A review from the PDE viewpoint of Hamilton-Jacobi-Bellman equations arising in optimal control with vectorial cost

Article

Accepted Version

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

A REVIEW FROM THE PDE VIEWPOINT OF HAMILTON-JACOBI-BELLMAN EQUATIONS ARISING IN OPTIMAL CONTROL WITH VECTORIAL COST

NIKOS KATZOURAKIS AND TRISTAN PRYER

ABSTRACT. This paper is a review of results on Optimisation which are perhaps not so standard in the PDE realm. To this end, we consider the problem of deriving the PDEs associated to the optimal control of a system of either ODEs or SDEs with respect to a *vector-valued* cost functional. Optimisation is considered with respect to a partial ordering generated by a given cone. Since in the vector case minima may not exist, we define vectorial value functions as (Pareto) minimals of the ordering. Our main objective is the derivation of the model PDEs which turn out to be parametric families of HJB single equations instead of systems of PDEs. However, this allows to utilise the theory of Viscosity Solutions.

1. INTRODUCTION

Let $A \subseteq \mathbb{R}^m$ be a compact set, $F : \mathbb{R}^n \times A \longrightarrow \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \times A \longrightarrow \mathbb{R}^{n \times m}$ continuous maps. For $x \in \mathbb{R}^n$, consider the following initial value problem for a system of stochastic differential equations (SDEs)

$$(1.1) \quad \begin{cases} \mathrm{d}\chi(s) = F\big(\chi(s), \alpha(s)\big) \, \mathrm{d}s + \sigma\big(\chi(s), \alpha(s)\big) \, \mathrm{d}W(s), & t < s < T, \\ \chi(t) = x, \end{cases}$$

in the Itô sense. Here α is a measurable map is the class $\mathcal{A}$, where

$$(1.2) \quad \mathcal{A} := \Big\{ \alpha \in L^\infty(0, T) \mid \alpha(t) \in A, \text{ for a.e. } t \in (0, T) \Big\}$$

and $W(s)$ is an m dimensional system of independent Wiener processes. Note that in the case $\sigma = 0$ the problem reduces to a system of ordinary differential equations (ODEs)

$$(1.3) \quad \begin{cases} \dot{\chi}(s) = F\big(\chi(s), \alpha(s)\big), & t < s < T, \\ \chi(t) = x. \end{cases}$$

In this paper we consider the problem of deriving the PDE associated to the optimal control of system (1.1) with respect to the *vectorial cost functional*

1

$$(1.4) \qquad C_{x,t}[\alpha] \; := \; \mathbb{E}\left[g\big(\chi(T)\big) \; + \; \int_t^T h\big(\chi(s), \alpha(s)\big) \, ds\right]$$

where

$$(1.5) \qquad h \; : \; \mathbb{R}^n \times A \longrightarrow \mathbb{R}^N, \quad g \; : \; \mathbb{R}^n \longrightarrow \mathbb{R}^N,$$

are given maps, called the *running cost* and the *terminal cost* respectively and $\mathbb{E}$ is the expectation defined with respect to the measure induced by the stochastic basis. In (1.4) χ denotes the stochastic (deterministic) flow map of (1.1) when $\sigma \not\equiv 0$ ($\sigma \equiv 0$) respectively, having suppressed the dependence in x, α:

$$(1.6) \qquad \chi(s) \; \equiv \; \chi\big(s, x, \alpha(s)\big).$$

We would like to clarify that *this is primarily a review paper which is aimed at PDE theorists who may not be experts of control theory. In particular, Section 2 is a review of results standard in the community of Optimisation which are lesser known in the PDE community. However, Sections 3 and 4 contain seemingly new results as we explain below. Our main objective is to derive the Hamilton-Jacobi and Hamilton-Jacobi-Bellman equations which are associated to the problem of vectorial optimisation with or without noise. It is a remarkable fact that instead of obtaining systems of HJB equations as one would expect in the vectorial case, we actually obtain parametric families of single equations via the method of scalarisation. Since the equations turn out to be single and not systems (but with parameters), we invoke the Crandall-Ishii-Lions theory of Viscosity Solutions. Further, we do not discuss the much more delicate question of uniqueness.*

In the sequel we will assume that F, g, h, σ are bounded and Lipschitz continuous with respect to $x \in \mathbb{R}^n$, uniformly with respect to $a \in A$:

$$(1.7) \qquad \left.\begin{aligned} |F(x,a)|, \; |g(x)|, \; |h(x,a)|, \; |\sigma(x,a)| \; &\leq \; C, \\ |F(x,a) - F(y,a)| \; &\leq \; C|x-y|, \\ |g(x) - g(y)| \; &\leq \; C|x-y|, \\ |h(x,a) - h(y,a)| \; &\leq \; C|x-y|, \\ |\sigma(x,a) - \sigma(y,a)| \; &\leq \; C|x-y|, \end{aligned}\right\} \text{ for all } x, y \in \mathbb{R}^n, \; a \in A.$$

For simplicity in the exposition we have made the *simplifying assumption* that σ is not matrix valued. Moreover, the bounds we assume are also non-optimal in order to allow us to focus on the main ideas rather than on technical complications.

In both the deterministic and stochastic cases of *scalar cost functional*, namely when $N = 1$, this problem is standard in Control Theory and there is an extensive literature, see for instance Evans [E], Fleming-Soner [FS], Baldi-Capuzzo Dolcetta [BCD], Lions [L], Fleming-Rishel [FR] and Kenneth [Ke]. When $C_{x,t}[\alpha] \in [0, \infty]$, optimisation with respect to this scalar cost functional is unambiguous: one seeks to minimise (1.4) over all admissible controls $\alpha \in \mathcal{A}$. Along the lines of (ordinary) dynamic programming, one may define the *value function*

$$(1.8) \qquad u(x,t) \; := \; \inf_{\alpha \in \mathcal{A}} C_{x,t}[\alpha], \quad x \in \mathbb{R}^n, \, 0 < t < T.$$

It then follows by standard PDE theory (see e.g. [E] for the case of vanishing white noise) that under assumption (1.7) the value function is Lipschitz continuous and

solves an initial value problem for a Hamilton–Jacobi–Bellman (HJB) PDE

(1.9)
$$\begin{cases} u_t \;+\; H(\cdot, Du, D^2 u) \;=\; 0, & \text{in } \mathbb{R}^n \times (0, T), \\ \qquad\qquad\qquad u \;=\; g, & \text{on } \mathbb{R}^n \times \{0\}, \end{cases}$$

in the Viscosity sense (see e.g. [CIL] for the theory of viscosity solutions, and for a more elementary introduction we refer to [K]). Here the Hamiltonian H is the function defined by

(1.10) $$H(x, p, Z) \;:=\; \min_{a \in A} \left\{ \frac{1}{2} \sigma(x, a) \sigma(x, a)^\top : Z \;+\; F(x, a) \cdot p \;+\; h(x, a) \right\}.$$

The HJB equation can be utilised to construct a *feedback control* α^* which optimally drives the dynamics of the flow generated by the system (1.1) and minimises both the running cost and the terminal cost. Roughly, at points of differentiability of u, α^* is defined by selecting for each $t < s < T$ the $\alpha(s)$ which realises the minimum:

$$H\Big(\chi^*(s), Du\big(\chi^*(s), \alpha^*(s)\big), D^2 u\big(\chi^*(s), \alpha^*(s)\big)\Big)$$
$$= \; \frac{1}{2} \sigma\big(\chi^*(s), \alpha^*(s)\big) \sigma\big(\chi^*(s), \alpha^*(s)\big)^\top : D^2 u\big(s, \alpha^*(s)\big)$$
$$+ \; F\big(\chi^*(s), \alpha^*(s)\big) \cdot Du\big(\chi^*(s), \alpha^*(s)\big) \;+\; h\big(\chi^*(s), \alpha^*(s)\big),$$

where

$$\chi^*(s) \;:=\; \chi\big(s, x, \alpha^*(s)\big).$$

Conversely in the deterministic case, given any HJ equation $u_t + H(\cdot, Du) = 0$ with $H(x, p)$ concave in p, one can always relate it to a scalar optimisation problem for an ODE system of the form (1.3) for some deterministic cost functional, by using the fact that concave functions can be written as infima of a family of affine functions.

In this paper we consider instead the case of $N \geq 2$ and we seek to optimise the vectorial cost functional (1.4). Vectorial Optimal Control and vectorial Dynamic Programming are extremely important in applications and have been extensively studied in the last 50 years, mostly in connection to real-world applications like in Economics/Finance, Mechanics/Engineering, Aeronautics, Automotive industry etc, see e.g. Guigue [G], Bellman-Fan [BF], Chen-Huang-Yang [CHY], Debreu [D], Henig [H], Isii [I], Luenberger [L1, L2, L3, L4], Olech [O], Salukvadze [S], Pareto [P], Boyd-Vanderberghe [BV] and references therein.

In the vectorial case, ones needs to be very careful regarding the meaning of "minimise the vector cost". In (finite-dimensional) Optimisation theory (see e.g. [BV]), it is fairly standard to consider minimisation with respect to a *partial ordering "$\leq_K$" generated by a convex cone $K \subseteq \mathbb{R}^N$ with some further properties*, that is for $\xi, \eta \in \mathbb{R}^N$, we define

$$\xi \leq_K \eta \quad \Longleftrightarrow \quad \eta - \xi \in K.$$

Vector-valued optimisation is extremely important in applications, and there a very active current research on the topic, mostly in connection to multi-criterion optimisation and Nash equilibria, see for instance [DD, M, MGGJ, GC, RBG, BKR, RK]. Let us also note that there exists a large number of contributions in Game Theory which are closely related to the problem considered here. For instance, two well known references in the area are by Basar and Olsder [BO] and by Abou-Kandil, Freiling, Ionescu and Jank [AFIJ], that also contain numerous relevant references.

In some of these works, the problem addressed herein has been widely investigated but only in particular *Linear-Quadratic* case. In this case the theory has been well established, with or without stochastic terms. The equations then simplify to coupled *Riccati equations* that are exactly HJB in the linear quadratic context. We will summarise this issue further in the sequel.

In the scalar case we typically have $K = [0, \infty)$, in the case of symmetric $N \times N$ matrices one might take

$$K = \mathbb{S}(N)^+ = \left\{ A \in \mathbb{R}^{N \times N} \mid A = A^\top \geq 0 \right\}.$$

In the case of $\mathbb{R}^N$, one simple choice of cone could be

$$K = \mathbb{R}_+^N = \left\{ \xi \in \mathbb{R}^N \mid \xi_i \geq 0, \ 1 \leq i \leq N \right\},$$

which results in the *component-wise ordering* of $\mathbb{R}^N$. The choice of ordering is determined by the priority of objectives in the case of cost functionals with multi-dimensional range. A typical difficulty of the vectorial case is that *minima of the partial ordering may not exist*, and except for some prominent (but otherwise ill-behaved) orderings like the lexicographic ordering, this is usually the case. By minimum with respect to the ordering $\leq_K$ over a set $S \subseteq \mathbb{R}^N$ we mean a point $\xi \in S$ satisfying

$$\xi \leq_K \eta, \quad \text{for all } \eta \in S.$$

(We do not define the vectorial "inf", but this can be done in the obvious way.) The way to overcome this difficulty is to seek instead for *minimals*, usually called *Pareto Minimals* ([P]). A point $\xi \in S$ is a (Pareto) Minimal of the set $S \subseteq \mathbb{R}^N$ with respect to the ordering $\leq_K$ when

$$\text{For any } \eta \in S \ : \ \eta \leq_K \xi \quad \Longrightarrow \quad \eta = \xi.$$

Minima and minimals coincide for global (linear) orderings, but in general *they do not*. Unlike minima, minimals always exist and correspond to choices for which "there is no better available choice", while minima, if they exist, correspond to "the best available choice". Once again, this distinction has no bearing in the case of linear orderings. A well known method in order to construct minimals of a partial ordering is the so-called *scalarisation method* which is recalled later. Roughly, the idea of scalarisation is that

a partial ordering can be recovered from a family of scalar orderings along projections on lines generated by the direction in the dual cone $K^* \subseteq \mathbb{R}^N$.

By using scalarisation, one can construct a *manifold of Pareto minimals* which corresponds to the manifold of "unimprovable choices" and, motivated by the applications in Financial Mathematics, is usually called *the trade-off manifold* ("manifold" here is meant in the loose and not the strict mathematical sense, since it may lack the usual locally euclidean structure).

In this paper we commence a program which is along the lines of the scalar theory. Our central goal is to identify the appropriate vectorial extension of the concept of value function and derive the respective vectorial analogues of the HJ and HJB equations which are connected to the deterministic and stochastic control problems. The solutions of these PDEs would allow to construct feedback controls which optimally drive the system (1.1). To the best of our knowledge, this line of development via PDE theory has not been pursued before.

Interestingly, it turns out that, via the method of scalarisation, instead of a system of HJ or HJB equations as one might expect due to the vectorial nature

of the cost, we obtain a *parametric family of HJ/HJB equations, where the parameters λ are the unit directions inside the dual cone K^* relative to the selected partial ordering.* The respective Viscosity Solutions of these HJ/HJB equations are projections of the family of vectorial value functions $\{u^\lambda\}$ along directions normal to certain supporting hyperplanes. The value functions are Pareto minimals with respect to the ordering and form the *trade-off manifold inside the space of maps* $\mathbb{R}^n \times (0, \infty) \longrightarrow \mathbb{R}^N$. This manifold gives rise to a respective manifold of feedback controls $\{\alpha^\lambda\}$. The study of the topological structure of these manifolds of optimal choices seems to be an interesting topic in itself, but will not be considered in this introductory work.

This paper is organised as follows: In Section 2 we collect basic facts about cones, orderings, minimals and viscosity solutions, including the scalarisation method and the existence of Pareto minimals in the case of (finite-dimensional) Optimisation theory.

In Section 3 we consider the case of optimally controlling the system (1.3) without white noise with respect to the deterministic version of the vectorial functional (1.5). By introducing the appropriate value functions as (Pareto) minimals of the cost with respect to a fixed ordering (Definition 6), we prove their existence as a consequence of the scalarisation method (Lemma 7). Next, we derive the analogue of the Hamilton–Jacobi equation which arises in the deterministic vector case and show that appropriate projections of the value functions along lines generated by the dual cone are viscosity solutions of a family of HJ equations (Theorem 11 and Propositions 10 and 9) parameterised by the directions in the dual cone.

In Section 4 we turn our attention to the problem of stochastic optimal control of (1.1) via PDE theory, and extend the results of Section 3 to the case of non-trivial white noise. The results of this section are in correspondence to those of Section 3, the main difference being that here we have a family of 2nd order Hamilton–Jacobi–Bellman equations parameterised by the directions inside the dual cone and whose solutions optimally drive the system (1.1).

Finally, in Section 5 we examine particular applications of the theory to Linear-Quadratic models, showing that for the vectorial cost functional in certain directions the HJB problem can be reduced to solving a one-parameter family of matrix valued Riccati equations.

2. Cones, Generalised Ordering, Minimals and Viscosity Solutions

In this section we collect some rudimentary material related to generalised ordering, cones, minima, minimals, scalarisation and viscosity solutions. These notions and results we recall herein can be found in different guises sparsely distributed inside our references (and mostly proofless). We recall them here for the sake of completeness of the exposition and for the convenience of the reader.

2.1. Generalised Orderings with Respect to Cones.
Let $K \subseteq \mathbb{R}^N$ be a non-empty set. K is called a *cone* when

$$\xi \in K \quad \text{implies} \quad t\xi \in K, \ \text{for any } t \geq 0,$$

that is when

$$tK = K, \quad \text{for all } t \geq 0.$$

A cone K is called a *Proper Cone* when

- K is a closed convex set,

- the topological interior of K is non-empty: $\mathrm{int}(K) \neq \emptyset$,

- K contains no line: $\xi \in K$ and $-\xi \in K$ implies $\xi = 0$.

Some examples of proper cones are the ones given in the introduction, that is

- $K = [0, \infty)$, in $\mathbb{R}^1$,

- $K = \mathbb{S}(N)^+$, in $\mathbb{R}^{N \times N}$,

- $K = \mathbb{R}^N_+$, in $\mathbb{R}^N$.

However, the lexicographic cone $K_{\mathrm{lex}} \subseteq \mathbb{R}^N$, defined by

$$K_{\mathrm{lex}} := \{0\} \bigcup \left\{ \xi \in \mathbb{R}^N \ \middle|\ \xi_i > 0, \ \forall i \in \{1, ..., N\} \right\} \bigcup$$
$$\left\{ \xi \in \mathbb{R}^N \ \middle|\ \exists k \in \{1, ..., N-1\} \ : \ \xi_1 = \cdots = \xi_k = 0, \ \xi_{k+1} > 0 \right\}$$

is not a proper cone. Every proper cone $K \subseteq \mathbb{R}^N$ induces a *partial ordering* "$\geq_K$", given by

$$\eta \geq_K \xi \quad \Longleftrightarrow \quad \eta - \xi \in K.$$

Obviously, $\eta \leq_K \xi$ means $-\eta \geq_K -\xi$. The respective *strict ordering* "$>_K$" is defined analogously:

$$\eta >_K \xi \quad \Longleftrightarrow \quad \eta - \xi \in \mathrm{int}(K),$$

but will not be used in this work. Properness of the cone implies that the relation $\geq_K \subseteq \mathbb{R}^N \times \mathbb{R}^N$ is actually a partial ordering compatible with the topological and linear structure of $\mathbb{R}^N$:
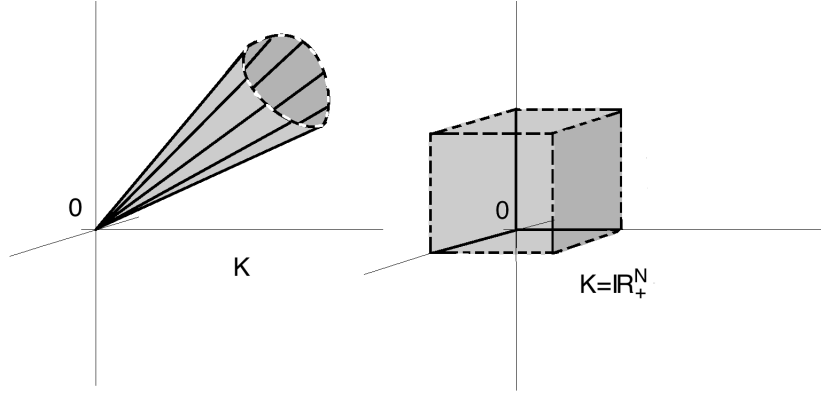

Figure 1.

- $\xi \leq_K \xi$,

- $\xi \leq_K \eta$ and $\eta \leq_K \zeta$ imply $\xi \leq_K \zeta$,

- $\xi \leq_K \eta$ and $\xi \geq_K \eta$ imply $\xi = \eta$,

- $\xi \leq_K \eta$ and $t \geq 0$ imply $t\xi \leq_K t\eta$,

- $\xi_m \leq_K \eta_m$ and $\xi_m \to \xi$, $\eta_m \to \eta$ as $m \to \infty$, imply $\xi \leq_K \eta$,

- $\xi' \leq_K \eta'$ and $\xi'' \leq_K \eta''$ imply $\xi' + \xi'' \leq_K \eta' + \eta''$.

2.2. **Mimina, Minimals and their Geometric Interpretation.** Let $S \subseteq \mathbb{R}^N$ be a non-empty set and suppose we are given a partial ordering "$\leq_K$" generated by a proper cone $K \subseteq \mathbb{R}^N$. By mimicking the scalar case, one may define the minimum of S with respect to the ordering "$\leq_K$" as a point $\xi \in S$ such that

$$\xi = \min S \quad \Longleftrightarrow \quad \xi \leq_K \eta, \quad \text{for all } \eta \in S.$$

One may also define the infimum of the set S as the minimum of the closure $\overline{S}$ of S, that is as the point $\xi \in \mathbb{R}^N$ such that

$$\xi = \inf S \quad \Longleftrightarrow \quad \xi = \min \overline{S}.$$

The minimum, if is exists, it is unique. In a more compact form, its definition reads

$$\xi = \min S \quad \Longleftrightarrow \quad S \subseteq \xi + K,$$

that is, ξ is the minimum of S if and only if it is contained in the translate of the cone K with vertex at ξ. The basic problem for the notion of minimum is that in general does not exist since only sets with very special structure possess it. The way to overcome this difficulty is to seek instead for *(Pareto) Minimals*. A point $\xi \in S$ is a (Pareto) Minimal of the set $S \subseteq \mathbb{R}^N$ with respect to the ordering $\leq_K$ when it satisfies

$$\left. \begin{array}{c} \eta \in S \text{ and} \\ \eta \leq_K \xi \end{array} \right\} \quad \Longrightarrow \quad \eta = \xi.$$
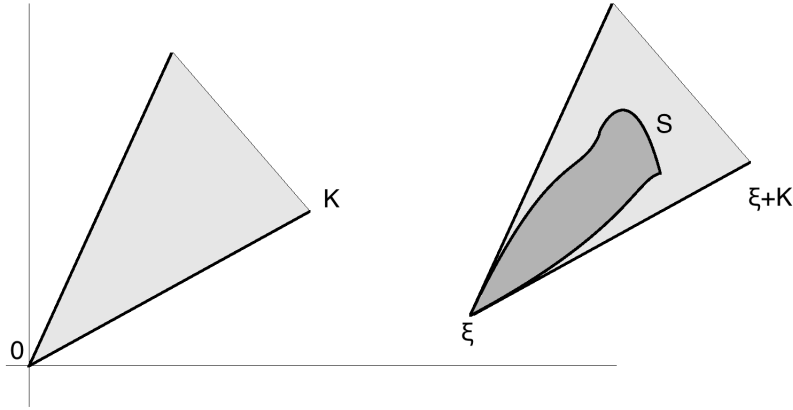

Figure 2.

Minima and minimals coincide for global (linear) orderings, but in general *they do not*. The geometric characterisation of minimals is

$$(\xi - K) \cap S = \{\xi\},$$

that is, ξ is a minimal of S if and only if the reflected translated cone $\xi - K$ with vertex at ξ intersects S only at ξ. Obviously one can more generally define the minimal of a set S as a point ξ *not necessarily contained in S* by considering the closure $\overline{S}$ in the place of S, but we will not go into that.
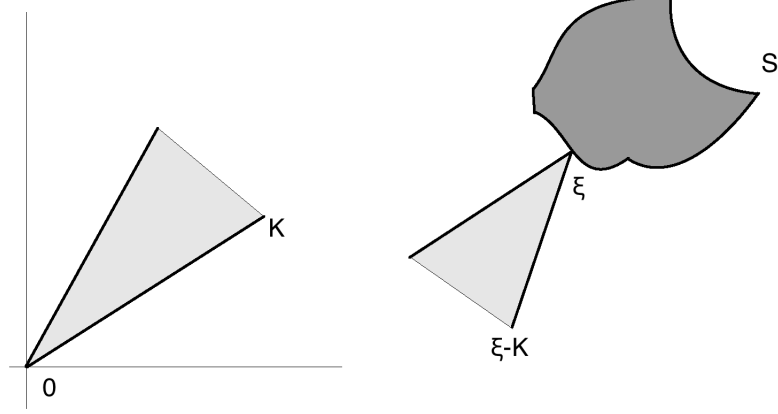
Figure 3.

We will shortly see that any closed set S possesses at least one minimal element with respect to an ordering generated by a proper cone.

2.3. **Dual Cones and Dual Inequalities.** A central concept in this context is duality. Given a cone $K \subseteq \mathbb{R}^N$, we define its dual cone as

$$K^* := \left\{ \eta \in \mathbb{R}^N \ \big| \ \eta \cdot \xi \geq 0, \ \forall \xi \in K \right\}.$$

Geometrically, $\eta \in K^*$ if and only if η is the inwards pointing vector to a halfspace supporting K at the origin (see Figure 4). As usual, dual objects satisfy better properties than the objects themselves. A cone is called *self-dual* if it coincides with its dual $K = K^*$. Simple properties of dual cones are

- K^* is closed and convex (although K might not be),

- $K' \subseteq K''$ implies $K''^* \subseteq K'^*$,

- if $\mathrm{int}(K) \neq \emptyset$, then K^* contains no non-trivial line,

- K^{**} coincides with the closed convex hull of K:

$$K^{**} = \overline{\mathrm{co}}\,(K)$$

The standard examples of cones given in the introduction are proper and self-dual:

- $[0, \infty) = ([0, \infty))^*$, in $\mathbb{R}^1$,

- $\mathbb{S}(N)^+ = (\mathbb{S}(N)^+)^*$, in $\mathbb{R}^{N \times N}$,

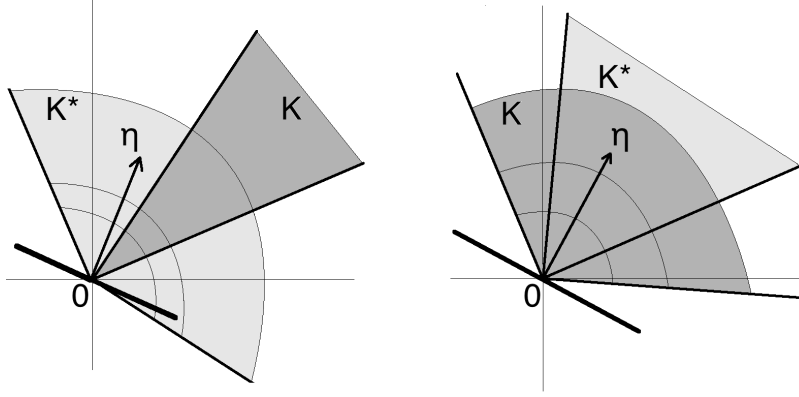- $\mathbb{R}_+^N = (\mathbb{R}_+^N)^*$, in $\mathbb{R}^N$.

Figure 4.

The only less trivial equality is the middle one, and follows by observing that $A \geq 0$ in $\mathbb{S}(N)$ if and only $A : B := A_{ij} B_{ij} \geq 0$ for all $B \in \mathbb{S}(N)^+$.

The dual cone induces a partial ordering itself on $\mathbb{R}^N$, in general different from the ordering induced by K, defined as

$$\xi \geq_{K^*} \eta \quad \Longleftrightarrow \quad \xi - \eta \in K^*.$$

One of the main utilities of the dual objects is that they *allow to characterise the ordering via a family of ordinary scalar orderings with respect to projections on the lines generated by directions in the dual cones.* Accordingly, we have

Lemma 1 (Orderings via duality). *Let $K \subseteq \mathbb{R}^N$ be a proper cone and K^* its dual. Then, for any $\xi, \eta \in \mathbb{R}^N$, we have the equivalence*

$$\xi \geq_K \eta \quad \Longleftrightarrow \quad \lambda \cdot \xi \geq \lambda \cdot \eta, \ \ for \ all \ \lambda \in K^*.$$

Proof of Lemma 1. We may assume $N \geq 2$, since the case $N = 1$ is trivial. Fix $\xi, \eta \in \mathbb{R}^N$ and suppose first that $\xi \geq_K \eta$. By definition, this means $\xi - \eta \in K$. Hence, by definition of K^*, for any $\lambda \in K^*$ we have $\lambda \cdot (\xi - \eta) \geq 0$, which is what we want.

Conversely, suppose that for any $\lambda \in K^*$ we have $\lambda \cdot (\xi - \eta) \geq 0$. For the sake of contradiction assume that $\xi - \eta \notin K$. Since K is a convex set, the projection on K

$$\mathrm{Proj}_K \ : \ \mathbb{R}^N \longrightarrow K,$$

is uniquely defined. Since $\xi - \eta \notin K$, we have $\mathrm{Proj}_K(\xi - \eta) \neq \xi - \eta$ and hence we may consider the 2-dimensional plane Π passing through the origin and the points $\xi - \eta$ and $\mathrm{Proj}_K(\xi - \eta)$. Consider now the orthogonal matrix $O \in \mathrm{O}(N, \mathbb{R})$ which leaves the $N-2$-dimensional orthogonal complement of Π invariant and coincides with the clockwise rotation by $\pi/2$ on Π with respect to the orientation generated by the frame $\{\mathrm{Proj}_K(\xi - \eta), \xi - \eta\}$ (See Figure 5). Then, we have

$$\lambda_0 := O \, \mathrm{Proj}_K(\xi - \eta) \ \in \ K^*,$$

since the orthogonal complement of λ_0 in $\mathbb{R}^N$ is a hyperplane which supports K at the origin. However, we have $\lambda_0 \cdot (\xi - \eta) < 0$ because the angle between $\xi - \eta$ and λ_0 is greater than $\pi/2$, which contradicts our assumption. Hence, we obtain

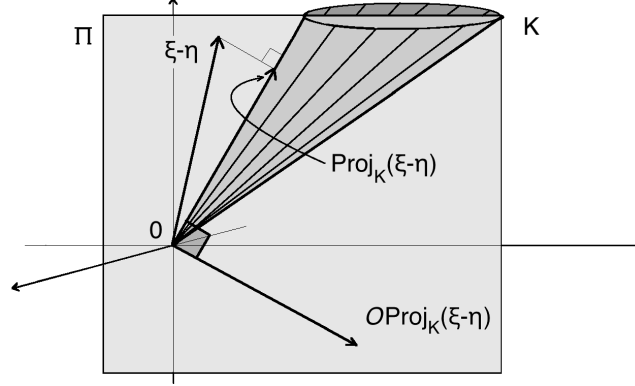$\xi - \eta \in K$, or equivalently $\xi \geq_K \eta$. $\hfill \square$

Figure 5.

2.4. **The Method of Scalarisation and Existence of Minimals.** Using the duality result of Lemma 1 above, we now derive characterisations of minima and minimals of a set with respect to an ordering generated by a proper cone.

Lemma 2 (Scalarisation of Minima and Minimals)**.** *Let $S \subseteq \mathbb{R}^N$ be a non-empty closed set and let $K \subseteq \mathbb{R}^N$ be a proper cone with K^* its dual cone.*
    Then,

   (1) *(Minima) ξ is the minimum of S with respect to the ordering $\leq_K$ if and only if for all $\lambda \in K^*$, ξ is the minimum of all the linear scalar functions*

$$\eta \mapsto \lambda \cdot \eta \; : \; \mathbb{R}^N \longrightarrow \mathbb{R},$$

   over the set S.

   (2) *(Minimals)*
   - *If for some $\lambda \in int(K^*)$, ξ is the minimum of the linear scalar function*

$$\eta \mapsto \lambda \cdot \eta \; : \; \mathbb{R}^N \longrightarrow \mathbb{R},$$

   over the set S, then ξ is a minimal of S with respect to the ordering $\leq_K$.
   - *Conversely, if in addition the set S is convex ($S = co(S)$) and ξ is a minimal element of S, then for any $\lambda \in K^*$, ξ is the minimum of the linear scalar function*

$$\eta \mapsto \lambda \cdot \eta \; : \; \mathbb{R}^N \longrightarrow \mathbb{R},$$

   over the set S.

Remark 3. In view of (1) above, it follows that minima of a set in general do not exist due to *obstructions* which can be rephrased as the requirement to have simultaneous minimisation of a family of linear function and the minimum being realised at the same point for all the functions. On the other hand, (2) says that for every direction strictly inside the dual cone, minimising the projection along this line leads to a minimal point for the set. The converse however to this statement is true in a weaker form and convexity plays a crucial role to that.

Lemma 2 leads immediately to the following important consequence:

Corollary 4 (Existence of Minimals). *Let $K \subseteq \mathbb{R}^N$ be a proper cone, K^* its dual cone and $\leq_K$ the ordering generated by K. Let also S be a compact non-empty set.*

Then, there exists at least one minimal element $\xi \in S$ of the set S with respect to the ordering $\leq_K$.

Moreover, for any $\lambda \in int(K^*)$, consider the supporting hyperplane of S which is normal to λ and such that λ points inside the halfspace which contains S. Then, the touching point belongs to the set of minimals of S.
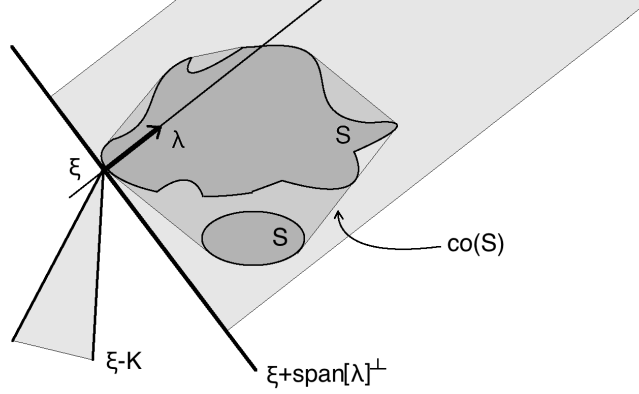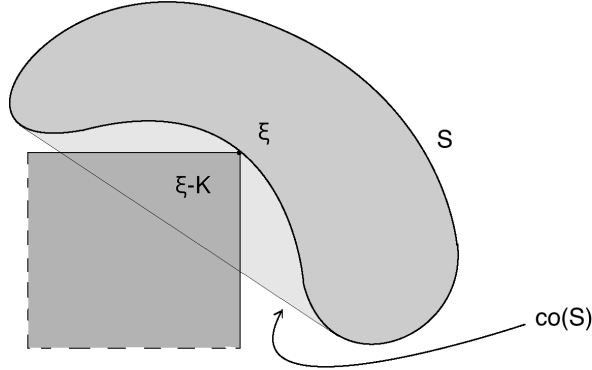

Figure 6.

It follows from the corollary that only the only the contact points of S with its convex hull $\overline{\text{co}}(S)$ are "attainable" candidate minimals by the scalarisation method, namely the set of points $S \cap \partial(\overline{\text{co}}(S))$.

Figure 7. ξ is a minimal with respect to "$\leq_{\mathbb{R}^2_+}$" in $\mathbb{R}^2$, but not attainable via scalarisation.

Proof of Lemma 2. We prove only the first statement of (2), which is the only one we will use in the sequel. The proof of the rest claims can be found e.g. in [BV], page 54. We fix $\lambda \in K^*$ and suppose that ξ is the minimum of the linear functional $\eta \mapsto \lambda \cdot \eta$ over S. We claim that $\{\xi\} = (\xi - K) \cap S$, which is equivalent to the statement that ξ is a minimal of S. If this is not the case, then there is an $\eta \in S$ with $\eta \neq \xi$ such that $\eta \in (\xi - K) \cap S$. Since $\eta \in \xi - K$ and $\xi \neq \eta$, we have that $\xi - \eta \in K \setminus \{0\}$. Since $\lambda \in int(K^*)$, the definition of the dual cone implies

that $\lambda \cdot (\xi - \eta) < 0$ (because λ is in the interior and can not be normal to $\xi - \eta$), which is a contradiction. The claim ensues. $\square$

2.5. **Viscosity Solutions of Hamilton-Jacobi-Bellman Equations.** Here we recall the definition of appropriate "weak" solutions for fully nonlinear (1st and) 2nd order degenerate elliptic (and parabolic) PDE. The notion is standard and can be found in many sources, e.g. in the standard reference [CIL]. However, the version below is taken from the introductory text [K]. The main difference is that degenerate ellipticity is assumed as monotonicity, instead of anti-monotonicity. Let

$$\mathscr{F} \; : \; (\Omega \subseteq \mathbb{R}^n) \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{S}(n) \longrightarrow \mathbb{R}$$

be a continuous function and consider the PDE

$$\mathscr{F}(\cdot, u, Du, D^2 u) \, = \, 0, \quad u \, : \, \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}.$$

We assume that F satisfies

$$X \, \leq \, Y \quad \Longrightarrow \quad \mathscr{F}(x, r, p, X) \, \leq \, \mathscr{F}(x, r, p, Y),$$

for all $(x, r, p) \in \Omega \times \mathbb{R} \times \mathbb{R}^n$, $X, Y \in \mathbb{S}(n)$.

Definition 5 (Viscosity Solutions). Let $u \in C^0(\Omega)$, $\Omega \subseteq \mathbb{R}^n$, and consider the (degenerate elliptic) PDE

$$\mathscr{F}(\cdot, u, Du, D^2 u) \, = \, 0.$$

(a) We say that u is a Viscosity Subsolution of the PDE (or a Viscosity Solution of $\mathscr{F}(\cdot, u, Du, D^2 u) \geq 0$) when

$$\left. \begin{array}{c} u - \psi \leq 0 = (u - \psi)(x_0) \\ \text{on a ball } \mathbb{B}_r(x_0) \subseteq \Omega, \\ x_0 \in \Omega, \; \psi \in C^2(\mathbb{R}^n) \end{array} \right\} \quad \Longrightarrow \quad \mathscr{F}(x_0, \psi(x_0), D\psi(x_0), D^2\psi(x_0)) \geq 0.$$

(b) We say that u is a Viscosity Supersolution of the PDE (or a Viscosity Solution of $\mathscr{F}(\cdot, u, Du, D^2 u) \leq 0$) when

$$\left. \begin{array}{c} u - \phi \geq 0 = (u - \phi)(y_0) \\ \text{on a ball } \mathbb{B}_r(y_0) \subseteq \Omega, \\ y_0 \in \Omega, \; \phi \in C^2(\mathbb{R}^n) \end{array} \right\} \quad \Longrightarrow \quad \mathscr{F}(y_0, \phi(y_0), D\phi(y_0), D^2\phi(y_0)) \leq 0.$$

(c) We say that u is a Viscosity Solution, when it is both a Viscosity Subsolution and a Viscosity Supersolution.

3. Deterministic Optimal Control and Hamilton-Jacobi PDE

In this section we consider the first main theme of this paper, namely the deterministic optimal control of the initial value problem[1]

(3.1)
$$\begin{cases} \dot{\chi}(s) \, = \, F(\chi(s), \alpha(s)), \quad t < s < T, \\ \chi(t) \, = \, x, \end{cases}$$

---

[1]Note that we conform with standard conventions as e.g. in the textbook [E] and (by a time reversal) we have an initial instead of a terminal condition.

where $x \in \mathbb{R}^n$, $A \subseteq \mathbb{R}^m$ is a compact set and $F : \mathbb{R}^n \times A \longrightarrow \mathbb{R}^n$ a continuous map. α is a measurable map is the class $\mathcal{A}$ given by (1.2). Optimal controllability is meant with respect to the vectorial cost functional

$$(3.2) \qquad C_{x,t}[\alpha] \; = \; g\Big(\chi\big(T,x,\alpha(T)\big)\Big) \; + \; \int_t^T h\Big(\chi\big(s,x,\alpha(s)\big),\alpha(s)\Big) \, \mathrm{d}s$$

where h, g are given maps and χ denotes the flow map of (3.1), which we may abbreviate to merely $\chi(s)$, suppressing the dependence in x and α. We assume that F, g, h satisfy the assumption (1.7).

We begin with the definition of the appropriate vectorial minimals we will use as our vectorial extension of the value function.

Definition 6 (Pareto Minimals of the Vectorial Cost). Let $A \subseteq \mathbb{R}^m$ be a compact set, $\mathcal{A}$ the class given by (1.2) and F, g, h given maps which satisfy the assumption (1.7) with $n, m, N \geq 1$.

Consider the initial value problem (3.1) and the vectorial cost functional (3.2). Suppose further that we are given a partial ordering $\leq_K$ on $\mathbb{R}^N$ generated by a proper cone $K \subseteq \mathbb{R}^N$.

We say that the map $u : \mathbb{R}^n \times (0,T) \longrightarrow \mathbb{R}^N$ is a vectorial value map (or a Pareto Minimal) of the cost (3.2) when for any $(x,t) \in \mathbb{R}^n \times (0,T)$, $u(x,t)$ is a Pareto Minimal of the set

$$S_{x,t} \; := \; \overline{\big\{C_{x,t}[\alpha] \mid \alpha \in \mathcal{A}\big\}} \subseteq \mathbb{R}^N,$$

that is, when

$u(x,t) \in S_{x,t}$ and for all $\eta \in S_{x,t}$ for which $\eta \leq_K u(x,t)$, we have $\eta = u(x,t)$.

By using the results of the previous section, in particular Corollary 4 and Lemma 2, we readily have the next

Lemma 7 (Existence of Pareto Minimals of the Vectorial Cost). *In the setting of Definition 6, there exists at least one value maps which is a Pareto Minimals of the vectorial cost. In addition, any direction $\lambda \in K^*$ generates a Pareto Minimal by considering the supporting hyperplane (of the convex hull of) $S_{x,t}$ which is normal to λ.*

Remark 8. We note that in general, not all the Pareto minimals of the ordering can be realised via scalarisation, but instead only those on the "extreme points" of the convex hull (see Section 2).

We now have the next result, regarding the vector value functions:

Proposition 9 (Existence and properties of the vectorial value function). *Given a partial ordering $\leq_K$ on $\mathbb{R}^N$ generated by a proper cone $K \subseteq \mathbb{R}^N$, let K^* be its dual cone (see Section 2). Then, for any $\lambda \in K^*$ there exists a vectorial value function*

$$(3.3) \qquad\qquad u^\lambda \; : \; \mathbb{R}^n \times (0,T) \longrightarrow \mathbb{R}^N$$

which is a Pareto Minimal of the functional (3.2) with respect to the ordering $\leq_K$ and also satisfies

$$(3.4) \qquad \lambda \cdot u^\lambda(x,t) \; = \; \inf_{\alpha \in \mathcal{A}} \big\{\lambda \cdot C_{x,t}[\alpha]\big\}, \quad x \in \mathbb{R}^n, \, 0 < t < T.$$

The scalar function $\lambda \cdot u^\lambda$ is the projection of the vectorial value map u^λ along the direction generated by a supporting hyperplane of $S_{x,t}$ normal to λ. In addition, $\lambda \cdot u^\lambda$ is bounded and Lipschitz continuous in both variables.

Proof of Proposition 9. The proof is a direct consequence of Definition 6, Remark 8, the scalarisation method of Section 2 and an application of the corresponding scalar result given in [E, Section 10.3.2]. $\qquad\square$

We now state a proposition whose proof is straightforward extension of the corresponding scalar result given in [E, Section 10.3.2] and is based on our previous analysis.

Proposition 10 (Deterministic dynamic optimality). *Let $A \subseteq \mathbb{R}^m$ be a compact set, $\mathcal{A}$ the class given by (1.2) and F, g, h are given maps which satisfy the assumption (1.7) with $n, m, N \geq 1$. Then, for any $\delta > 0$ such that $t + \delta \leq T$, we have*

$$(3.5) \quad \lambda \cdot u^\lambda(x,t) \,=\, \inf_{\alpha \in A} \left\{ \lambda \cdot u^\lambda\big(\chi(t+\delta), t+\delta\big) \,+\, \int_t^{t+\delta} \lambda \cdot h(\chi(s), \alpha(s))\,\mathrm{d}s \right\},$$

where χ is the solution of the differential equation (3.1).

The above result is a consequence of the dynamic programming principle. Roughly speaking, it states that optimal cost through the entire time interval $[t, T]$ can be achieved by running optimally in $[t, t+\delta]$ and then restarting the problem at time $t + \delta$ with initial conditions $\chi(t + \delta)$.

We now come to the main result of this section, which is the following:

Theorem 11 (Vectorial Optimal Control, Pareto Minimals, Viscosity Solutions of HJ PDE). *Let the conditions of Propositions 9 and 10 hold. Consider the initial value problem (3.1) and the vectorial cost functional (3.2).*

Then, for any unit direction $\lambda \in K^*$, the projection $\lambda \cdot u^\lambda(x,t)$ is a scalar function and a Viscosity Solution of the initial value problem

$$\begin{cases} v_t \,+\, H^\lambda(\cdot, Dv) \,=\, 0, & in\ \mathbb{R}^n \times (0, T), \\ \hspace{3.2cm} v \,=\, \lambda \cdot g, & on\ \mathbb{R}^n \times \{0\}. \end{cases}$$

Here the Hamiltonian H^λ is defined by

$$H^\lambda(x, p) \,:=\, \min_{a \in A} \big\{ F(x,a) \cdot p \,+\, \lambda \cdot h(x,a) \big\}.$$

Proof of Theorem 11. The proof is a direct consequence of the definition of Minimals, the results of duality/scalarisation of Section 2, Propositions 9 and 10 and the standard scalar control theory (see e.g. Evans [E], pages 550-560). It is also a special case of the proof of Theorem 14 which will be proved is some detail in the subsequent section in the more general stochastic case with non-trivial noise. $\quad\square$

4. Stochastic Optimal Control and Hamilton-Jacobi-Bellman PDE

In this section we extend the ideas of Section 3 to the stochastic case. For the rudimentary facts of the theory of SDEs needed in this paper we refer to Evans [E2].

We consider now the case of optimal control of the system of stochastic differential equation

$$(4.1) \qquad \begin{cases} \mathrm{d}\chi(s) = F\big(\chi(s), \alpha(s)\big)\,\mathrm{d}s \; + \; \sigma\big(\chi(s), \alpha(s)\big)\,\mathrm{d}W(s), \quad t < s < T, \\ \quad \chi(t) = x, \end{cases}$$

where we use the same notation as in Section 3. The extra ingredients now are $W(s)$ which is an m dimensional system of independent Wiener processes and $\sigma : \mathbb{R}^n \times A \longrightarrow \mathbb{R}^{n \times m}$ which is a continuous map. In the stochastic case, the flow map χ has to be interpreted as a stochastic process. This process is defined over a *probability space* which is a triple $(\mathbb{R}^n, \mathscr{A}, P)$, with $\mathscr{A}$ a σ–algebra of subsets of $\mathbb{R}^n$ and P a probability measure over $\mathbb{R}^n$.

The stochastic process χ is said to solve the stochastic differential equation (4.1) if

$$(4.2) \qquad \chi(s) \;=\; x + \int_t^s F\big(\chi(r), \alpha(r)\big)\,\mathrm{d}r \; + \; \int_t^s \sigma\big(\chi(r), \alpha(r)\big)\,\mathrm{d}W(r),$$

interpreted as an Itô integral, holds almost surely for all $s \in (t, T)$.

The stochastic version of the vectorial cost functional has the same form as the deterministic, however is given in terms of an expectation

$$(4.3) \qquad C_{x,t}[\alpha] \;=\; \mathbb{E}\left[g\Big(\chi\big(T, x, \alpha(T)\big)\Big) \; + \; \int_t^T h\Big(\chi\big(s, x, \alpha(s)\big), \alpha(s)\Big) ds \right],$$

which is defined as

$$(4.4) \qquad\qquad\qquad \mathbb{E}[X] := \int_{\mathbb{R}^n} X \,\mathrm{d}P.$$

The following two Propositions are the stochastic equivalents to Propositions 9 and 10:

Proposition 12 (Existence and properties of the vectorial value function)**.** *Given a partial ordering $\leq_K$ on $\mathbb{R}^N$ generated by a proper cone $K \subseteq \mathbb{R}^N$, let K^* be its dual cone (see Section 2). Then, for any $\lambda \in K^*$ there exists a vectorial value function*

$$u^\lambda \;:\; \mathbb{R}^n \times (0, T) \longrightarrow \mathbb{R}^N$$

which is a Pareto Minimal of the functional (4.3) with respect to the ordering $\leq_K$ and also satisfies

$$(4.5) \qquad \lambda \cdot u^\lambda(x, t) \;=\; \inf_{\alpha \in \mathcal{A}} \big\{ \lambda \cdot C_{x,t}[\alpha] \big\}, \quad x \in \mathbb{R}^n, \, 0 < t < T.$$

In addition, the value function is bounded and Lipschitz continuous in both variables.

Proposition 13 (Stochastic dynamic optimality)**.** *Let $A \subseteq \mathbb{R}^m$ be a compact set, $\mathcal{A}$ the class given by (1.2) and F, g, h, σ given maps which satisfy the assumption (1.7) with $n, m, N \geq 1$. A consequence of the dynamic programming principle is that for any $\delta > 0$ such that $t + \delta \leq T$, we have*

$$\lambda \cdot u^\lambda(x, t) \;=\; \inf_{\alpha \in A} \left\{ \mathbb{E}\left[\lambda \cdot u^\lambda(\chi(t+\delta), t+\delta) \; + \; \int_t^{t+\delta} \lambda \cdot h(\chi(s), \alpha(s))\,\mathrm{d}s \right] \right\},$$

where χ is the solution to the system of stochastic differential equations (4.1).

The main result of this section is:

Theorem 14 (Viscosity Solutions of HJB PDE). *Consider the stochastic initial value problem (4.1) and the vectorial cost functional (4.3). For each $\lambda \in K^*$, the projection $\lambda \cdot u^\lambda(x,t)$ is a scalar function and a Viscosity Solution of the initial value problem*

$$(4.6) \qquad \begin{cases} v_t + H^\lambda(\cdot, Dv, D^2v) = 0, & in \ \mathbb{R}^n \times (0,T), \\ v = \lambda \cdot g, & on \ \mathbb{R}^n \times \{0\}, \end{cases}$$

where the Hamiltonian H^λ is defined by

$$(4.7) \quad H^\lambda(x, p, Z) := \min_{a \in A} \left\{ \frac{1}{2} \sigma(x,a)\sigma(x,a)^\top : Z + F(x,a) \cdot p + \lambda \cdot h(x,a) \right\}.$$

Proof. The proof of this result is just an extension of [E, p.557 Thm 2] to the stochastic vectorial case. We include it here for completeness. Throughout this proof for convenience we will denote $u := \lambda \cdot u^\lambda$. We begin by noting that (4.5) implies

$$(4.8) \qquad u(x,T) \ = \ \inf_{a \in A} \left\{ \mathbb{E}[\lambda \cdot g(\chi(T))] \right\} \ = \ \lambda \cdot g(x).$$

To show that u is the viscosity solution of (4.6) we must verify the conditions given in Definition 5. To that end, without loss of generality let $v \in C^2(\mathbb{R}^n \times (0,T))$ and assume there exist (x_0, t_0) such that

$$(4.9) \qquad u(x,t) - v(x,t) \leq u(x_0,t_0) - v(x_0,t_0), \ \text{when} \ |x - x_0| + |t - t_0| \leq \epsilon.$$

We now want to show that

$$(4.10) \qquad v_t(x_0,t_0) \ + \ H^\lambda(x_0, Dv, D^2v) \ \geq \ 0.$$

Assume for the sake of contradiction that there exists an $a \in A$ such that

$$(4.11) \qquad \begin{aligned} \mathscr{L}v(x_0,t_0) &:= \ v_t(x_0,t_0) \ + \ Dv(x_0,t_0) \cdot F(x_0,a) \\ &\quad + \frac{1}{2} D^2v(x_0,t_0) : \left(\sigma(x_0,a)\sigma(x_0,a)^\top \right) \ + \ \lambda \cdot h(x_0,a) \ < \ 0 \end{aligned}$$

for $|x - x_0| + |t - t_0| \leq \epsilon$. Let $\chi(s)$ be the solution of the stochastic differential equation

$$(4.12) \qquad \begin{aligned} \mathrm{d}\chi(s) &= \ F\big(\chi(s),a\big)\,\mathrm{d}s \ + \ \sigma\big(\chi(s),a\big)\,\mathrm{d}W(s), \quad t < s < T, \\ \chi(t_0) &= \ x_0, \end{aligned}$$

where a is now taken to be a contant control. Now choose $\delta \in (0, \epsilon)$ such that

$$(4.13) \qquad |\chi(s) - x_0| \ \leq \ \epsilon, \quad \text{when } s \in [t_0, t_0 + \delta],$$

then by (4.11) we know

$$(4.14) \qquad \mathscr{L}v(\chi(s), s) \ < \ 0, \ \text{ for } s \in [t_0, t_0 + \delta].$$

Using (4.9) we have that

$$(4.15) \qquad \begin{aligned} u\big(\chi(t_0 + \delta), t_0 + \delta\big) \ - \ u(x_0,t_0) \ &\leq \ v\big(\chi(t_0 + \delta), t_0 + \delta\big) \ - \ v(x_0,t_0) \\ &\leq \ \int_{t_0}^{t_0+\delta} \mathrm{d}v(\chi(s), s). \end{aligned}$$

From Proposition 13, we have

$$(4.16) \qquad u(x_0,t_0) \ \leq \ \mathbb{E}\left[u(\chi(t_0 + \delta), t_0 + \delta) + \int_{t_0}^{t_0+\delta} \lambda \cdot h(\chi(s),a)\,\mathrm{d}s \right].$$

Combining (4.15) and (4.16) and dividing through by δ we see

$$
\begin{aligned}
(4.17) \qquad 0 \;&\leq\; \frac{1}{\delta}\mathbb{E}\left[\int_{t_0}^{t_0+\delta} \mathrm{d}v(\chi(s),s)\right] \;+\; \frac{1}{\delta}\mathbb{E}\left[\int_{t_0}^{t_0+\delta} \lambda\cdot h(\chi(s),a)\,\mathrm{d}s\right] \\
&=:\; \mathscr{E}_1(\delta) \;+\; \mathscr{E}_2(\delta).
\end{aligned}
$$

The first term, in view of the Itô chain rule is

$$
\begin{aligned}
(4.18)\qquad \mathscr{E}_1(\delta) \;&=\; \frac{1}{\delta}\mathbb{E}\left[\int_{t_0}^{t_0+\delta} \mathrm{d}\big(v(\chi(s),s)\big)\right] \\
&=\; \frac{1}{\delta}\mathbb{E}\bigg[\int_{t_0}^{t_0+\delta} v_t(\chi(s),s)\,\mathrm{d}s + Dv(\chi(s),s)\cdot \mathrm{d}\chi(s) \\
&\qquad\qquad + \frac{1}{2}D^2v(\chi(s),s):\big(\mathrm{d}\chi(s)\,\mathrm{d}\chi(s)^\top\big)\bigg] \\
&=\; \frac{1}{\delta}\mathbb{E}\bigg[\int_{t_0}^{t_0+\delta} v_t(\chi(s),s) + Dv(\chi(s),s)\cdot F(\chi(s),s) \\
&\qquad\qquad + \frac{1}{2}D^2v(\chi(s),s):\big(\sigma(\chi(s),a)\sigma(\chi(s),a)^\top\big)\,\mathrm{d}s \\
&\qquad\qquad + Dv(\chi(s),s)\cdot\sigma(\chi(s),a)\,\mathrm{d}W(s)\bigg].
\end{aligned}
$$

Now using Fubini's theorem we have

$$
\begin{aligned}
(4.19)\qquad \lim_{\delta\to 0}\mathscr{E}_1(\delta) \;&=\; v_t(x_0,t_0) \;+\; Dv(x_0,t_0)\cdot F(x_0,a) \\
&\qquad + \frac{1}{2}D^2v(x_0,t_0):\big(\sigma(x_0,a)\sigma(x_0,a)^\top\big).
\end{aligned}
$$

The second term, again by Fubini's theorem, gives

$$
\begin{aligned}
(4.20)\qquad \lim_{\delta\to 0}\mathscr{E}_2(\delta) \;&=\; \lim_{\delta\to 0}\frac{1}{\delta}\mathbb{E}\left[\int_{t_0}^{t_0+\delta} \lambda\cdot h(\chi(s),a)\,\mathrm{d}s\right] \\
&=\; \lambda\cdot h(x_0,a).
\end{aligned}
$$

Substituting (4.19) and (4.20) into (4.17), we have

$$
(4.21)\qquad\qquad\qquad 0 \;\leq\; \mathscr{L}v(x_0,t_0),
$$

which contradicts (4.11). Consequently, u is a Viscosity Subsolution of (4.6). In a similar fashion one can show that u is a Viscosity Supersolution to (4.6), hence u is a Viscosity Solution of (4.6). $\qquad\square$

5. Applications to Linear-Quadratic problems

We conclude this exposition with an application of the main ideas to a sample problem, that of Linear-Quadratic models. This is a prototypical example arising in optimal control applicable in many areas, data assimilation being one such example [GM]. To demonstrate the approach and the differences between the scalar case, i.e., when the cost functional is scalar ($N = 1$) and the non scalar case, i.e., when the cost functional is vectorial ($N > 1$). We will present both cases.

In the following suppose the optimal control problem with dynamics governed by the SDE (4.1) together with cost functional (4.3) take the specific form

$$d\chi(s) = (A\chi(s) + B\alpha(s))\,ds + \sigma\,dW(s)$$

(5.1)
$$C_{x,t}[\alpha] = \mathbb{E}\left[\chi(T)^\top Q_T\,\chi(T) + \int_t^T \frac{1}{2}\alpha(s)^\top R\,\alpha(s) + \frac{1}{2}\chi(s)^\top Q\,\chi(s)ds\right].$$

This is precisely the Linear-Quadratic model so named as the dynamics are described via a *linear* SDE and *quadratic* cost functional.

Theorem 15 (Linear-Quadratic models)**.** *Let $N = 1$, then suppose $A, B, Q, Q_T, R \in \mathbb{R}^{n^2}$ are symmetric, positive definite matrices and the noise $\sigma \in \mathbb{R}^n$ is additive Gaussian. Then the solution to the stochastic Linear-Quadratic model (5.1) is itself quadratic and takes the form:*

(5.2)
$$u(x,t) = \frac{1}{2}x^\top U(t)\,x + b(t),$$

if and only if $b(t) \in \mathbb{R}, U(t) \in \mathbb{R}^{n^2}$ solve the following (backward in time) initial value problems

(5.3)
$$\begin{cases} -\dot{b}(s) = \frac{1}{2}\sigma\sigma^\top : U(s), & t < s < T, \\ b(T) = 0. \end{cases}$$
$$\begin{cases} -\dot{U}(s) = Q + 2\,A^\top U(s) - U(s)^\top B(R^{-1})^\top B^\top U(s), & t < s < T, \\ U(T) = Q_T. \end{cases}$$

Proof. Making use of Theorem 14 we have that (as $N = 1$) the solution, u, is a viscosity solution to (4.6). Now using the specific form of u from (5.2) we may compute that

(5.4)
$$u_t(x,s) = \frac{1}{2}x^\top \dot{U}(s)\,x + \dot{b}(s)$$
$$Du(x,s) = U(s)\,x$$
$$D^2 u(x,s) = U(s).$$

Substituting this into the Hamilton-Jacobi-Bellman equation (4.6) we see that

(5.5)
$$0 = \frac{1}{2}x^\top \dot{U}(s)\,x + \dot{b}(s) + \min_{a\in A}\left\{\frac{1}{2}\sigma\sigma^\top : U(s) + (A\,x + B\,a)^\top U(s)\,x\right.$$
$$\left. + \frac{1}{2}a^\top R\,a + \frac{1}{2}x^\top Q\,x\right\}.$$

Since the Hamiltonian is quadratic in a then the minimum can be explicitly computed. Indeed,

(5.6)
$$a^*(x,s) = -R^{-1}B^\top U(s)\,x$$

is the minimiser of the Hamiltonian, thus we see, using elementary but tedious linear algebra that

(5.7)

$$
\begin{aligned}
0 &= \frac{1}{2} x^\top \dot{U}(s)\, x + \dot{b}(s) + \frac{1}{2}\sigma\sigma^\top : U(s) + \big(A\,x + B\left(-R^{-1}\, B^\top\, U(s)\, x\right)\big)^\top U(s)\, x \\
&\quad + \frac{1}{2}\big(R^{-1}\, B^\top\, U(s)\, x\big)^\top R\left(R^{-1}\, B^\top\, U(s)\, x\right) + \frac{1}{2} x^\top\, Q\, x \\
&= \frac{1}{2} x^\top\, \dot{U}(s)\, x + \dot{b}(s) + \frac{1}{2}\sigma\sigma^\top : U(s) \\
&\quad + x^\top \left[2 A^\top U(s) - U(s)^\top B\big(R^\top\big)^{-1} B^\top U(s) + q\right] x.
\end{aligned}
$$

Now since we require the equality to hold for all x the value function u can only be quadratic if and only if the backward in time ODEs are satisfied. The end time conditions occur from matching the final time value of the cost functional, that is, considering $C_{x,T}$ in (5.1), concluding the proof. $\qquad\square$

Remark 16 (The deterministic LQ model is equivalent)**.** Notice the equation for U in (5.3) is a matrix valued Riccati equation and is independent of σ. This means that the optimal control (a^* in the proof of Theorem 15) must be the same in the deterministic case and hence the same Riccati equation can be derived. In this scenario u is called the Linear-Quadratic Regulator.

Theorem 17 (The vectorial case)**.** *Let $N > 1$, then suppose $A, B \in \mathbb{R}^{n^2}$ are symmetric, positive definite matrices and the noise σ is constant in time additive Gaussian. In addition suppose that $Q, Q_T, R \in \mathbb{R}^{n^2 N}$. Now for every $\lambda \in K^*$ such that $\lambda \cdot Q, \lambda \cdot Q_T, \lambda \cdot R \in \mathbb{R}^{n^2}$ remain symmetric positive definite matrices, the projection $\lambda \cdot u^\lambda$ is quadratic and takes the form*

(5.8)
$$
u^\lambda(x,t) = \frac{1}{2} x^\top\, U^\lambda(t)\, x + b(t),
$$

if and only if $b(t) \in \mathbb{R}, U^\lambda(t) \in \mathbb{R}^{n^2}$ solve the following (backward in time) initial value problems

(5.9)
$$
\begin{cases}
-\dot{b}(s) = \frac{1}{2}\sigma\sigma^\top : U(s), & t < s < T, \\
b(T) = 0.
\end{cases}
$$
$$
\begin{cases}
-\dot{U}^\lambda(s) = \lambda \cdot Q + 2\, A^\top\, U^\lambda(s) - U^\lambda(s)^\top B\left(\big(\lambda \cdot R\big)^{-1}\right)^\top B^\top\, U^\lambda(s), & t < s < T, \\
U^\lambda(T) = \lambda \cdot Q_T.
\end{cases}
$$

Proof. The proof of this fact consists of applying the same arguments as that of Theorem 15, together with those presented in Theorem 14. $\qquad\square$

6. Conclusion

In this work we have summarised an approach aimed at proving existence of solutions to optimal control problems with *vectorial* cost functionals. The main idea of this approach is that, given a vectorial cost functional, one can *select* a direction in $\mathbb{R}^N$ along which to minimise the cost as long as that direction does not leave the problem degenerate. Then for any direction one prove existence of a viscosity solution and, in the case of LQ control, write down solutions. It was

our primary intention of deriving the model equations which we conjectured to be systems of HJB equations but discovered, using this approach, are scalarised equations. It was not the goal of this exposition to show the uniqueness of such a solution, which is completely nontrivial, if indeed true at all.

References

AFIJ. H. Abou-Kandil, G. Freiling, V. Ionescu, G. Jank, *Matrix Riccati Equations in Control and Systems Theory*, Birkhaüser, 2003.

BCD. M. Bardi, I. Capuzzo-Dolcetta, *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*, Modern Birkhäuser Classics, 2008.

BO. T. Basar, G.J. Olsder, *Dynamic Noncooperative Game Theory*, Classics in Applied Mathematics 23, 2nd Edition, SIAM, 1999.

BF. R. Bellman, K. Fan, *On systems of linear inequalities in Hermitian matrix variables*, in: V. L. Klee (editor), *Convexity*, vol. VII, Proceedings of the Symposia in Pure Mathematics, Amer. Math. Soc. 1963, 1 11.

BV. S. Boyd, L.Vandenberghe, *Convex Optimization*, Cambridge University Press, 7th printing, 2009.

BKR. R. S. Burachik, C. Y. Kaya, M. M. Rizvi, *A New Scalarization Technique to Approximate Pareto Fronts of Problems with Disconnected Feasible Sets*, Journal of Optimization Theory and Applications 162 (2014), 428–446.

CIL. M. G. Crandall, H. Ishii, P.L. Lions, *User's Guide to Viscosity Solutions of 2nd Order Partial Differential Equations*, Bulletin of the AMS 27 (1992), 1 - 67.

CHY. G.Chen, X. Huang, X. Yang, *Vector Optimization: Set-valued and Variational Analysis*, Lecture Notes in Economics and Mathematical Systems, Springer, 2005.

DD. I. Das, J.E. Dennis, *Normal-Boundary Intersection: A New Method for Generating the Pareto Surface in Nonlinear Multicriteria Optimization Problems*, SIAM J. Optim., 8 (2006), 631-657.

D. G. Debreu, *Theory of Value: An Axiomatic Analysis of Economic Equilibrium*, Yale University Press, 1959.

E. L.C. Evans, *Partial differential equations*, AMS, Graduate Studies in Mathematics Vol. 19, 1998.

E2. L.C. Evans, *An Introduction to Stochastic Differential Equations*, AMS pulications, ISBN-13: 978-1-4704-1054-4, 2013.

FR. W. H. Fleming, R. W. Rishel, *Deterministic and Stochastic Optimal Control*, Stochastic Modelling and Applied Probability, Springer, corrected 2nd Printing, 1983.

FS. W. H. Fleming, H. M. Soner, *Controlled Markov Processes and Viscosity Solutions*, Springer, Stochastic Modelling and Applied Probability, Springer, 2nd edition, 2010.

GC. D. Ghosh, D. Chakraborty, *A new Pareto set generating method for multi-criteria optimization problems*, Operations Research Letters Vol. 42 (2014), 514-521.

GM. M. Ghil, P. Malanotte-Rizzoli, *Data assimilation in meteorology and oceanography*, Adv. Geophys Vol. 33, 141-266 (1991).

G. A. Guigue, *Set-Valued Return Function and Generalized Solutions for Multiobjective Optimal Control Problems*, SIAM Journal on Control and Optimization,51, 2379 - 2405 (2013).

H. M. I. Henig, *Vector-valued dynamic programming*, SIAM J. Control and Optimization, Vol. 21, No. 3 (1983) 490 - 499.

I. K. Isii, *Inequalities of the types of Chebyshev and Cramér-Rao and mathematical programming*, Annals of The Institute of Statistical Mathematics 16 (1964), 277 - 293.

K. N. Katzourakis, *An introduction to Viscosity Solutions for fully nonlinear PDE with applications to Calculus of Variations in L^∞*, Springer Brief monograph, BCAM series, in press.

Ke. H. J. Kenneth, *Stochastic Optimal Control Theory with Application in Self-Tuning Control*, Lecture Notes in Control and Information Sciences, Vol. 117, Springer, 1989.

L. P. L. Lions, *Generalized solutions of Hamilton-Jacobi equations*, Research Notes in Mathematics 69, Pitman, 1982.

L1. D. G. Luenberger, *Quasi-convex programming*, SIAM Journal on Applied Mathematics 16 (1968), 1090 - 1095.

L2. D. G. Luenberger, *Optimization by Vector Space Methods*, John Wiley & Sons, 1969.

L3. D. G. Luenberger, *Linear and Nonlinear Programming*, Addison-Wesley, second edition, 1984.

L4. D. G. Luenberger, *Microeconomic Theory*, McGraw-Hill, 1995.

MGGJ. B. Martin, A. Goldsztejn, L. Granvilliers, C. Jermann, *On continuation methods for non-linear bi-objective optimization: towards a certified interval-based approach*, Journal of Global Optimization (2014) DOI: 10.1007/s10898-014-0201-3.

M. E. S. Maskin, *The theory of implementation in Nash equilibria: a survey*, in *Social Goals and Social Organization: Essays in Memory of Elisha Pazner*, by L. Hurwicz (Ed.), D. Schmeidler (Ed.), H. Sonnenschein (Ed.), Cambridge University Press, 1985.

O. C. Olech, *Existence Theorems for Optimal Problems with Vector-Valued Cost Function*, Transactions of the Amer. Math. Soc. Vol. 136 (1969), 159 - 180.

P. V. Pareto, *Manual of Political Economy*, A. M. Kelley Publishers, 1971. Translated from the French edition, first published in Italian in 1906.

RK. N. Rastegar, E. Khorram, *A combined scalarizing method for multiobjective programming problems*, European Journal of Operational Research 236 (2014), 229237.

RBG. G. Reynoso-Meza, J. Sanchis, X. Blasco, S. Garcia-Nieto, *Physical programming for preference driven evolutionary multi-objective optimization*, Applied Soft Computing Vol. 24 (2014), 341-362.

S. M. E. Salukvadze, *Vector-Valued Optimization Problems in Control Theory*, Academic Press, 1979.

T. D. F. M. Torres, *A Noether Theorem on Unimprovable Conservation Laws for Vector-Valued Optimization Problems in Control Theory*, ArXiv preprint.

Department of Mathematics and Statistics, University of Reading, Whiteknights, PO Box 220, RG6 6AX, Reading, UK

E-mail address: n.katzourakis@reading.ac.uk, t.pryer@reading.ac.uk