynomial growth of degree $q \in [1,p]$, we can find a positive constant $c_2'(m,f) > 0$ such that $\left| f_K(\mathbf{x}+\mathbf{y}) \right| \le c_2'(m,f) |\mathbf{y}|^p$, for all $|\mathbf{x}| \le m$ and $\mathbf{y} \in E^c$. Thus

$$|f_K(\mathbf{x}+\mathbf{y}) - f_K(\mathbf{x}) - \mathbf{y} \cdot \nabla f_K(\mathbf{x})| \le c_2(m,f) |\mathbf{y}|^p, \quad \mathbf{y} \in E^c, \quad (\text{A.17})$$

where $c_2(m,f) := \left(c_2'(m,f) + \sup_{|\mathbf{z}| \le m} \{ |f(\mathbf{z})| + |\nabla f(\mathbf{z})\}| \right)$.

Since $f_K = f$ on $[-2K, 2K]$ and $\left| \mathbf{X}_{s-}^{\boldsymbol{\alpha^x}} \right| \le m < K$ $\mathbb{P}^\alpha - a.s.$ for all $s \le \tau_m$, the estimates (A.16) - (A.17) imply that

$$| \left(G^\nu f_K\right)(\mathbf{X}_{s-}^{\boldsymbol{\alpha^x}})| \le c(m,f) \int_{\mathbb{R}_0^n} |\mathbf{y}|^2 \vee |\mathbf{y}|^p \nu_s(\mathrm{d}\mathbf{y}) \nu_s(\mathrm{d}\mathbf{y}), \quad \text{for all} \ \ K \ge m, \quad (\text{A.18})$$

where $c(m,f) := 2 \max\{c_1(m,f), c_2(m,f)\}$. Estimates (A.15) and (A.18), together with (A.10) yield

$$\left| (L^{\boldsymbol{\alpha^x}}_s f_K)(\mathbf{X}_{s-}^{\boldsymbol{\alpha^x}}) \right| \le c_0(m,f) \left(|\boldsymbol{\mu}_s| + \|\boldsymbol{\sigma}_s\|^2 \right) + c(m,f) \int_{\mathbb{R}_0^n} |\mathbf{y}|^2 \vee |\mathbf{y}|^p \nu_s(\mathrm{d}\mathbf{y}) \nu_s(\mathrm{d}\mathbf{y}).$$

Estimate (A.13) follows by setting $C(m,f) := \max\{c_0(m,f), c(m,f)\}$ (recall definition of $Q_s^{p,\boldsymbol{\alpha^x}}$ in (3.3)).

Now, to prove the convergence (A.14), thanks to (A.10) and the equality in (A.15), it is sufficient to prove the $\mathbb{P}^\alpha - a.s.$ convergence, for all $t \in \mathbb{R}_+$,

$$\int_0^{t \wedge \tau_m} \left(G^{\nu_s} f_K\right)(\mathbf{X}_{s-}^{\boldsymbol{\alpha^x}}) \, \mathrm{d}s \ \rightarrow \ \int_0^{t \wedge \tau_m} \left(G^{\nu_s} f\right)(\mathbf{X}_{s-}^{\boldsymbol{\alpha^x}}) \, \mathrm{d}s \quad \text{as } K \to \infty. \quad (\text{A.19})$$

Since (A.18) holds for f_K and $\int_{\mathbb{R}_0^n} |\mathbf{y}|^2 \vee |\mathbf{y}|^p \nu_s(\mathrm{d}\mathbf{y}) < +\infty$ (as $\nu_s \in \mathcal{M}_p$, recall definition (2.1)), the DCT implies

$$\left(G^{\nu_s} f_K\right)(\mathbf{X}_{s-}^{\boldsymbol{\alpha^x}}) \ \rightarrow \ \left(G^{\nu_s} f\right)(\mathbf{X}_{s-}^{\boldsymbol{\alpha^x}}) \quad \text{as } K \to \infty. \quad (\text{A.20})$$

Moreover, since $\int_0^{t \wedge \tau_m} \mathrm{d}s \int_{\mathbb{R}_0^n} |\mathbf{y}|^2 \vee |\mathbf{y}|^p \nu_s(\mathbf{y}) \, \mathrm{d}\mathbf{y} < m$ (by definition of S_m and because $\tau_m \le S_m$), DCT implies (A.19), as required. ∎

Theorem A.3. *Let $p \ge 2$ and take any admissible pair $(\mathbf{X}^{\boldsymbol{\alpha^x}}, \boldsymbol{\alpha^x})$, $\boldsymbol{\alpha^x} \in \mathcal{A}_{\mathbf{x}}^p$ defined on the filtered probability space $(\Omega^\alpha, \mathcal{F}^\alpha, \mathbb{F}^\alpha := (\mathcal{F}_t^\alpha), \mathbb{P}^\alpha)$. Then, for each bounded function $f \in C^2(\mathbb{R}^d)$ with polynomial growth of degree $q \in [1,p]$, the process $M^{f,\boldsymbol{\alpha^x}} = (M_t^{f,\boldsymbol{\alpha^x}})_{t \in \mathbb{R}_+}$ defined by*

$$M_t^{f,\boldsymbol{\alpha^x}} := f(\mathbf{X}_t^{\boldsymbol{\alpha^x}}) - \int_0^t \left(L^{\boldsymbol{\alpha^x}}_s f\right)(\mathbf{X}_{s-}^{\boldsymbol{\alpha^x}}) \, \mathrm{d}s, \quad t \in \mathbb{R}_+, \quad (\text{A.21})$$

is a local $(\mathbb{F}^\alpha, \mathbb{P}^\alpha)$-martingale.

Proof. Let $\boldsymbol{\alpha^x} = (\boldsymbol{\sigma}, \nu, \boldsymbol{\mu}) \in \mathcal{A}_{\mathbf{x}}^p$, $\mathbf{X}^{\boldsymbol{\alpha^x}}$ and f be as in the statement. Take $m \in \mathbb{N}$ and consider a sequence $\{f_K\}_{K > m}$ defined via (A.7). Since $f_K \in C_c^2(\mathbb{R}^n)$, condition (**H1**) implies that the process $M^{f_K} = (M_t^{f_K})_{t \in \mathbb{R}_+}$, where

$$M_t^{f_K} := f_K(\mathbf{X}_t^{\boldsymbol{\alpha^x}}) - \int_0^t \left(L^{\boldsymbol{\alpha^x}}_s f_K\right)(\mathbf{X}_{s-}^{\boldsymbol{\alpha^x}}) \, \mathrm{d}s, \quad t \in \mathbb{R}_+, \quad (\text{A.22})$$

is a local $(\mathbb{F}^\alpha, \mathbb{P}^\alpha)$-martingale.

Let $\tau_m := T_m \wedge S_m$, where T_m and S_m are defined via (A.8) and (A.9), respectively. Note that the stopped process $M_\cdot^{f_K, \tau_m} := M_{\cdot \wedge \tau_m}^{f_K}$ is also a local $(\mathbb{F}^\alpha, \mathbb{P}^\alpha)$-martingale [26, Corollary 3.6, Chapter II, p. 71]. Since (by construction) $f_K \left(\mathbf{X}_{t \wedge \tau_m}^{\boldsymbol{\alpha}^{\mathbf{x}}} \right) \to f(\mathbf{X}_{t \wedge \tau_m}^{\boldsymbol{\alpha}^{\mathbf{x}}})$ as $K \to \infty$ and, further, (A.14) in Lemma A.1 also holds, we obtain that $M_t^{f, \tau_m} = \lim_{K \to \infty} M_t^{f_K, \tau_m}$ $\mathbb{P}^\alpha - a.s.$ for all $t \in \mathbb{R}_+$, where $M_\cdot^{f, \tau_m} := M_{\cdot \wedge \tau_m}^f$ is the stopped version of the process in (A.21).

Let us now prove that, for each m, the process M^{f, τ_m} is a local $(\mathbb{F}^\alpha, \mathbb{P}^\alpha)$-martingalem. Take $s, t \in \mathbb{R}_+$, $s < t$ and $B \in \mathcal{F}_s^\alpha$. It is sufficient to show that $\mathbb{E}_{\mathbf{x}}^\alpha \left[(M_t^{f, \tau_m} - M_s^{f, \tau_m}) 1_B \right] = 0$. Since

$$\left(M_t^{f, \tau_m} - M_s^{f, \tau_m} \right) 1_B = \left(M_t^{f, \tau_m} - M_t^{f_K, \tau_m} \right) 1_B + \left(M_t^{f_K, \tau_m} - M_s^{f_K, \tau_m} \right) 1_B + \left(M_s^{f_K, \tau_m} - M_s^{f, \tau_m} \right) 1_B, \tag{A.23}$$

by taking expectations and using that, for each K and m, the second term vanishes because M^{f_K, τ_m} is a true $(\mathbb{F}^\alpha, \mathbb{P}^\alpha)$-martingale, we only need to prove that, for each t, $M_t^{f_K, \tau_m} \to M_t^{f, \tau_m}$ in $L^1(\mathbb{P}^\alpha)$ as $K \to \infty$.

Using (A.13) in Lemma A.1, we get that, for fixed $t \in \mathbb{R}_+$ and $m \in \mathbb{N}$, $\sup_K \left| M_t^{f_K, \tau_m} \right| \leq |f| + mtC(m, f)$ (because f is bdd, $|f_K| \leq |f|$ and by definition of τ_m). Therefore, for each $t \in \mathbb{R}_+$ and $m \in \mathbb{N}$, the family of r.v.'s $\mathfrak{M} := \left\{ M_t^{f_K, \tau_m} : K > m \right\}$ is uniformly integrable (UI) (see, for example, [25, Chapter 1, p.8]).

Therefore, there exists a process M^m such that $\mathbb{P}^\alpha - a.s.$ $M_{t \wedge \tau_m}^{f_K} \to M_t^m$ for all $t \in \mathbb{R}_+$, as $K \to \infty$. It follows that M_t^m is integrable for each $t \in \mathbb{R}_+$ and, further, $M_{t \wedge \tau_m}^{f_K} \to M_t^m$ in $L^1(\mathbb{P}^\alpha)$ as $K \to \infty$. The previous implies then that the stopped process M^{f, τ_m} is an $(\mathbb{F}^\alpha, \mathbb{P}^\alpha)$-adapted martingale. Hence, by Theorem 50 in [25, Chapter I, p. 38], we conclude that M^f is an $(\mathbb{F}^\alpha, \mathbb{P}^\alpha)$-local martingale, as required. ∎

A.3. Proof of Proposition 4.4.

Proof. Let $(\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}^{\mathbf{x}})$ be an admissible pair. Observe first that $\mathbf{B}$ (resp. $\mathbf{C}$) is predictable as, by definition, it is a Lebesgue integral of the locally integrable processes $\boldsymbol{\mu}$ and ν (resp. $\boldsymbol{\sigma}$). By (**H2**), for each $t \geq r$, $\int_r^t Q_s^{p, \boldsymbol{\alpha}^{r, \xi}} \, \mathrm{d}s$ is the Lebesgue integral of a locally integrable process, hence the dominated convergence theorem (DCT) implies that $\mathbb{P}^\alpha - a.s.$ the paths $t \mapsto \int_r^t Q_s^{p, \boldsymbol{\alpha}^{r, \xi}} \, \mathrm{d}s$ are continuous. Thus, the process $\int_0^\cdot Q_s^{p, \boldsymbol{\alpha}^{r, \xi}} \, \mathrm{d}s$ is predictable. As for the random measure η, by [10, Definition 1.6, Chapter II.1a, p. 66] we need to prove that, for any predictable function $W(\omega, s, \mathbf{x})$ on $\Omega^\alpha \times \mathbb{R}_+ \times \mathbb{R}_0^n$, the integral process $W * \eta$ is also predictable, where

$$W * \eta_t(\omega) := \int_{(0, t] \times \mathbb{R}_0^n} W(\omega, s, \mathbf{x}) \eta(\omega; \mathrm{d}s, \mathrm{d}\mathbf{x}) \tag{A.24}$$

if $\int_{[0, t] \times \mathbb{R}_0^n} |W(\omega, s, \mathbf{x})| \eta(\omega; \mathrm{d}s, \mathrm{d}\mathbf{x})$ is finite, and equal to $+\infty$ otherwise. Since $\eta(\mathrm{d}s, \mathrm{d}\mathbf{y}) = \mathrm{d}s \otimes \nu_s(\mathrm{d}\mathbf{y})$, for each t and ω, the integral $W * \eta_t(\omega)$ is a Lebesgue integral of the product of two predictable processes: W and $\boldsymbol{\sigma}$. Hence, $W * \eta$ is predictable.

Therefore, by [10, Theorem II. 2.42 p. 86], we only need to show that, for each bounded function $f \in C^2(\mathbb{R}^d)$, the process

$$N_t^{f, \boldsymbol{\alpha}^{\mathbf{x}}} := f(\mathbf{X}_t^{\boldsymbol{\alpha}^{\mathbf{x}}}) - f(\mathbf{X}_0^{\boldsymbol{\alpha}^{\mathbf{x}}}) - \int_0^t \sum_{i=1}^d \partial_i f(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{\mathbf{x}}}) \, \mathrm{d}B_s^j - \frac{1}{2} \int_0^t \sum_{i,j=1}^d \partial_{ij}^2 f(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{\mathbf{x}}}) \, \mathrm{d}C_{ij}(s)$$

$$- \int_{[0, t] \times \mathbb{R}_0^n} f(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{\mathbf{x}}} + \mathbf{y}) - f(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{\mathbf{x}}}) - h(\mathbf{y}) \cdot \nabla f(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{\mathbf{x}}}) \eta(\mathrm{d}s, \mathrm{d}\mathbf{y})$$

is a local martingale. By the definition of $L^{\boldsymbol{\alpha}_s^{\mathbf{x}}}$ and η, we have the equality

$$N_t^{f, \boldsymbol{\alpha}^{\mathbf{x}}} = f(\mathbf{X}_t^{\boldsymbol{\alpha}^{\mathbf{x}}}) - f(\mathbf{X}_0^{\boldsymbol{\alpha}^{\mathbf{x}}}) - \int_0^t \left(L^{\boldsymbol{\alpha}_s^{\mathbf{x}}} f \right) (\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{\mathbf{x}}}) \, \mathrm{d}s, \quad t \in \mathbb{R}_+.$$

Therefore, the result follows from Theorem A.3. ∎

A.4. **Proof of Proposition 4.8.** For the proof of this result, we will need the following auxiliary lemma.

Lemma A.4. *Theorem A.3 is also valid for any function* $f \in C^2(\mathbb{R}^n)$ *satisfying the polynomial growth* $|f(\mathbf{x})| \le C(|\mathbf{x}|^q + 1)$ *for some* $C > 0$ *and* $q \in [2, p]$.

Proof. Let $(\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}^{\mathbf{x}})$ be an admissible pair with $\boldsymbol{\alpha}^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}^p$ and let $f \in C^2(\mathbb{R}^n)$ be as in the statement. Define

$$W(\omega, s, \mathbf{y}) := f(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{\mathbf{x}}} + \mathbf{y}) - f(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{\mathbf{x}}}) - \sum_{i=1}^n \partial_i f(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{\mathbf{x}}}) y_i, \quad (\omega, s, \mathbf{y}) \in \Omega^\alpha \times \mathbb{R}_+ \times \mathbb{R}_0^n.$$

Notice that all processes $f(\mathbf{X}_-^{\boldsymbol{\alpha}^{\mathbf{x}}} + \mathbf{y})$, $f(\mathbf{X}_-^{\boldsymbol{\alpha}^{\mathbf{x}}})$ and $\partial_i f(\mathbf{X}_-^{\boldsymbol{\alpha}^{\mathbf{x}}})$ are left-continuous with right limits, so they are locally bounded and predictable. The latter implies that, for each $\mathbf{y}$, the process W is also predictable. Since (by Theorem **??**) $\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}$ is a semimartingale, the generalised Itô formula [10, Theorem 4.57, Chapter I.4e, p. 57] implies that $f(\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}})$ is also a semimartingale satisfying

$$f(\mathbf{X}_t^{\boldsymbol{\alpha}^{\mathbf{x}}}) = f(\mathbf{x}) + \sum_{i=1}^n \int_{0+}^t \partial_i f(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{\mathbf{x}}}) \, \mathrm{d}X_s^i + \frac{1}{2} \sum_{1 \le i,j \le n} \int_{0+}^t \partial_{ij} f(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{\mathbf{x}}}) \, \mathrm{d}\langle X^{i,c}, X^{j,c} \rangle + W * \eta_t^X$$

$$= f(\mathbf{x}) + \sum_{i=1}^n \int_{0+}^t \partial_i f(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{\mathbf{x}}})(u^i + \mu_i(s)) \, \mathrm{d}s + \frac{1}{2} \sum_{1 \le i,j \le n} \int_{0+}^t \partial_{ij} f(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{\mathbf{x}}})(\boldsymbol{\sigma}_s^T \boldsymbol{\sigma}_s)_{ij} \, \mathrm{d}s + W * \eta_t + N_t$$

$$= f(\mathbf{x}) + \int_0^t L^{\boldsymbol{\alpha}_s^{\mathbf{x}}} f\left(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{\mathbf{x}}}\right) + N_t \tag{A.25}$$

where $N_t := \sum_{i=1}^n \int_{0+}^t \partial_i f(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{\mathbf{x}}}) \, \mathrm{d}M_t^i + W * (\eta^X - \eta)_t$ is a local martingale (recall that M^i is the local martingale in the decomposition of X^i and η is the compensator of the random measure η^X). Second equality follows from Theorem **??** and the third equality from the definition of $L^{\boldsymbol{\alpha}_s^{\mathbf{x}}}$ and the fact that $W * \eta_t(\omega) = \int_{[0,t] \times \mathbb{R}_0^n} W(\omega, s, \mathbf{y}) \, \mathrm{d}s \otimes \nu_s(\omega; \mathrm{d}\mathbf{y})$ (see definition in (A.24)). Let us observe that f having a polynomial growth of degree $q \le p$ with $p \ge 2$ is a key assumption to guarantee that $W * \eta_t$ is well-defined. The latter holds because ν_s takes values in $\mathcal{M}_p$ and thus it is a measure with finite second moments inside the unitary ball B_1 and finite pth-moments in B_1^c. This concludes the proof. ∎

Proof. (of Proposition 4.8)

(i) Equality (4.3) is the canonical representation (relative to h) for special semimartingales (see [10, II.2c, Theorem 2.34, p.84]) and follows from Proposition 4.4 and [10, II.2c, Corollary 2.38, p.84].

(ii) Since $\left|\mathbf{X}_s^{\boldsymbol{\alpha}^{\mathbf{x}},c}\right|^q \le C \sum_{i=1}^n \left|X_s^{i,c}\right|^q$ and (by Corollary (4.4)) the equality $\langle X^{i,c}, X^{i,c} \rangle_t = \int_0^t \sum_{k=1}^n \sigma_{ik}^2(s) \, \mathrm{d}s$ holds, the Burkholder-Davis-Gundy inequality and the fact that $\sum_{k=1}^n |a_i|^r \le c \left(\sum_{k=1}^n |a_i|\right)^r$ for $r > 1$ and some constant $c > 0$, imply

$$\mathbb{E}_{\mathbf{x}}^\alpha \sup_{0 \le s \le t} \left|\mathbf{X}_s^{\boldsymbol{\alpha}^{\mathbf{x}},c}\right|^q \le C \mathbb{E}_{\mathbf{x}}^\alpha \left[\sum_{i=1}^n \left(\sum_{k=1}^n \int_0^t \sigma_{ik}^2(s) \, \mathrm{d}s\right)^{q/2}\right]$$

$$\le C \mathbb{E}_{\mathbf{x}}^\alpha \left[\left(\sum_{i=1}^n \sum_{k=1}^n \int_0^t \sigma_{ik}^2(s) \, \mathrm{d}s\right)^{q/2}\right] = C \mathbb{E}_{\mathbf{x}}^\alpha \left(\int_0^t \|\boldsymbol{\sigma}_s\|^2 \, \mathrm{d}s\right)^{q/2},$$

as required.

iii) To deal with the running maximum of the discontinuous martingale part of $\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}$, we consider the controlled process $\mathbf{Y}$ obtained by taking the policy $\beta = (0, \nu, -\mathbf{u}) \in \mathcal{A}_{\mathbf{x}}^p$, where ν is the same

process in the control $\boldsymbol{\alpha}^{\mathbf{x}} = (\boldsymbol{\sigma}, \nu, \boldsymbol{\sigma})$ and $\mathbf{u}$ is the vector corresponding to the operator L^a defined in (3.1). Therefore, the statement (i) proved above implies that

$$\mathbf{Y}_t = \mathbf{x} + \int_0^t \int_{\mathbb{R}_0^n} \mathbf{y} \left(\eta^{Y,\beta} - \eta \right) (\mathrm{d}s, \mathrm{d}\mathbf{y}), \tag{A.26}$$

where $\eta^{Y,\beta}$ is the integer-valued random measure associated with the jumps of $\mathbf{Y}$ and the random measure η is its predictable compensator. Hence, $\mathbf{Y}$ is a local martingale and, by Corollary (4.4), $\eta(\omega, \mathrm{d}s, \mathrm{d}\mathbf{y}) = \mathrm{d}s \otimes \nu_s(\omega, \mathrm{d}\mathbf{y})$. Thus, to obtain the estimate for the process $\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}, d}$, we only need to estimate $|\mathbf{Y}_t|^q$ for $\mathbf{x} = 0$.

Define

$$F(\omega, s, \mathbf{y}) := |\mathbf{Y}_{s-} + \mathbf{y}|^q - |\mathbf{Y}_{s-}|^q - \mathbf{y}^T \nabla |\mathbf{Y}_{s-}|^q, \quad (\omega, s, \mathbf{y}) \in \Omega^\alpha \times \mathbb{R}_+ \times \mathbb{R}_0^n, \tag{A.27}$$

and

$$G(\omega, s, \mathbf{y}) := \mathbf{y}^T \nabla |\mathbf{Y}_{s-}|^q \quad (\omega, s, \mathbf{y}) \in \Omega^\alpha \times \mathbb{R}_+ \times \mathbb{R}_0^n. \tag{A.28}$$

Observe that the processes $|\mathbf{Y}_-|^q$ and $\nabla |\mathbf{Y}_-|^q$ are left-continuous with right limits. The previous implies that, for each $\mathbf{y}$, both processes F and G are also predictable.

Itô 's formula applied to $h : \mathbf{y} \mapsto |\mathbf{y}|^q$ implies that

$$|\mathbf{Y}_t|^q = G \star (\eta^{Y,\beta} - \eta)_t + F \star \eta_t^{Y,\beta} = |\mathbf{x}|^q + G \star (\eta^{Y,\beta} - \eta)_t + F \star (\eta^{Y,\beta} - \eta)_t + F \star \eta_t, \tag{A.29}$$

where we have used that η is the compensator of $\eta^{Y,\beta}$. Notation $\star$ stands for the stochastic integral defined in (A.24).

Observe now that the process N defined by

$$\begin{aligned} N_t &:= (F + G) \star (\eta^{Y,\beta} - \eta)_t \\ &= \int_0^t \int_{\mathbb{R}_0^n} \left(|\mathbf{Y}_{s-} + \mathbf{y}|^q - |\mathbf{Y}_{s-}|^q \right) (\eta^{Y,\beta} - \eta)(\mathrm{d}s, \mathrm{d}\mathbf{y}), \end{aligned}$$

is a local martingale. Without loss of generality, let us assume that $N = |F + G| \star (\eta^{Y,\beta} - \eta)$ is a true martingale (otherwise one can proceed by considering an appropriate localising sequence).

Hence, since $\sup_{0 \le s \le t} |N_s| \le |F + G| \star (\eta^{Y,\beta} - \eta)_t$, we obtain that $\mathbb{E}_0^\beta \left(\sup_{0 \le s \le t} |N_s| \right) = 0$ and, thus, the equality (A.29) implies

$$\mathbb{E}_0^\beta \left(\sup_{0 \le s \le t} |\mathbf{Y}_s|^q \right) \le \mathbb{E}_0^\beta \left(\sup_{0 \le s \le t} F \star \eta_s \right). \tag{A.30}$$

To estimate the right hand side above, we can now proceed as in the proof of the Kunita's inequalities for Lévy -type stochastic integrals given in [2, Theorem 4.4.23, p. 265]. Namely, using the definition of F and Taylor's theorem one can find $\theta \in (0, 1)$ such that

$$\begin{aligned} F \star \eta_t &= \int_0^t \mathrm{d}s \int_{\mathbb{R}_0^n} \left(|\mathbf{Y}_{s-} + \mathbf{y}|^q + |\mathbf{Y}_{s-}|^q + \sum_{i=1}^n q |\mathbf{Y}_{s-}|^{q-2} \mathbf{Y}_{s-} y_i \right) \nu_s(\mathrm{d}\mathbf{y}) \\ &\le \frac{1}{2} \int_0^t \mathrm{d}s \int_{\mathbb{R}_0^n} \sum_{1 \le i,j \le n} \partial_{ij}^2 |\mathbf{Y}_{s-} + \theta \mathbf{y}|^q |y_i y_j| \nu_s(\mathrm{d}\mathbf{y}) \\ &\le \frac{1}{4} \int_0^t \mathrm{d}s \int_{\mathbb{R}_0^n} \sum_{1 \le i,j \le n} \left(\delta_{ij} q |\mathbf{Y}_{s-} + \theta \mathbf{y}|^{q-2} + q(q-2) |\mathbf{Y}_{s-} + \theta y_i| |\mathbf{Y}_{s-} + \theta y_j| |\mathbf{Y}_{s-}|^{q-4} \right) |\mathbf{y}|^2 \nu_s(\mathrm{d}\mathbf{y}) \\ &\le C \int_0^t \mathrm{d}s \int_{\mathbb{R}_0^n} \left(|\mathbf{Y}_{s-}| + |\mathbf{y}| \right)^{q-2} |\mathbf{y}|^2 \nu_s(\mathrm{d}\mathbf{y}) \\ &\le C \int_0^t \mathrm{d}s \int_{\mathbb{R}_0^n} \left(|\mathbf{Y}_{s-}|^{q-2} + |\mathbf{y}|^{q-2} \right) |\mathbf{y}|^2 \nu_s(\mathrm{d}\mathbf{y}) \\ &\le C \left\{ \int_0^t \mathrm{d}s \int_{|\mathbf{y}| \le 1} \left(|\mathbf{Y}_{s-}|^{q-2} + 1 \right) |\mathbf{y}|^2 \nu_s(\mathrm{d}\mathbf{y}) + \int_0^t \mathrm{d}s \int_{|\mathbf{y}| \ge 1} \left(|\mathbf{Y}_{s-}|^{q-2} + |\mathbf{y}|^{q-2} \right) |\mathbf{y}|^2 \nu_s(\mathrm{d}\mathbf{y}) \right\} \end{aligned}$$

Therefore, on taking expectations we get $\mathbb{E}_0^\beta \left(\sup_{0 \le s \le t} F \ast \eta_s \right) \le H_1 + H_2$, where

$$H_1 := C\, \mathbb{E}_0^\beta \left\{ \sup_{0 \le s \le t} \frac{1}{\gamma} |\mathbf{Y}_{s-}|^{q-2} \int_0^t \mathrm{d}s \int_{\mathbb{R}_0^n} \gamma |\mathbf{y}|^2 \nu_s(\mathrm{d}\mathbf{y}) \right\},$$

$$H_2 := C\, \mathbb{E}_0^\beta \left\{ \int_0^t \mathrm{d}s \int_{\mathbb{R}_0^n} |\mathbf{y}|^2 \vee |\mathbf{y}|^q \, \nu_s(\mathrm{d}\mathbf{y}) \right\},$$

for any $\gamma > 1$. Using Hölder 's inequality with the conjugate values $p' = q/(q-2)$ and $q' = q/2$, it follows that

$$H_1 \le \frac{C}{\gamma} \left\{ \mathbb{E}_0^\beta \sup_{0 \le s \le t} |\mathbf{Y}_{s-}|^q \right\}^{1-2/q} \left\{ \mathbb{E}_0^\beta \left(\int_0^t \mathrm{d}s \int_{\mathbb{R}_0^n} \gamma |\mathbf{y}|^2 \nu_s(\mathrm{d}\mathbf{y}) \right)^{q/2} \right\}^{2/q}$$

$$\le \frac{(q-2)C}{\gamma q} \mathbb{E}_0^\beta \left(\sup_{0 \le s \le t} |\mathbf{Y}_{s-}|^q \right) + \frac{2C\gamma^{q/2}}{q} \mathbb{E}_0^\beta \left(\int_0^t \mathrm{d}s \int_{\mathbb{R}_0^n} \gamma |\mathbf{y}|^2 \nu_s(\mathrm{d}\mathbf{y}) \right)^{q/2}$$

$$\le \frac{2C\gamma^{q/2}}{q} \mathbb{E}_0^\beta \left(\int_0^t \mathrm{d}s \int_{\mathbb{R}_0^n} \gamma |\mathbf{y}|^2 \nu_s(\mathrm{d}\mathbf{y}) \right)^{q/2}$$

whenever γ is chosen to satisfy $(q-2)C < \gamma q$. Using the previous estimates into (A.30) yields

$$\mathbb{E}_0^\beta \left(\sup_{0 \le s \le t} |\mathbf{Y}_s|^q \right) \le C_1 \left\{ \mathbb{E}_0^\beta \left(\int_0^t \mathrm{d}s \int_{\mathbb{R}_0^n} |\mathbf{y}|^2 \nu_s(\mathrm{d}\mathbf{y}) \right)^{q/2} + \mathbb{E}_0^\beta \left(\int_0^t \mathrm{d}s \int_{\mathbb{R}_0^n} |\mathbf{y}|^2 \vee |\mathbf{y}|^q \, \nu_s(\mathrm{d}\mathbf{y}) \right) \right\},$$

for some positive constant $C_1 > 0$. Applying the previous result to $\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}, d}$ and rearranging terms, we obtain the inequality required in (4.6). ∎

Appendix B. Proofs of Results in Sections 5 and 6

B.1. **Proof of Lemma 5.2.**

Proof. Take $\boldsymbol{\alpha}^{r,\mathbf{x}} \in \mathcal{A}_{r,\mathbf{x}}^p$, then a simple change of variables yields

$$J^\alpha(r, \mathbf{x}) = \mathbb{E}\left[\int_0^\infty e^{-\int_0^u q(\mathbf{X}_{m+r}^\alpha, \boldsymbol{\alpha}_{m+r}) \mathrm{d}m} f\left(\mathbf{X}_{u+r}^\alpha, \boldsymbol{\alpha}_{u+r} \right) \mathrm{d}u \Big| \mathbf{X}_r^\alpha = \mathbf{x} \right]$$

$$= \mathbb{E}\left[\int_0^\infty e^{-\int_0^u q(\mathbf{Y}_m^{\tilde{\alpha}}, \tilde{\boldsymbol{\alpha}}_m) \mathrm{d}m} f\left(\mathbf{Y}_u^{\tilde{\alpha}}, \tilde{\boldsymbol{\alpha}}_u \right) \mathrm{d}u \Big| \mathbf{Y}_0^{\tilde{\alpha}} = \mathbf{x} \right] = J^{\tilde{\alpha}}(\mathbf{x}), \tag{B.1}$$

where $\left(\mathbf{Y}^{\tilde{\boldsymbol{\alpha}}^{\mathbf{x}}}, \tilde{\boldsymbol{\alpha}}^{\mathbf{x}} \right) = \left(\mathbf{X}_{\cdot+r}^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}_{\cdot+r}^{\mathbf{x}} \right)$. Using that $\boldsymbol{\alpha}^{r,\mathbf{x}} \in \mathcal{A}_{r,\mathbf{x}}^p$, the definition of $\tilde{\boldsymbol{\alpha}}^{\mathbf{x}}$ implies that $\tilde{\boldsymbol{\alpha}}^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}^p$ with corresponding controlled process $\mathbf{Y}^{\tilde{\boldsymbol{\alpha}}^{\mathbf{x}}}$. Indeed, the construction of the pair $\left(\mathbf{Y}^{\tilde{\boldsymbol{\alpha}}^{\mathbf{x}}}, \tilde{\boldsymbol{\alpha}}^{\mathbf{x}} \right)$ is obtained from the corresponding canonical process $\left(\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}^{\mathbf{x}} \right)$ by shifting appropriately. The validity of (3.2) and (3.3) follow straightforwardly from the corresponding conditions on $\left(\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}^{\mathbf{x}} \right)$. Therefore, taking the infimum over $\boldsymbol{\alpha}^{r,\mathbf{x}} \in \mathcal{A}_{r,\mathbf{x}}^p$ in the first equality of (B.1) and then taking the infimum over $\hat{\alpha} \in \mathcal{A}_{\mathbf{x}}^p$ in the second equality of (B.1), yields $v(r, \mathbf{x}) = v(0, \mathbf{x}) = V(\mathbf{x})$, as required. ∎

B.2. **Proof of Lemma 5.5.**

Proof. Denote by $W(\mathbf{x})$ the right-hand side in (5.8). Let us first prove the inequality $V \le W$. Let $(\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}^{\mathbf{x}})$ be an admissible pair and define $\eta_t^\alpha := \mathbb{P}^{\boldsymbol{\alpha}} \circ \left(\mathbf{X}_t^{\boldsymbol{\alpha}^{\mathbf{x}}} \right)^{-1}$. Take $\epsilon > 0$ and let $\boldsymbol{\beta}_t^\alpha$ be an ϵ-optimal control in $\mathcal{A}_{t, \eta_t^\alpha}^p$ with corresponding control process $\mathbf{Y}^{\boldsymbol{\beta}_t^\alpha}$. Define $\mathbf{u}^{\mathbf{x}, \oplus_t} := \boldsymbol{\alpha}^{\mathbf{x}} \oplus_t \boldsymbol{\beta}_t^\alpha$ as given in (3.4). Since, by Lemma 3.4, $\mathbf{u}^{\mathbf{x}, \oplus_t}$ is also an admissible control in $\mathcal{A}_{\mathbf{x}}^p$, there exists $(\Omega^u, \mathcal{F}^u, (\mathcal{F}_t^u), \mathbb{P}^u)$ in which the corresponding admissible pair $(\mathbf{Z}^{\mathbf{u}^{\mathbf{x}, \oplus_t}}, \mathbf{u}^{\mathbf{x}, \oplus_t})$ is defined.

Set $J^u(\cdot) := J(\mathbf{Z}^{\mathbf{u}^{\cdot,\oplus t}}, \mathbf{u}^{\cdot,\oplus t})$. Then $V(\mathbf{x}) \leq J^u(\mathbf{x}) = A + B$, where

$$A := \mathbb{E}_{\mathbf{x}}^{\mathbf{u}}\left[\int_0^t e^{-\int_0^s q\left(\mathbf{Z}_l^{\oplus t}, \mathbf{u}_l^{\oplus t}\right)\mathrm{d}l} f\left(\mathbf{Z}_s^{\oplus t}, \mathbf{u}_s^{\oplus t}\right)\mathrm{d}s\right],$$

$$B := \mathbb{E}_{\mathbf{x}}^{\mathbf{u}}\left[\int_t^\infty e^{-\int_0^s q\left(\mathbf{Z}_l^{\oplus t}, \mathbf{u}_l^{\oplus t}\right)\mathrm{d}l} f\left(\mathbf{Z}_s^{\oplus t}, \mathbf{u}_s^{\oplus t}\right)\mathrm{d}s\right]. \tag{B.2}$$

Since the control process $\mathbf{Z}^{\mathbf{u}^{\mathbf{x},\oplus t}}$ satisfies that $\mathbf{Z}^{\mathbf{u}^{\mathbf{x},\oplus t}} = \mathbf{X}^{\alpha^{\mathbf{x}}}$ on $[0,t)$, we have

$$A = \mathbb{E}_{\mathbf{x}}^{\alpha}\left[\int_0^t e^{-\int_0^s q(\mathbf{X}_l^{\alpha}, \alpha_l)\mathrm{d}l} f\left(\mathbf{X}_s^{\alpha}, \alpha_s\right)\mathrm{d}s\right], \tag{B.3}$$

whereas the equality $\mathbf{Z}^{\mathbf{u}^{\mathbf{x},\oplus t}} = \mathbf{Y}^{\beta_t^{\alpha}}$ on $[t,\infty)$ and properties of conditional expectation yield

$$B = \mathbb{E}_{\mathbf{x}}^{\mathbf{u}}\left[e^{-\int_0^t q\left(\mathbf{Z}_l^{\oplus t}, \mathbf{u}_l^{\oplus t}\right)\mathrm{d}l} \int_t^\infty e^{-\int_t^s q\left(\mathbf{Z}_l^{\oplus t}, \mathbf{u}_l^{\oplus t}\right)\mathrm{d}l} f\left(\mathbf{Z}_s^{\oplus t}, \mathbf{u}_s^{\oplus t}\right)\mathrm{d}s\right]$$

$$= \mathbb{E}_{\mathbf{x}}^{\mathbf{u}}\left[e^{-\int_0^t q\left(\mathbf{Z}_l^{\oplus t}, \mathbf{u}_l^{\oplus t}\right)\mathrm{d}l}\mathbb{E}_{\mathbf{x}}^{\mathbf{u}}\left(\int_t^\infty e^{-\int_t^s q\left(\mathbf{Z}_l^{\oplus t}, \mathbf{u}_l^{\oplus t}\right)\mathrm{d}l} f\left(\mathbf{Z}_s^{\oplus t}, \mathbf{u}_s^{\oplus t}\right)\mathrm{d}s\Big|\mathcal{F}_t^u\right)\right]$$

$$= \mathbb{E}_{\mathbf{x}}^{\mathbf{u}}\left[e^{-\int_0^t q\left(\mathbf{Z}_l^{\oplus t}, \mathbf{u}_l^{\oplus t}\right)\mathrm{d}l}\mathbb{E}^{\beta_t^{\alpha}}\left(\int_t^\infty e^{-\int_t^s q\left(\mathbf{Y}_l^{\beta_t^{\alpha}}, \beta_t^{\alpha}(l)\right)\mathrm{d}l} f\left(\mathbf{Y}_s^{\beta_t^{\alpha}}, \beta_t^{\alpha}(s)\right)\mathrm{d}s\Big|\mathbf{X}_t^{\alpha}\right)\right]$$

$$= \mathbb{E}_{\mathbf{x}}^{\alpha}\left[e^{-\int_0^t q(\mathbf{X}_l^{\alpha}, \alpha_l)\mathrm{d}l} J\left(\tilde{\mathbf{Y}}^{\tilde{\beta}^{\alpha}}, \tilde{\beta}^{\alpha}\right)\right], \tag{B.4}$$

where $\left(\tilde{\mathbf{Y}}_{\cdot}^{\tilde{\beta}^{\alpha}}, \tilde{\beta}_{\cdot}^{\alpha}\right) := \left(\mathbf{Y}_{\cdot+t}^{\beta_t^{\alpha}}, \beta_t^{\alpha}(\cdot + t)\right)$. Similarly as we did in the proof of Lemma 5.2, we obtain that the process $\tilde{\beta}^{\alpha}$ is admissible and belongs to $\mathcal{A}_{\eta_t^{\alpha}}^p$. Moreover, it is not difficult to see that $\tilde{\beta}^{\alpha}$ is an ϵ-optimal policy as well. Hence, (B.3) and (B.4),together with the equality $\gamma_t^{\alpha^{\mathbf{x}}} := \exp\{-\int_0^s q\left(\mathbf{X}_l^{\alpha}, \alpha_l\right)\mathrm{d}l\}$, imply that

$$V(\mathbf{x}) < A + \mathbb{E}_{\mathbf{x}}^{\alpha}\left[e^{-\gamma_t^{\alpha^{\mathbf{x}}}} V\left(\mathbf{X}_t^{\alpha}\right)\right] + \epsilon = W(\mathbf{x}) + \epsilon. \tag{B.5}$$

Letting $\epsilon \downarrow 0$ and then taking the infimum over all policies in $\mathcal{A}_{\mathbf{x}}^p$ yield the desired inequality $V \leq W$.

To prove the reverse inequality, take an ϵ-optimal policy $\alpha^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}^p$. Again, properties of conditional expectation yield

$$V(\mathbf{x}) + \epsilon > J^{\alpha}(\mathbf{x}) \tag{B.6}$$

$$= \mathbb{E}_{\mathbf{x}}^{\alpha}\left[\left(\int_0^t + \int_t^\infty\right) e^{-\gamma_s^{\alpha}} f\left(\mathbf{X}_s^{\alpha}, \alpha_s\right)\mathrm{d}s\right]$$

$$= \mathbb{E}_{\mathbf{x}}^{\alpha}\left[\int_0^t e^{-\gamma_s^{\alpha}} f\left(\mathbf{X}_s^{\alpha}, \alpha_s\right)\mathrm{d}s\right] + \mathbb{E}_{\mathbf{x}}^{\alpha}\left[e^{-\gamma_t^{\alpha}}\mathbb{E}_{\mathbf{x}}^{\alpha}\left[\int_t^\infty e^{-\int_t^s q(\mathbf{X}_l^{\alpha}, \alpha_l)\mathrm{d}l} f\left(\mathbf{X}_s^{\alpha}, \alpha_s\right)\mathrm{d}s\Big|\mathcal{F}_t^{\alpha}\right]\right]$$

$$= \mathbb{E}_{\mathbf{x}}^{\alpha}\left[\int_0^t e^{-\gamma_s^{\alpha}} f\left(\mathbf{X}_s^{\alpha}, \alpha_s\right)\mathrm{d}s\right] + \mathbb{E}_{\mathbf{x}}^{\alpha}\left[e^{-\gamma_t^{\alpha}}\mathbb{E}_{\mathbf{x}}^{\alpha}\left[\int_t^\infty e^{-\int_t^s q(\mathbf{X}_l^{\alpha}, \alpha_l)\mathrm{d}l} f\left(\mathbf{X}_s^{\alpha}, \alpha_s\right)\mathrm{d}s\Big|\mathbf{X}_t^{\alpha}\right]\right]$$

$$= \mathbb{E}_{\mathbf{x}}^{\alpha}\left[\int_0^t e^{-\gamma_s^{\alpha}} f\left(\mathbf{X}_s^{\alpha}, \alpha_s\right)\mathrm{d}s\right] + \mathbb{E}_{\mathbf{x}}^{\alpha}\left[e^{-\gamma_t^{\alpha}} J\left(\tilde{\mathbf{X}}^{\tilde{\alpha}}, \tilde{\alpha}\right)\right]$$

$$\geq \mathbb{E}_{\mathbf{x}}^{\alpha}\left[\int_0^t e^{-\gamma_s^{\alpha}} f\left(\mathbf{X}_s^{\alpha}, \alpha_s\right)\mathrm{d}s\right] + \mathbb{E}_{\mathbf{x}}^{\alpha}\left[e^{-\gamma_t^{\alpha}} V\left(\mathbf{X}_t^{\alpha}\right)\right] = W(\mathbf{x}),$$

where $\left(\tilde{\mathbf{X}}_{\cdot}^{\tilde{\alpha}}, \tilde{\alpha}_{\cdot}\right) := (\mathbf{X}_{\cdot+t}^{\alpha}, \beta_t^{\alpha}(\cdot + t))$ is an admissible pair with $\tilde{\alpha} \in \mathcal{A}_{\eta_t^{\alpha}}^p$, η_t^{α} being the law of $\mathbf{X}_t^{\alpha^{\mathbf{x}}}$. By letting $\epsilon \downarrow 0$, we get $V \geq W$, as required. ∎

B.3. Proof of Lemma 5.6. .

Proof. We need to prove that for any admissible $\alpha^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}^p$, the inequality $\mathbb{E}_{\mathbf{x}}^{\alpha}\left[S_t^{V,\alpha}\big|\mathcal{F}_s^{\alpha}\right] \geq S_s^{V,\alpha}$ holds for all $t \geq s \geq 0$.

Define $\theta_\alpha(r,t) := e^{-\int_r^t q\left(\mathbf{X}_s^{\alpha^{\mathbf{x}}}, \alpha_s^{\mathbf{x}}\right)\mathrm{d}s}$ for $0 \le r \le t$. Notice that $\theta_\alpha(0,t) = e^{-\gamma_t^{\alpha^{\mathbf{x}}}}$. Now, fix some arbitrary admissible control $\boldsymbol{\alpha}^{\mathbf{x}}$, then

$$
\begin{aligned}
\mathbb{E}_{\mathbf{x}}^\alpha\left[S_t^{V,\boldsymbol{\alpha}}\Big|\mathcal{F}_s^\alpha\right] &= \mathbb{E}_{\mathbf{x}}^\alpha\left[\left(\int_0^s + \int_s^t\right)\theta_\alpha(0,s)f\left(\mathbf{X}_r^{\boldsymbol{\alpha}},\boldsymbol{\alpha}_r\right)\mathrm{d}r + \theta_\alpha(0,t)V\left(\mathbf{X}_t^{\boldsymbol{\alpha}}\right)\right] \\
&= \int_0^s \theta_\alpha(0,r)f\left(\mathbf{X}_r^{\boldsymbol{\alpha}},\boldsymbol{\alpha}_r\right)\mathrm{d}r + \\
&\quad + \theta_\alpha(0,s)\mathbb{E}\left[\int_s^t \theta_\alpha(s,r)f\left(\mathbf{X}_r^{\boldsymbol{\alpha}},\boldsymbol{\alpha}_r\right)\mathrm{d}r + \theta_\alpha(s,t)V\left(\mathbf{X}_t^{\boldsymbol{\alpha}}\right)\ \Big|\ \mathcal{F}_s^\alpha\right] \\
&= \int_0^s \theta_\alpha(0,r)f\left(\mathbf{X}_r^{\boldsymbol{\alpha}},\boldsymbol{\alpha}_r\right)\mathrm{d}r + \\
&\quad + \theta_\alpha(0,s)\mathbb{E}\left[\int_s^t \theta_\alpha(s,r)f\left(\mathbf{X}_r^{\boldsymbol{\alpha}},\boldsymbol{\alpha}_r\right)\mathrm{d}r + \theta_\alpha(s,t)V\left(\mathbf{X}_t^{\boldsymbol{\alpha}}\right)\ \Big|\ \mathbf{X}_s^{\boldsymbol{\alpha}}\right]. \quad (\mathrm{B.7})
\end{aligned}
$$

The last equality follows by conditioning on $\mathbf{X}_t^{\alpha^{\mathbf{x}}}$ and then by using the law of iterated conditional expectation. Let $\hat{\boldsymbol{\alpha}}$ be the restriction of $\boldsymbol{\alpha}^{\mathbf{x}}$ on $[s,\infty)$. Then, it is not difficult to see that $\hat{\boldsymbol{\alpha}}$ is an admissible policy in $\mathcal{A}_{s,\eta_s^\alpha}$ where $\eta_s^\alpha := \mathbb{P}^\alpha \circ \left(\mathbf{X}_s^{\alpha^{\mathbf{x}}}\right)^{-1}$, i.e. η_s^α is the law of $\mathbf{X}_s^{\alpha^{\mathbf{x}}}$. Hence, by the DPP (Lemma 5.5) we obtain that

$$
\mathbb{E}\left[\int_s^t \theta_\alpha(s,r)f\left(\mathbf{X}_r^{\alpha^{\mathbf{x}}},\boldsymbol{\alpha}_r^{\mathbf{x}}\right)\mathrm{d}r + \theta_\alpha(s,t)V\left(\mathbf{X}_t^{\alpha^{\mathbf{x}}}\right)\ \Big|\ \mathbf{X}_s^{\alpha^{\mathbf{x}}}\right] \ge V\left(\mathbf{X}_s^{\alpha^{\mathbf{x}}}\right),
$$

which plugged into (B.7) yields

$$
\mathbb{E}_{\mathbf{x}}^\alpha\left[S_t^{V,\boldsymbol{\alpha}}\Big|\mathcal{F}_s^\alpha\right] = \int_0^s \theta_\alpha(0,r)f\left(\mathbf{X}_r^{\boldsymbol{\alpha}},\boldsymbol{\alpha}_r\right)\mathrm{d}r + \theta_\alpha(0,s)V\left(\mathbf{X}_s^{\alpha^{\mathbf{x}}}\right) = S_s^{V,\alpha^{\mathbf{x}}}, \quad (\mathrm{B.8})
$$

which in turn implies that $S^{V,\alpha^{\mathbf{x}}}$ is a $\mathbb{P}^\alpha$-submartingale. On the other hand, if $\boldsymbol{\alpha}^{\mathbf{x}}$ is an optimal policy, then (B.8) ensures an equality in (B.8) which then yields $\mathbb{E}_{\mathbf{x}}^\alpha\left[S_t^{V,\boldsymbol{\alpha}}\Big|\mathcal{F}_s^\alpha\right] = S_s^{V,\alpha^{\mathbf{x}}}$. Therefore, $S^{V,\alpha^{\mathbf{x}}}$ is a $\mathbb{P}^\alpha$-martingale for any optimal policy with finite payoff. ∎

B.4. **Proof of Lemma 5.10.**

Proof. i) Take an arbitrary policy $\boldsymbol{\alpha}^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}^p$. Assume that $J^\alpha(\mathbf{x}) < +\infty$ as, otherwise, the inequality $\phi(x) \le J^\alpha(\mathbf{x})$ follows immediately. Since $S^{\phi,\alpha^{\mathbf{x}}}$ is a submartingale by assumption, $\phi(\mathbf{x}) \le \mathbb{E}_{\mathbf{x}}^\alpha\left[S_t^{\phi,\boldsymbol{\alpha}}\right]$. Therefore,

$$
\phi(\mathbf{x}) \le \mathbb{E}_{\mathbf{x}}^\alpha\left[\int_0^t e^{-\gamma_s^\alpha}f(\mathbf{X}_s^{\boldsymbol{\alpha}},\boldsymbol{\alpha}_s)\,\mathrm{d}s\right] + \mathbb{E}_{\mathbf{x}}^\alpha\left[e^{-\gamma_t^\alpha}\phi(\mathbf{X}_t^{\boldsymbol{\alpha}})\right]. \quad (\mathrm{B.9})
$$

Note that, as $t \to \infty$, the first expectation in (B.14) converges to $J^\alpha(\mathbf{x})$ (by the MCT). Hence, letting $t \to \infty$ in (B.14) and using the transversality condition (5.10) imply that $\phi(\mathbf{x}) \le J^\alpha(\mathbf{x})$. We have used the fact that $\liminf(a_n + b_n) = \liminf a_n + \liminf b_n$ whenever one of the sequences is convergent.

ii) Take $\hat{\boldsymbol{\alpha}}^{\mathbf{x}}$ be an optimal policy. By assumption (**SC**), $\phi(\mathbf{x}) = \mathbb{E}_{\mathbf{x}}^{\hat{\alpha}}\left[S_t^{\hat{\alpha},\phi}\right]$ and, thus, the same arguments above yield $\phi(\mathbf{x}) = J^{\hat{\alpha}}(\mathbf{x})$, which in turn implies that $V(\mathbf{x}) = \phi(\mathbf{x})$, as required.

iii) Using statement *i*) and the definition of V, as well as condition (**nC**), it follows that $\phi(\mathbf{x}) \le V(\mathbf{x}) \le J^{\alpha^n}(\mathbf{x}) < \phi(x) + \frac{1}{n}$, for all $n \ge 1$. Hence, letting $n \to \infty$, we get $\phi = V$, as required. ∎

B.5. **Proof of Theorem 5.13.**

Proof. i) Take an arbitrary policy $\boldsymbol{\alpha}^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}^p$. Assume that $J^\alpha(\mathbf{x}) < +\infty$ as, otherwise, the inequality $\phi(x) \le J^\alpha(\mathbf{x})$ follows immediately. Then there exists a complete, filtered probability space $(\Omega^\alpha, \mathcal{F}^\alpha, (\mathcal{F}_t^\alpha), \mathbb{P}^\alpha)$ and an $(\mathcal{F}_t^\alpha)$-adapted pair process $(\mathbf{X}^{\alpha^{\mathbf{x}}}, \boldsymbol{\alpha}^{\mathbf{x}})$ defined on it such that the process $\mathbf{X}^{\alpha^{\mathbf{x}}} := (X^1, \dots, X^n)^T$, started at $\mathbf{x} = (x_1, \dots, x_n)$, is càdlàg .

Let $\phi \in C^2(\mathbb{R}^n)$ be as in the statement and set

$$S_t^{\phi,\boldsymbol{\alpha}^{\mathbf{x}}} := \int_0^t e^{-\gamma_s^{\boldsymbol{\alpha}^{\mathbf{x}}}} f(\mathbf{X}_s^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}_s^{\mathbf{x}})\,\mathrm{d}s + e^{-\gamma_t^{\boldsymbol{\alpha}^{\mathbf{x}}}} \phi\left(\mathbf{X}_t^{\boldsymbol{\alpha}^{\mathbf{x}}}\right), \tag{B.10}$$

Since ϕ has polynomial growth of degree $q \le \max\{2, p\}$, Lemma A.4 guarantees that

$$\phi(\mathbf{X}_t^{\boldsymbol{\alpha}^{\mathbf{x}}}) = \int_0^t L^{\boldsymbol{\alpha}_s^{\mathbf{x}}} \phi(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{\mathbf{x}}})\,\mathrm{d}s + M_t^{\phi},$$

for some local martingale $M^{\phi} = (M_t^{\phi})_{t \in \mathbb{R}_+}$. Hence, the integration by parts formula [25, Corollary 2, p. 68] yields

$$e^{-\gamma_t^{\boldsymbol{\alpha}^{\mathbf{x}}}} \phi\left(\mathbf{X}_t^{\boldsymbol{\alpha}^{\mathbf{x}}}\right) = \phi(\mathbf{x}) + \int_0^t e^{-\gamma_s^{\boldsymbol{\alpha}^{\mathbf{x}}}} \left(L^{\boldsymbol{\alpha}_s^{\mathbf{x}}} \phi(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{\mathbf{x}}}) - q\left(\mathbf{X}_s^{\boldsymbol{\alpha}^{\mathbf{x}}}, \boldsymbol{\alpha}_s^{\mathbf{x}}\right) \phi(\mathbf{X}_{s-}^{\boldsymbol{\alpha}^{\mathbf{x}}}) \right)\mathrm{d}s + N_t, \tag{B.11}$$

where $N = (N_t)_{t \in \mathbb{R}_+}$ is the local martingale given by $N_t := \int_0^t e^{-\gamma_s^{\boldsymbol{\alpha}^{\mathbf{x}}}}\,\mathrm{d}M_s^{\phi}$. It is not difficult to see that, for $\mathbf{a} \in A$, the mapping $\mathbf{x} \mapsto L^{\mathbf{a}}\phi(\mathbf{x})$ is continuous (this follows from the fact that $\phi \in C^2(\mathbb{R}^n)$ has polynomial growth of degree $q \le p$ and each $\nu \in \mathcal{M}_p$ has finite second moments in the unitary ball B_1 and pth-moments outside B_1). Thus, since the paths of $\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}$ are càdlàg (so they have at most a countable number of discontinuities), the integral in (B.11) is $\mathbb{P}^{\alpha} - a.s.$ equal to the one but with s instead of $s-$.

Substituting (B.11) into (B.10) yields

$$S_t^{\phi,\boldsymbol{\alpha}^{\mathbf{x}}} := \phi(x) + \int_0^t e^{-\gamma_s^{\boldsymbol{\alpha}^{\mathbf{x}}}} \left(f(\cdot, \boldsymbol{\alpha}_s^{\mathbf{x}}) + (L^{\boldsymbol{\alpha}_s^{\mathbf{x}}} \phi)(\cdot) - q(\cdot, \boldsymbol{\alpha}_s^{\mathbf{x}}) \phi(\cdot) \right)(\mathbf{X}_s^{\boldsymbol{\alpha}^{\mathbf{x}}})\,\mathrm{d}s + N_t. \tag{B.12}$$

Thus the process $S^{\phi,\boldsymbol{\alpha}^{\mathbf{x}}}$ is a local submartingale as the integral term in (B.12) is non-negative thanks to assumption (**HJB**). Let $\{T_m\}_{n \ge 0}$ be a localising sequence for the local martingale N. Then, for each m, the stopped process $(S_{t \wedge T_m}^{\phi,\boldsymbol{\alpha}^{\mathbf{x}}})_{t \in \mathbb{R}_+}$ is a submartingale and thus $\phi(\mathbf{x}) \le \mathbb{E}_{\mathbf{x}}^{\alpha}\left[S_{t \wedge T_m}^{\phi,\boldsymbol{\alpha}}\right]$. Hence,

$$\phi(\mathbf{x}) \le \mathbb{E}_{\mathbf{x}}^{\alpha}\left[\int_0^{t \wedge T_m} e^{-\gamma_s^{\boldsymbol{\alpha}}} f(\mathbf{X}_s^{\boldsymbol{\alpha}}, \boldsymbol{\alpha}_s)\,\mathrm{d}s\right] + \mathbb{E}_{\mathbf{x}}^{\alpha}\left[e^{-\gamma_{t \wedge T_m}^{\boldsymbol{\alpha}}} \phi\left(\mathbf{X}_{t \wedge T_m}^{\boldsymbol{\alpha}}\right)\right]. \tag{B.13}$$

The uniform integrability condition (**UI**) yields $\lim_{m \to \infty} \mathbb{E}_{\mathbf{x}}^{\alpha}\left[e^{-\gamma_{t \wedge T_m}^{\boldsymbol{\alpha}}} \phi\left(\mathbf{X}_{t \wedge T_m}^{\boldsymbol{\alpha}}\right)\right] = \mathbb{E}_{\mathbf{x}}^{\alpha}\left[e^{-\gamma_t^{\boldsymbol{\alpha}}} \phi\left(\mathbf{X}_t^{\boldsymbol{\alpha}}\right)\right]$, whereas the MCT implies (by letting $m \to \infty$) that

$$\phi(\mathbf{x}) \le \mathbb{E}_{\mathbf{x}}^{\alpha}\left[\int_0^t e^{-\gamma_s^{\boldsymbol{\alpha}}} f(\mathbf{X}_s^{\boldsymbol{\alpha}}, \boldsymbol{\alpha}_s)\,\mathrm{d}s\right] + \mathbb{E}_{\mathbf{x}}^{\alpha}\left[e^{-\gamma_t^{\boldsymbol{\alpha}}} \phi\left(\mathbf{X}_t^{\boldsymbol{\alpha}}\right)\right]. \tag{B.14}$$

Hence, letting $t \to \infty$ in (B.14), the MCT and the transversality condition (**TC**) yield $\phi(\mathbf{x}) \le J^{\alpha}(\mathbf{x})$, as required.

ii) Suppose now that, for every $\mathbf{x} \in \mathbb{R}^n$, there exists an admissible pair $(\mathbf{X}^{\hat{\boldsymbol{\alpha}}^{\mathbf{x}}}, \hat{\boldsymbol{\alpha}}^{\mathbf{x}})$ such that the triplet $(\hat{\boldsymbol{\sigma}}, \hat{\nu}, \hat{\boldsymbol{\mu}})$ satisfies (5.12). To prove the optimality of $\hat{\boldsymbol{\alpha}}^{\mathbf{x}}$, it remains to prove that $\phi(\mathbf{x}) = J^{\hat{\boldsymbol{\alpha}}}(\mathbf{x})$. Similar calculations than above imply that the equality

$$S_t^{\phi,\hat{\boldsymbol{\alpha}}^{\mathbf{x}}} = \int_0^t e^{-\gamma_s^{\hat{\boldsymbol{\alpha}}^{\mathbf{x}}}} f(\mathbf{X}_s^{\hat{\boldsymbol{\alpha}}^{\mathbf{x}}}, \hat{\boldsymbol{\alpha}}_s^{\mathbf{x}})\,\mathrm{d}s + e^{-\gamma_t^{\hat{\boldsymbol{\alpha}}^{\mathbf{x}}}} \phi\left(\mathbf{X}_t^{\hat{\boldsymbol{\alpha}}^{\mathbf{x}}}\right), \quad t \in \mathbb{R}_+, \tag{B.15}$$

can be rewritten as

$$S_t^{\phi,\hat{\boldsymbol{\alpha}}^{\mathbf{x}}} = \phi(\mathbf{x}) + \int_0^t e^{-\gamma_s^{\hat{\boldsymbol{\alpha}}^{\mathbf{x}}}} \left[f(\mathbf{X}_s^{\hat{\boldsymbol{\alpha}}^{\mathbf{x}}}, \hat{\boldsymbol{\alpha}}_s^{\mathbf{x}}) + L^{\hat{\boldsymbol{\alpha}}_s^{\mathbf{x}}} \phi(\mathbf{X}_{s-}^{\hat{\boldsymbol{\alpha}}^{\mathbf{x}}}) - q\left(\mathbf{X}_s^{\hat{\boldsymbol{\alpha}}^{\mathbf{x}}}, \hat{\boldsymbol{\alpha}}_s^{\mathbf{x}}\right) \phi(\mathbf{X}_{s-}^{\hat{\boldsymbol{\alpha}}^{\mathbf{x}}}) \right]\mathrm{d}s + \hat{M}_t^{\phi},$$

where $\hat{M}^{\phi}$ is some local martingale. The càdlàg property of $\mathbf{X}_s^{\hat{\boldsymbol{\alpha}}^{\mathbf{x}}}$ and the fact that ϕ solves (5.12) ensure the equality $S_t^{\phi,\hat{\boldsymbol{\alpha}}^{\mathbf{x}}} = \phi(\mathbf{x}) + \hat{M}_t^{\phi}$, which then implies that $S_t^{\phi,\hat{\boldsymbol{\alpha}}^{\mathbf{x}}}$ is a local martingale. By repeating the same arguments as before (localising and taking the corresponding limits), we obtain the equality $\phi(\mathbf{x}) = J^{\hat{\boldsymbol{\alpha}}}(\mathbf{x})$, which implies both that $\hat{\boldsymbol{\alpha}}^{\mathbf{x}}$ is optimal and that $\phi(\mathbf{x})$ is the value function. ∎

B.6. **Proof of Lemma 5.18.**

Proof. Let $C > 0$, $\mathbf{a}_0 = (\boldsymbol{\sigma}, \nu, \boldsymbol{\mu}) \in A$, $p \geq 2$ and f be as in the statement. Take the admissible pair $(\mathbf{X}^{\tilde{\boldsymbol{\alpha}}^x}, \tilde{\boldsymbol{\alpha}}^x)$, where $\mathbf{X}^{\tilde{\boldsymbol{\alpha}}^x}$ is the Lévy process, starting at $\mathbf{x} \in \mathbb{R}^n$, corresponding to the constant policy $\tilde{\boldsymbol{\alpha}}_t^x = \mathbf{a}_0$ for all $t \in \mathbb{R}_+$ (see Remark 3.2). The assumption $|f(\mathbf{x}, \mathbf{a}_0)| \leq C(1 + |\mathbf{x}|^q)$ for all $\mathbf{x} \in \mathbb{R}^n$ and the definition of $J^{\tilde{\boldsymbol{\alpha}}}(\mathbf{x})$, yield

$$J^{\tilde{\boldsymbol{\alpha}}}(\mathbf{x}) \leq C \left(\frac{1}{\delta} + \int_0^\infty e^{-\delta t} \mathbb{E}_{\mathbf{x}}^{\alpha} |\mathbf{X}_t^{\tilde{\boldsymbol{\alpha}}^x}|^q \, \mathrm{d}t \right), \tag{B.16}$$

where $\delta > 0$ is the lower bound of the function q in the discount factor $\gamma_t^{\boldsymbol{\alpha}^{\mathbf{x}}}$. Since $\mathbf{X}^{\tilde{\boldsymbol{\alpha}}^x}$ is a Lévy process with jump intensity measure in $\mathcal{M}_p$ (recall definition in (2.1)), Proposition (4.8) ensures that, for each $t \in \mathbb{R}^n$, $\mathbb{E}[|\mathbf{X}_t^{\tilde{\boldsymbol{\alpha}}^x}|^q] \leq K(|\mathbf{x}|^q + t^q)$, for some $K > 0$ depending on the fixed constants μ, σ and the measure ν. Plugging the previous expression into (B.16) implies that $J^{\tilde{\boldsymbol{\alpha}}}(\mathbf{x}) \leq \tilde{C}(1 + |\mathbf{x}|^q)$ for some constant $\tilde{C} > 0$. Therefore, taking the infimum over all admissible policies implies, by definition of the value function, that $V(\mathbf{x}) \leq \tilde{C}(1 + |\mathbf{x}|^q)$, as required. ∎

B.7. **Proof of Lemma 5.20.** We first recall the following.

Lemma B.1. *If $b = \liminf_{t \to \infty} f(t)$, then for all $\epsilon > 0$, there exists t_0 such that $f(t) > b - \epsilon$ for all $t \geq t_0$.*

Proof. By definition, $b = \liminf_{t \to \infty} f(t) := \lim_{t \to \infty} A_t$, where $A_t := \inf\{f(s) : s \geq t\}$. Thus, by definition of limit, for all $\epsilon > 0$, there exists t_0 such that $|A_{t_0} - b| < \epsilon$, thus $b - \epsilon < A_{t_0}$. Since A_t is an increasing sequence in t, then $b - \epsilon < A_{t_0} \leq A_t$ for all $t \geq t_0$. Also, by definition of A_t, it follows that $A_t \leq f(t)$, which in turns implies $b - \epsilon < f(t)$ for all $t \geq t_0$, as required. ∎

Proof. (of Lemma 5.20) Define the mapping $g^{h, \boldsymbol{\alpha}^{\mathbf{x}}} : t \mapsto \mathbb{E}_{\mathbf{x}}^{\alpha}[e^{-\gamma_t^{\boldsymbol{\alpha}}} h(\mathbf{X}_t^{\boldsymbol{\alpha}})]$ for any nonnegative function h on $\mathbb{R}^n$. Take $\phi \in C^2(\mathbb{R}^n)$ and f as in the statement. Suppose that (5.10) does not hold. That is, $J^{\alpha}(\mathbf{x}) < \infty$ and $\liminf_{t \to \infty} g_t^{\phi, \boldsymbol{\alpha}} = \gamma$ for some constant $\gamma > 0$. Then

$$+\infty > J^{\alpha}(\mathbf{x}) \geq \mathbb{E}_{\mathbf{x}}^{\alpha} \left[\int_{t_0}^\infty e^{-\gamma_t^{\boldsymbol{\alpha}}} f(\mathbf{X}_t^{\boldsymbol{\alpha}}, \boldsymbol{\alpha}_t) \, \mathrm{d}t \right] \geq c \left(\frac{e^{-bt_0}}{b} + \int_{t_0}^\infty g_t^{|\cdot|^p, \boldsymbol{\alpha}^{\mathbf{x}}} \, \mathrm{d}t \right),$$

where $b > 0$ is the upper bound of the function q defining the discounting factor $\gamma_t^{\boldsymbol{\alpha}^{\mathbf{x}}}$. Notice the use of the lower bound of $|f(x, a)|$ as well as Tonelli's theorem to interchange the integral and the expectation in the right hand side above. Since ϕ is of polynomial growth of degree $p \geq 2$, there exists $C > 0$ such that $\int_{t_0}^\infty g_t^{\phi, \boldsymbol{\alpha}^{\mathbf{x}}} \, \mathrm{d}t \leq C(\frac{e^{-bt_0}}{b} + \int_{t_0}^\infty g_t^{|\cdot|^p, \boldsymbol{\alpha}^{\mathbf{x}}} \, \mathrm{d}t)$. Moreover, by Lemma B.1, for $\epsilon = \gamma/2$, there exists $t_0 \geq 0$ such that $g_t^{\phi, \boldsymbol{\alpha}^{\mathbf{x}}} > \gamma/2$ for all $t \geq t_0$, and this implies that $\int_{t_0}^\infty g_t^{\phi, \boldsymbol{\alpha}^{\mathbf{x}}} \, \mathrm{d}t$ is not finite, which in turn implies (by the inequalities above) that the payoff function $J^{\alpha}(\mathbf{x})$ is not finite. The latter yields a contradiction and we thus conclude that (5.10) holds. ∎

B.8. **Proof of Theorem 6.2.**

Proof. The proof follows the same lines as the one for Theorem 5.13. The only change is made at justifying the limiting step in (B.13) to obtain the inequality (B.14), which now is ensured by the DCT and the finiteness of the expectation of the running maximum of $|\mathbf{X}^{\boldsymbol{\alpha}^{\mathbf{x}}}|^q$. The latter assertion holds true due to Proposition 4.8. ∎

C.1. **Proof of Lemma 8.1.**

Proof. i) The convexity of f yields

$$\theta\psi(x) + (1-\theta)\psi(y) = \mathbb{E}\left(\int_0^\infty e^{-(q+1)t}\left[\theta f(x+B_t) + (1-\theta)f(y+B_t)\right]dt\right)$$
$$\geq \mathbb{E}\left(\int_0^\infty e^{-(q+1)t}f(\theta x + (1-\theta)y + B_t)\,dt\right)$$
$$= \psi(\theta x + (1-\theta)y),$$

establishing that ψ is convex. Symmetry follows from the symmetry of f and of the normal distribution. Finally, convexity and symmetry show that $\psi(x) = \frac{1}{2}\psi(x) + \frac{1}{2}\psi(-x) \geq \psi(0)$, establishing that zero gives the global minimum of ψ.

ii) The polynomial growth of f ensures that $\psi(x)$ is finite for each $x \in \mathbb{R}$ and, further, it implies that ψ has the same polynomial growth. It is not difficult to see that ψ satisfies

$$\frac{1}{2}\psi'' - (q+1)\psi + f = 0. \tag{C.1}$$

We will now show that $\phi := \psi + c$ solves the HJB equation:

$$\inf_{\nu \in \mathcal{M}_{\leq 1}}\left\{\frac{1}{2}h''(x) + \int\left(h(x+y) - h(x)\right)\nu(dy) - qh(x) + f(x)\right\} = 0. \tag{C.2}$$

Note that, for each $x \in \mathbb{R}$ and $\nu \in \mathcal{M}_{\leq 1}$,

$$\frac{1}{2}\phi''(x) + \int\left(\phi(x+y) - \phi(x)\right)\nu(dy) - q\phi(x) + f(x)$$
$$\geq \frac{1}{2}\psi''(x) + (\psi(0) - \psi(x))\nu(\mathbb{R}) - q\psi(x) - \psi(0) + f(x)$$
$$\geq \frac{1}{2}\psi''(x) + (\psi(0) - \psi(x)) - q\psi(x) - \psi(0) + f(x)$$
$$= \frac{1}{2}\psi''(x) - \psi(x)(1+q) + f(x) \geq 0, \tag{C.3}$$

where we used that $\psi(z) \geq \psi(0)$ for all $z \in \mathbb{R}$ (because of statement i) above), and $cq = \psi(0)$ by definition. The last inequality in (C.3) follows from (C.1). Hence, taking the infimum over all $\nu \in \mathcal{M}_{\leq 1}$ establishes that ϕ satisfies condition (**HJB**). Observe now that (8.5) and Proposition (4.8) imply that $\mathbb{E}\left[\left|X_t^{\alpha^x}\right|^p\right] \leq C\left(|x|^p + t^p\right)$ for some positive constant $C = C(\beta, p)$, which then implies (**TC**). Furthermore, $\sup_{0 \leq s \leq t}\left|X_s^{\alpha^x}\right|^p \in L^1(\mathbb{P}^\alpha)$ which establishes condition (**UI**). Moreover, since the infimum (C.2) is attained at $\hat{\nu} = \delta_{-x}$ for each $x \in \mathbb{R}$, Theorem (6.2) establishes the equality $V = \phi$, as required.

∎

C.2. **Proof of Theorem 8.3.** We will need the following result.

Theorem C.1. *Define the operator* $G^q h = \frac{1}{2}h'' - qh$. *Then,* ϕ *satisfies*

$$G^q\phi + f = 0, \quad on \quad (0, \hat{b}), \tag{C.4}$$

and

$$G^q\phi + f + \kappa - (\phi - \phi(0)) = 0, \quad on \quad (\hat{b}, \infty), \tag{C.5}$$

and ϕ *is increasing on* $\mathbb{R}_+$.

Proof. (of Theorem C.1) We first prove that ϕ satisfies (C.4)-(C.5). Let B^x be a Brownian motion started at x. Define the stopping times $\tau_b := \inf\{t \geq 0 : |B_t^x| = b\}$ and $\tau_0 := \inf\{t \geq 0 : B_t^{b,x} = 0\}$, for each $x \in (-b, b)$. Observe that $\tau_b = \tau_0$ in distribution.

Using the strong Markov property of the Brownian motion, ϕ in (8.10) can be rewritten as

$$\phi(x) = \mathbb{E}_x\left[\int_0^{\tau_0} e^{-qt}\left(f(B_t^{b,x}) + \kappa 1_{|B_t^x| \geq b}\right) dt\right] + \phi(0)\mathbb{E}_x\left[e^{-q\tau_0}\right],$$

$$= \mathbb{E}_x\left[\int_0^{\infty} e^{-\left(qt + \int_0^t 1_{\{|B_s^{b,x}| \geq b\}} ds\right)}\left(f(B_t^{b,x}) + \kappa 1_{|B_t^x| \geq b}\right) dt\right] + \phi(0)\mathbb{E}_x\left[e^{-q\tau_0}\right]$$

Using that

$$\mathbb{E}_x\left[e^{-q\tau_0}\right] = \mathbb{E}_x\left[\int_0^{\infty} 1_{\{|B_s^{b,x}| \geq b\}} e^{-\left(qt + \int_0^t 1_{\{|B_s^{b,x}| \geq b\}} ds\right)} dt\right],$$

it follows that

$$\phi(x) = \mathbb{E}_x\left[\int_0^{\infty} e^{-\left(qt + \int_0^t 1_{\{|B_s^{b,x}| \geq b\}} ds\right)}\left(f(B_t^{b,x}) + \left(\kappa + \phi(0)\right)1_{|B_t^x| \geq b}\right) dt\right], \quad x \in \mathbb{R}$$

and

$$\phi(0) = \frac{\mathbb{E}_x\left[\int_0^{\infty} e^{-\left(qt + \int_0^t 1_{\{|B_s^{b,x}| \geq b\}}\right)}\left(f(B_t^{b,x}) + \kappa 1_{|B_t^x| \geq b}\right) dt\right]}{1 - \mathbb{E}_x\left[\int_0^{\infty} e^{-\left(qt + \int_0^t 1_{\{|B_s^{b,x}| \geq b\}}\right)} 1_{\{|B_s^{b,x}| \geq b\}}\right]}.$$

Therefore, the stationary Feynman-Kac formula implies that ϕ solves

$$\frac{1}{2}\phi''(x) - (q + 1_{|x| \geq b})\phi(x) + f(x) + (\kappa + \phi(0))1_{|x| \geq b} = 0,$$

that is,

$$\begin{cases} \frac{1}{2}\phi''(x) - q\phi(x) + f(x) = 0 & x \in (-b, b) \\ \frac{1}{2}\phi''(x) - (q + 1)\phi(x) + f(x) + \kappa + \phi(0) = 0 & x \in (-b, b)^c \end{cases}$$

The existence of $\hat{b}$ can be justified as follows. Observe that $\mathbb{E}_x f(B_t) \uparrow \infty$ as $x \to \infty$, $\phi_b(x) \to \infty$ for any b, and $\phi_b(b) \geq c_q f(b-1)$ for some constant c_q. Moreover,

$$\phi_b(0) = \mathbb{E}_0 \int_0^{\tau_b} e^{-qt} f(B_t)\, dt + \mathbb{E}_0 e^{-q\tau_b} \phi_b(b) \tag{C.6}$$

$$\leq \mathbb{E}_0 \int_0^{\infty} f(B_t)e^{-qt}\, dt + c_b\phi_b(b), \tag{C.7}$$

where $0 < c_b < 1$. Thus, $\phi_b(b) - \phi_b(0) \geq (1 - c_b)\phi_b(b) - d \to \infty$ as $b \to \infty$ and, further, $\phi_b(b) - \phi_b(0)$ is continuous as a function of b.

Now, to prove that ϕ is increasing, we proceed in six stages:

(1) Show that $\phi \in C^2(\mathbb{R})$.

(2) Show that $\phi(x) - \epsilon x \to \infty$ as $x \to \infty$ for some $\epsilon > 0$.

(3) Show that ϕ' has no negative minimum on the domain $(0, \hat{b})$ or on the domain $(\hat{b}, \infty)$.

(4) Show that $\liminf_{x \to \infty} \phi'(x) \geq 0$.

(5) Deduce that either $\phi' \geq 0$ on $\mathbb{R}_+$ or

    a) ϕ' attains its unique negative minimum on $[0, \hat{b}]$ at $\hat{b}$ and

    b) ϕ' attains its unique negative minimum on $[\hat{b}, \infty]$ at $\hat{b}$.

(6) Deduce a contradiction from (C.4) and (C.5).

Proof of:

(1) Since ϕ is clearly positive and satisfies (C.4) on $(0, \hat{b})$, it follows that ϕ is C^2 on $(0, \hat{b})$. Similarly, since ϕ satisfies (C.5) on $(\hat{b}, \infty)$, it follows that ϕ is C^2 on $(\hat{b}, \infty)$. A standard martingale argument based on the Itô -Tanaka formula shows that ϕ is C^1 on $\mathbb{R}$. It then follows from the characterization of $\hat{b}$ that ϕ'' does not have a discontinuity at $\hat{b}$.

(2) A simple argument show that, for $x \geq 1$,

$$\phi(x) \geq e^{-(q+1)}\mathbb{E}_x\left[\int_0^1 f(B_t)\,\mathrm{d}t\right] \geq e^{-(q+1)}\mathbb{P}_x\left(\inf_{0 \leq t \leq 1} B_t \geq x - 1\right) \geq c_q f(x-1), \qquad \text{(C.8)}$$

for some $c_q > 0$. Since f is convex, increasing on $\mathbb{R}_+$, it is of at least linear growth on $\mathbb{R}_+$, and so the result follows.

(3) Denote ϕ' by ψ. It follows from differentiating (C.4) and (C.5) that

$$\frac{1}{2}\psi'' - q\psi + f' = 0, \quad \text{on} \quad (0, \hat{b}), \qquad \text{(C.9)}$$

and

$$\frac{1}{2}\psi'' - (q+1)\psi + f' = 0, \quad \text{on} \quad (\hat{b}, \infty). \qquad \text{(C.10)}$$

Since $f' > 0$ on $\mathbb{R}_+$ the result follows from the strong minimum principle applied separately on each domain.

(4) On $(\hat{b}, \infty)$, $\frac{1}{2}\psi'' - (q+1)\psi = -f' < 0$. It follows from the strong minimum principle that ψ has no negative minimum on $(\hat{b}, \infty)$. Consequently, if $m := \liminf \psi < 0$, then once ψ becomes negative it must decrease monotonically to m. But then $\lim \psi = -\infty$ which contradicts the positivity of ϕ.

(5) Note that $\psi(0) = 0$ by symmetry of ϕ. So if ψ has a negative minimum on $[0, \hat{b}]$ it follows from (3) that it must be attained at $\hat{b}$. Similarly for the negative minimum on $[\hat{b}, \infty)$.

(6) Suppose that ψ goes below 0. Then from (5) we must have $0 < l < \hat{b} < r < \infty$ such that

$$\{x : \psi(x) < 0\} = (l, r)$$

and ψ is decreasing on $[l, \hat{b}]$ and increasing on $[\hat{b}, r]$. It follows that $\phi'' = \psi'$ is negative on $(l, \hat{b})$ and zero at $\hat{b}$. Now define $h = q\phi - f$ and notice that $h(0) \geq 0$ since f is minimised at 0, while $h(\hat{b}) = \frac{1}{2}\phi''(\hat{b}) = 0$. Note that $G^q h = -\frac{1}{2}f'' \leq 0$ (since f is convex) on $(0, \hat{b})$ so by the weak minimum principle the (negative) minimum of h on $[0, \hat{b}]$ is attained at the boundary. But the boundary values are non-negative (since $h = \frac{1}{2}\phi'''$ on $[0, \hat{b}]$) so we deduce a contradiction.

∎

Proof. (of Theorem 8.3) Observe that ϕ satisfies the HJB equation because ϕ solves (C.4)-(C.5) and ϕ is increasing on $\mathbb{R}_+$ (by Theorem C.1). Condition (**TC**) is satisfied by Corollary 5.21. The validity of (**UI**) follows by Proposition 4.8 as each $\nu \in \mathcal{M}_{\leq 1}$. Therefore, the existence of the process $B^{b,x}$ satisfying (5.12) and Theorem 5.13 imply the result. ∎

C.3. **Proof of Theorem 8.5.** For this proof, we need the following preliminary result.

Lemma C.2. *The process* $\hat{\alpha}^{\mathbf{x}} = (\hat{\alpha}_t^{\mathbf{x}})_{t \in \mathbb{R}_+}$ *as defined in* (8.13) *is an admissible policy in* $\hat{\mathcal{A}}_{\mathbf{x}}^2$. *Furthermore,* $J^{\hat{\alpha}}(\mathbf{x}) < \infty$, *for each* $\mathbf{x} \in \mathbb{R}^n$.

Proof. First notice that (**H2**) holds directly as $\hat{\alpha}_s^{\mathbf{x}} = (\hat{\sigma}, \hat{\nu}, \hat{\mu}(\hat{\mathbf{X}}_s^{\mathbf{x}}))$ for all s, and, further, $\hat{\nu} \in \mathcal{M}$. To prove (**H1**), it is enough to guarantee the existence of a filtered probability space $(\hat{\Omega}, \hat{\mathcal{F}}, (\hat{\mathcal{F}})_{t \geq 0}, \hat{\mathbb{P}})$ supporting the process $\hat{\mathbf{X}}^{\mathbf{x}}$, $\mathbf{x} \in \mathbb{R}^n$. This is, however, just a direct consequence of [29, Theorem 3.1]. Furthermore, Proposition 1.7 in [6, Chapter 4] guarantees that, for every

function h in the domain of the (infinitesimal) generator $\hat{L}$ (and hence for every $h \in C_c^2(\mathbb{R}^n)$), the process M^h defined by $M_t^h := h(\hat{\mathbf{X}}_t^{\mathbf{x}}) - \int_0^t (\hat{L}h)(\hat{\mathbf{X}}_s^{\mathbf{x}}) \, \mathrm{d}s$ is an $(\hat{\mathcal{F}})_t$-martingale. Thus, the equality

$$(\hat{L}h)(\hat{\mathbf{X}}_s^{\mathbf{x}}) = (L^{\hat{\alpha}_s^{\mathbf{x}}} h)(\hat{\mathbf{X}}_s^{\mathbf{x}}), \quad \hat{\mathbb{P}} - a.s., \tag{C.11}$$

implies (3.2). Therefore, $\hat{\boldsymbol{\alpha}}^{\mathbf{x}} \in \mathcal{A}_{\mathbf{x}}$ and its associated controlled process $\mathbf{X}^{\hat{\boldsymbol{\alpha}}^{\mathbf{x}}}$ is given by $\hat{\mathbf{X}}^{\mathbf{x}}$.

For the second part, let us recall that $|\mathbf{G}\mathbf{x}| \leq \|\mathbf{G}\| \, |\mathbf{x}|$ for any matrix $\mathbf{G} \in \mathrm{M}_{n \times n}(\mathbb{R})$, $\mathbf{x} \in \mathbb{R}^n$, whereas $|\mathbf{x}^T \mathbf{G} \mathbf{x}| \leq \lambda_n |\mathbf{x}|$ for any positive definite matrix $\mathbf{G}$ with eigenvalues $0 \leq \lambda_1 \leq \ldots \leq \lambda_n$. Hence, due to the quadratic form of the payoff function and the fact that $\hat{\mu}$ is linear in $\mathbf{x}$, to prove (5.10) it is enough to show the inequality

$$\mathbb{E}_{\mathbf{x}}^{\alpha} \left[|\hat{\mathbf{X}}_t^{\mathbf{x}}|^2 \right] \leq C(|\mathbf{x}|^2 + 1), \quad t \in \mathbb{R}_+, \ \mathbf{x} \in \mathbb{R}^n, \tag{C.12}$$

for some positive constant C independent of both t and $\mathbf{x}$.

By Theorem 2.12 in [5], $\hat{\mathbf{X}}^{\mathbf{x}}$ is a $\mathbb{R}^n$-valued semimartingale satisfying the stochastic differential equation

$$\mathrm{d}\hat{\mathbf{X}}_t^{\mathbf{x}} = -\mathbf{Q}\hat{\mathbf{X}}_t^{\mathbf{x}} \, \mathrm{d}t + (\mathbf{u} + \mathbf{v}) \, \mathrm{d}t + \hat{\boldsymbol{\sigma}} \, \mathrm{d}\mathbf{W}_t + \mathrm{d}\mathbf{Z}_t, \qquad \hat{\mathbf{X}}_0^{\mathbf{x}} = \mathbf{x}, \tag{C.13}$$

where $\mathbf{W} = (\mathbf{W}_t)_{t \in \mathbb{R}_+}$ is a standard n-dimensional Brownian motion and $\mathbf{Z} = (\mathbf{Z}_t)_{t \in \mathbb{R}_+}$ is a $\mathbb{R}^n$–valued pure-jump Lévy martingale with quadratic variation $[\mathbf{Z}]$ taking values in $\mathrm{M}_{n \times n}(\mathbb{R})$, where the quadratic covariation entries $[Z_i, Z_j]$, $1 \leq i, j \leq d$, are given by

$$[Z_i, Z_j]_t = t \int_{\mathbb{R}_0^n} y_i y_j \hat{\nu}(\mathrm{d}y), \quad t \in \mathbb{R}_+.$$

Using Itô 's formula [25, Theorem 33, Chapter 7, p. 81] one can verify that the solution $\hat{\mathbf{X}}^{\mathbf{x}}$ to (C.13) is given by

$$\hat{\mathbf{X}}_t^{\mathbf{x}} = e^{-\mathbf{Q}t} \mathbf{x} + \int_0^t e^{-\mathbf{Q}(t-s)} (\mathbf{u} + \mathbf{v}) \, \mathrm{d}s + \mathbf{U}_t + \mathbf{V}_t, \tag{C.14}$$

where

$$\mathbf{U}_t := \int_0^t \hat{\boldsymbol{\sigma}} e^{-\mathbf{Q}(t-s)} \, \mathrm{d}\mathbf{W}(s) \quad \text{and} \quad \mathbf{V}_t := \int_0^t e^{-\mathbf{Q}(t-s)} \, \mathrm{d}\mathbf{Z}(s).$$

Since both $\mathbf{W}$ and $\mathbf{Z}$ are martingales and, further,

$$\mathbb{E}_{\mathbf{x}}^{\alpha} \{ [\mathbf{U}]_t \} \leq \frac{n \|\hat{\boldsymbol{\sigma}}\|^2}{\|\mathbf{Q}\|} < \infty \quad \text{and} \quad \mathbb{E}_{\mathbf{x}}^{\alpha} \{ [\mathbf{V}]_t \} \leq \frac{n}{\|\mathbf{Q}\|} \int_{\mathbb{R}_0^n} |\mathbf{y}|^2 \hat{\nu}(\mathrm{d}\mathbf{y}) < \infty,$$

it follows that $\mathbf{U}$ and $\mathbf{V}$ are also true martingales. Moreover, by Theorem 29 in [25, p. 75] and the generalised Ito isometry, it follows that $\mathbb{E}_{\mathbf{x}}^{\alpha} \{ \mathbf{U}_t^2 \} = \mathbb{E}_{\mathbf{x}}^{\alpha} \{ [\mathbf{U}]_t \}$ and $\mathbb{E}_{\mathbf{x}}^{\alpha} \{ \mathbf{V}_t^2 \} = \mathbb{E}_{\mathbf{x}}^{\alpha} \{ [\mathbf{V}]_t \}$. Set $C := \max \{ 4, |\mathbf{u} + \mathbf{v}|^2 / \|\mathbf{Q}\| + n \delta^* / \|\mathbf{Q}\| \}$ and $\delta^* := \|\hat{\boldsymbol{\sigma}}\|^2 + \int_{\mathbb{R}_0^n} |\mathbf{y}|^2 \hat{\nu}(\mathrm{d}\mathbf{y})$. Then, (C.14) and the above estimates imply (C.12), which in turn implies that (**H3**) holds and, hence, $\hat{\boldsymbol{\alpha}}^{\mathbf{x}} \in \hat{\mathcal{A}}_{\mathbf{x}}^2$. Finally, note that estimate (C.12) also ensures $J(\hat{\mathbf{X}}^{\mathbf{x}}, \hat{\boldsymbol{\alpha}}^{\mathbf{x}}) < \infty$, as required. ∎

Proof. (of Theorem 8.5)
By Theorem 6.2, we need to prove that the admissible pair $(\hat{\mathbf{X}}^{\mathbf{x}}, \hat{\boldsymbol{\alpha}}^{\mathbf{x}})$ satisfies (5.12) with $\phi(\mathbf{x}) := \mathbf{x}^T \mathbf{B} \mathbf{x} + \mathbf{x}^T \mathbf{c} + d$, i.e. ϕ solves the HJB equation

$$\inf_{\mathbf{a} = (\boldsymbol{\sigma}, \nu, \boldsymbol{\mu}) \in \mathbf{A}} \left\{ L^{\mathbf{a}} \phi(\mathbf{x}) - q\phi(\mathbf{x}) + (\mathbf{x}^T \Lambda \mathbf{x} + \boldsymbol{\mu}^T \Theta \boldsymbol{\mu}) \right\} = 0. \tag{C.15}$$

We will see then that, for each $\mathbf{x} \in \mathbb{R}^n$, the triplet $(\hat{\boldsymbol{\sigma}}, \hat{\nu}, \hat{\mu}(\mathbf{x}))$ is a minimiser of (C.15).

Since $\nabla \phi(\mathbf{x}) = 2\mathbf{B}\mathbf{x} + \mathbf{c}$ and $\mathrm{H}\phi(\mathbf{x}) = 2\mathbf{B}$, it follows that

$$L^{\mathbf{a}} \phi(\mathbf{x}) = (\mathbf{u} + \boldsymbol{\mu})^T (2\mathbf{B}\mathbf{x} + \mathbf{c}) + \mathrm{Tr}(\boldsymbol{\sigma}^T \mathbf{B} \boldsymbol{\sigma}) + \int_{\mathbb{R}_0^n} \mathbf{y}^T \mathbf{B} \mathbf{y} \, \nu(\mathrm{d}\mathbf{y}). \tag{C.16}$$

Thus, the minimal infinitesimal variance $\hat{\delta}$ in (8.12) yields

$$\mathbf{u}^T (2\mathbf{B}\mathbf{x} + \mathbf{c}) + \hat{\delta} - q(\mathbf{x}^T \mathbf{B}\mathbf{x} + \mathbf{x}^T \mathbf{c} + d) + \mathbf{x}^T \Lambda \mathbf{x} + \inf_{\boldsymbol{\mu} \in \mathbb{R}^n} \left\{ \boldsymbol{\mu}^T (2\mathbf{B}\mathbf{x} + \mathbf{c}) + \boldsymbol{\mu}^T \Theta \boldsymbol{\mu} \right\} = 0. \tag{C.17}$$

Let $g(\boldsymbol{\mu}) := \boldsymbol{\mu}^T(2\mathbf{B}\mathbf{x} + \mathbf{c}) + \boldsymbol{\mu}^T\Theta\boldsymbol{\mu}$. Then

$$\mu^*(\mathbf{x}) := -\frac{1}{2}\Theta^{-1}(2\mathbf{B}\mathbf{x} + \mathbf{c}) \tag{C.18}$$

minimises g for every $\mathbf{x} \in \mathbb{R}^n$ and, further, $g(\mu^*(\mathbf{x})) = -(2\mathbf{B}\mathbf{x} + \mathbf{c})^T\Theta^{-1}(2\mathbf{B}\mathbf{x} + \mathbf{c})/4$, which (after rearranging terms and substituting into equation (C.17)) yields

$$\mathbf{x}^T\left(\Lambda - q\mathbf{B} - \mathbf{B}^T\Theta^{-1}\mathbf{B}\right)\mathbf{x} + \left(2\mathbf{u}^T\mathbf{B} - q\mathbf{c}^T - \mathbf{c}^T\Theta^{-1}\mathbf{B}\right)\mathbf{x} + \left(\mathbf{u}^T\mathbf{c} + \hat{\delta} - qd - \frac{\mathbf{c}^T\Theta^{-1}\mathbf{c}}{4}\right) = 0. \tag{C.19}$$

Observe that Θ^{-1} exists as Θ is a positive definite matrix. Since equation (C.19) should hold for every $\mathbf{x} \in \mathbb{R}^n$, we can now verify that $\mathbf{B}$, $\mathbf{c}$ and d as defined in (8.15) solve the corresponding system of equations. Moreover, using the definitions of $\mathbf{c}$ and $\mathbf{v}$, we can see that $\mu^*(\mathbf{x})$ coincides with $\hat{\mu}(\mathbf{x}) = -\mathbf{Q}\mathbf{x} + \mathbf{v}$, as required.

Let us also observe that the previous calculations imply that the function ϕ satisfies

$$\hat{L}\phi(\mathbf{x}) - q\phi(\mathbf{x}) + \mathbf{x}^T\Lambda\mathbf{x} + \boldsymbol{\mu}^T\Theta\boldsymbol{\mu} = 0, \quad \text{for each } \mathbf{x} \in \mathbb{R}_+, \tag{C.20}$$

where $\hat{L}$ is the operator defined in (8.14). Since, by Lemma 5.20, the transversality condition (5.10) holds, Theorem 6.2 implies the optimality of the family of admissible policies $\{\hat{\boldsymbol{\alpha}}^{\mathbf{x}} : \mathbf{x} \in \mathbb{R}^n\}$, as desired. ∎

REFERENCES

[1] Y. Aït-Sahalia and J. Jacod. *High-Frequency Financial Econometrics.*, volume 816 of *Lect. Notes in Math.* Princeton: Princeton University Press., 2014.

[2] D. Applebaum. *Lévy Processes and Stochastic Calculus.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2 edition, 2009.

[3] V. S. Borkar. Controlled diffusion processes. *Probability Surveys*, 2:213–244, 2005.

[4] M. H. A. Davis. Martingale methods in stochastic control. *Stochastic Control Theory and Stochastic Differential Systems*, pages 85–117, 1979.

[5] D. Duffie, D. Filipovi, and W. Schachermayer. Affine processes and applications in finance. *Ann. Appl. Probab.*, 13(3):984–1053, 08 2003.

[6] S. N. Ethier and T. G. Kurtz. *Markov processes.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.

[7] W. H. Fleming and Soner H. M. *Controlled Markov Processes and Viscosity Solutions.* Springer-Verlag, Berlin, 1993.

[8] W. H. Fleming and Rishel R. W. *Deterministic and Stochastic Optimal Control*, volume 1 of *Applications of Mathematics.* Springer-Verlag, Berlin Heidelberg New York, 1975.

[9] S. D. Jacka and A.á Mijatovic. On the policy improvement algorithm in continuous time. *Stochastics*, 89(1):348–359, 2017.

[10] J. Jacod and A. N. Shiryaev. *Limits theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [A Series of Comprehensive Studies in Mathematics].* Springer-Verlag, Berlin, 1987.

[11] Kalman. When is a linear control system optimal? *J. Basic Eng.*, 86:51–60, 1967.

[12] I. Karatzas and I.M. Zamfirescu. Martingale approach to stochastic control with discretionary stopping. *Appl. Math. Optim.*, 53:163–184, 2006.

[13] N. Karoui. *Les aspects probabilistes du contrôl stochastique*, volume 816 of *Lect. Notes in Math.* Springer-Verlag, 1981.

[14] Nguyen D. Karoui N. and M. Jeanblanc-Picqué. Compactification methods in the control of degenerate diffusions: existence of an optimal control. *Stochastics*, 20(3):169–219, 1987.

[15] T. G. Kurtz. Martingale problems for controlled processes. *In: Germani A. (eds) Stochastic Modelling and Filtering. Lecture Notes in Control and Information Sciences*, 91:75–90, 1987.

[16] K. Martensson. On the matrix riccati equation. *Information Sciences*, 3:17–49, 1971.

[17] K. Martensson. New approaches of the numerical solution of optimal control problems. *Department of Automatic Control, Lund Institute of Technology (LTH)*, 1972.

[18] B. L. Miller. Finite state continuous time markov decision processes with a finite planning horizon. *SIAM J. Control*, 6:266–280, 1968.

[19] B. Oksendal. *Stochastic Differential Equations.* Springer, Berlin Heidelberg New York, 6th edition, 2003.

[20] B. Oksendal and A. Sulem. *Applied Stochastic Control of Jump Diffusions.* Springer-Verlag, Berlin Heidelberg, second edition, 2007.

[21] H. Pham. On some recent aspects of stochastic control and their applications. *Probability Surveys*, 2:506–549, 2005.

[22] H. Pham. *Continuous-time Stochastic Control and Optimization with Financial Applications*. Stochastic Modelling and Applied Probability. Springer, 2009.

[23] S. R. Pliska. Controlled jump processes. *Stochastic Processes and their Applications*, 3:259–282, 1975.

[24] J. E. Protter. Matrix quadratic solutions. *J. SIAM Appl. Math.*, 14(3):496–501, 1966.

[25] P. E. Protter. *Stochastic integration and differential equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2005. Second edition. Version 2.1, Corrected third printing.

[26] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[27] L.C.G. Rogers and D. Williams. *Diffusions, Markov processes and martingales*, volume 2. Cambridge University Press, 2 edition, 2000.

[28] K. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1 edition, 1999.

[29] K. Sato and M. Yamazato. Operator-selfdecomposable distributions as limit distributions of processes of ornstein-uhlenbeck type. *Stoch. Proc. Appl.*, 17:73–100, 1984.

[30] L. R. Schilling and A. Schnurr. The symbol associated with the solution of a stochastic differential equation. *Electronic Journal of Probability*, 15:1369–1393, 2010.

[31] L. D. Stone. Necessary and sufficient conditions for optimal control of semi-markov jump processes. *SIAM J. Control*, 11:187–201, 1973.

[32] Wonham. On matrix quadratic equations and matrix riccati equations. *Technical Report 67-5, Division of Applied Mathematics, Brown University, Providence, R. I.*

[33] J. Yong and X. Y. Zhou. *Stochastic Controls: Hamiltonian Systems and HJB equations*, volume 43 of *Applications of Mathematics. Stochastic Modelling and Applied Probability*. Springer-Verlag New York, Inc., 1999.

[1,2,3]DEPARTMENT OF STATISTICS, UNIVERSITY OF WARWICK, COVENTRY CV4 7AL UK