A FRAMEWORK FOR THE DYNAMIC PROGRAMMING PRINCIPLE AND MARTINGALE-GENERATED CONTROL CORRESPONDENCES

ROMAN FAYVISOVICH AND GORDAN ŽITKOVIĆ

ABSTRACT. We construct an abstract framework in which the dynamic programming principle (DPP) can be readily proven. It encompasses a broad range of common stochastic control problems in the weak formulation, and deals with problems in the "martingale formulation" with particular ease. We give two illustrations; first, we establish the DPP for general controlled diffusions and show that their value functions are viscosity solutions of the associated Hamilton-Jacobi-Bellman equations under minimal conditions. After that, we show how to treat singular control on the example of the classical monotone-follower problem.

1. INTRODUCTION

The goal of this paper is creating a probabilistic framework in which the dynamic programming principle (DPP) can be easily proved. To be useful, such a framework needs to be sufficiently powerful, so as to encompass as many stochastic control problems as possible, but also sufficiently simple, so that it is easily applied in a given situation. On a deeper level, our intention is to identify the fundamental properties stochastic control problems and their setups need to have in order for the DPP to hold. One of the many interesting things about (proving) the DPP is that its validity depends both on topological/measure theoretic properties of the underlying spaces (such as the Polish structure) and structural properties of the control problem (such as the ability to concatenate controls). A large part of this paper is a study of their interplay in the setting of filtered probability spaces and general formulations of stochastic-control problems.

Even though the dynamic programming principle has been introduced in the mid 20th century, or even earlier, (we point the reader to [Žit14] for a short historical overview), research related to DPP - especially in continuous time - underwent somewhat of a renaissance in the past several decades (see., e.g., [EK81], [Bor89], [FS93], [ST02a], [ST02b], [BV10], [BT11] [BN12], [ET13a], [ET13b] and [Žit14]).

1.1. **Our contributions.** Our starting point is the paper [Žit14] which focuses on two specific control problems and shows that they both satisfy the DPP. Therein, the so-called controlled Markov families

Date: June 1, 2019.

2010 *Mathematics Subject Classification.* 93E20, 60G44, 60J25.

Acknowledgments: The authors would like to thank Mihai Sîrbu and Kasper Larsen for valuable conversations and acknowledge the support by the National Science Foundation under Grants DMS-0706947 (2007 - 2010), DMS-09556194 (2010 - 2015), DMS-1107465 (2012 - 2017) and DMS-1516165 (2015 - 2018). Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation (NSF).

(families of *sets* of probability measures indexed by the elements of a state space) are introduced, and DPP is formulated as a natural analogue of the Markov property in that setting. That formulation helps identify three separate properties (already present in the literature, see, e.g., [ET13b, NvH13, Žit14]) of a controlled Markov family, called *analyticity, concatenability* and *disintegrability,* under which the DPP holds. On their own, these three properties do not amount to much more than a rephrasing of the DPP without making it much easier to establish. The present paper takes up the task of providing wide sufficient conditions for each of these three and, thus, for the validity of the DPP.

1.1.1. *Truncation- and truncation-concatenation spaces.* We begin by introducing the structure of a *truncated space* (T-space) which carries the structure of a "measurably-filtered space" with each $\mathcal{F}_t$ generated by a single, albeit, Polish-valued, random variable. Perhaps unexpectedly at first, virtually all (uncompleted) concrete filtrations used in probability and stochastic control turn out to be T-spaces; moreover, we show that some perks of canonical spaces C and D (such as Galmarino's test) extend to all T-spaces. Another added benefit is that sigma-algebras $\mathcal{F}_\tau$ corresponding to stopping times inherit the property of being generated by a single, Polish-valued random variable. This observation simplifies many of our proofs and provides further insight into the structure of T-spaces. Moreover, many natural constructions (such as products or subspaces) work well in the T-space context. This is particular important for our purposes as control problems come in a variety of forms, but are invariably built out of a smaller number of "probabilistic building blocks". In the same, categorical, worldview, a natural and useful notion of a morphism between T-spaces can be introduced.

If one adds a time-indexed family of binary operations to a T-space and imposes appropriate measurability and compatibility requirements, one obtains the structure of a *truncation-concatenation space* (TC-space). The idea is to abstract away the main properties that define the operation of concatenation in the context of the DPP. In addition to the model case of pasting of (right-) continuous paths, many other forms of concatenation are covered by TC-spaces. Indeed, while the state spaces of control problems typically involve the spaces of (right-, left-, ...) continuous trajectories, the spaces of controls are much less regular and need a more flexible framework. Just like in the case of T-spaces, one defines products, subspaces and structure-preserving maps (morphisms) between TC-spaces. Morphisms into the model space $D_{\mathbb{R}}$ of càdlàg trajectories play an especially important role later when we deal with martingale-generated controlled Markov families.

Once TC-spaces are set up, control problems are represented by *control correspondences*, i.e., correspondences that map each element of the sample space into a set of probability measures on it. In this context, one defines the notion of a value function of a control problem, as well as the properties of analyticity, concatenability and disintegrability which, together, imply (an abstract) DPP. It is, perhaps, interesting to note that no notion of a state is needed for the abstract DPP to hold. It can be introduced explicitly, as we often do, but its role is abstractly taken over by the notion of compatibility used to define a TC space.

1.1.2. *Martingale-generated control correspondences.* Our central claim is that truncation-concatenation spaces, together with a shift operator (which can be thought of as a partial inverse of concatenation and plays a central role in the study of disintegrability), provide a convenient framework on which a variety of

stochastic control problems can be posed and analyzed. Of course, the validity of the DPP will depend on the nature of the problem itself, but, as we show in examples, this amounts to a verification of a small number of easily checked intuitive conditions. Focusing mainly on control problems in their weak formulation, and the derived control correspondences, we identify two important cases in which these conditions are especially easy to check. One is when the probability of the future evolution is controlled directly, without the need for an intermediate "control process", as is the case, e.g., with pure singular-control problems. In the other, much larger, family of cases, explicit control processes are typically present, but their structure is such that access to the totality of all possible controlled dynamics is possible via a system of "well-behaved" constraints. Such constraints are often expressible in terms of the (local) martingale property of a class of real-valued càdlàg processes. The control correspondences constructed in this way are said to be *martingale-generated* as they correspond, loosely, to what is known as the martingale formulation of optimal control in the literature. The second third of the paper focuses on martingale-generated control correspondences on TC spaces and provides sufficient conditions on the structure of the constrains (by interpreting them as morphisms into the model space $D_{\mathbb{R}}$) for the DPP to hold.

1.1.3. *Examples.* The final third of the paper contains two examples meant to illustrate the versatility of our framework. The first one is the classical controlled-diffusion case which we consider in the weak formulation and place it in our setting as a martingale-generated control correspondence. We show that sufficient conditions established in the previous section apply in this case, and conclude that the DPP holds under minimal conditions on the coefficients and the form of the controls. We also demonstrate that value functions of such control problems are viscosity solutions of the corresponding Hamilton-Jacobi-Bellman equations, under slightly stronger conditions (continuity of coefficients and admissibility of locally constant controls). This partially generalizes several recent results in the literature, such as the "stochastic Perron" method of Bayraktar and Sîrbu (introduced in [BS12]) or the work of Bouchard and Touzi on the "weak DPP" (see [BT11]). The same class of problems - under a somewhat different set of assumptions - has already been treated by the authors of [ET13a, ET13b]. Like the present paper, they rely on the ability to pose an equivalent controlled martingale problem on a suitable canonical space and characterize the resulting control correspondence using at most countably many test functions.

Our second example is of singular type, and features a mildly generalized Monotone-Follower problem. Here, we not only show how to establish the DPP for a singular-control problem in our framework, but also showcase its flexibility. Indeed, we split the variables into two groups and deal with one directly, and with the other using the martingale-generated approach. These two are considered separate control problems (with separate control correspondences) until the very last moment when they are easily merged.

1.2. **Notation and conventions.** Both probabilistic and analytic tools - which often come with less-than-perfectly compatible notations and terminology - are used in this paper. For the convenience of the reader, we outline some of our major choices and conventions below.

Both probabilistic $\mathbb{E}^{\mathbb{P}}[X]$ and analytic $\int G\, d\mu$ notation for integration will be used. The former will appear mostly in examples, and the latter in the abstract part.

Many of our probability spaces come with **Polish** (completely metrizable, separable) sample spaces and Borel probability measures. When the Polish structure is present, measurability will always refer to the associated Borel σ-algebra, denoted by $\mathsf{Borel}(\Omega)$. The set of all probability measures on $\mathsf{Borel}(\Omega)$ is denoted by $\mathsf{Prob}(\Omega)$.

A subset A of a Polish space Ω is called **analytic** if it can be realized as a projection of a Borel subset of $\Omega \times \mathbb{R}$ onto Ω. We remind the reader that analytic subsets of Polish spaces are closed under countable unions, intersections and products, but not necessarily under complements. It will be important for us that each analytic set is in the **universal** σ-algebra - denoted by $\mathsf{Univ}(\Omega)$ - i.e., the family of all sets which belong to the completion $(\mathsf{Borel}(\Omega))^*_\mu$ for each $\mu \in \mathsf{Prob}(\Omega)$. We refer the reader to [Sri98] for all the necessary details concerning descriptive set theory (see also [BS78] for a thorough treatment of related topics in the context of the dynamic programming principle).

We topologize $\mathsf{Prob}(\Omega)$ with the topology of (probabilist's) weak convergence. This way, $\mathsf{Prob}(\Omega)$ becomes a Polish space. The following well-known fact, proved in a standard way via the monotone-class theorem, will be used throughout without mention: Let U and V be Polish spaces and let $f : U \times V \to [0, \infty]$ be a Borel-measurable function. The map

$$U \times \mathsf{Prob}(V) \ni (x, \mu) \mapsto \mathbb{E}^{\mu}[f(x, \cdot)] = \int_V f(x, y)\, \mu(dy)$$

is Borel measurable.

A probability measure defined on $\mathsf{Borel}(\Omega)$ admits a natural extension to $\mathsf{Univ}(\Omega)$. Similarly, our kernels *will always be universally measurable*. More precisely, for Polish spaces $\Omega, \tilde{\Omega}$, a map $\nu : \Omega \times \mathsf{Borel}(\tilde{\Omega}) \to [0, 1]$ is called a **kernel** if $\nu(\omega, \cdot) \in \mathsf{Prob}(\tilde{\Omega})$ for each $\omega \in \Omega$ and $\nu(\cdot, B)$ is a universally-measurable map on Ω, for each $B \in \mathsf{Borel}(\tilde{\Omega})$. Depending on the situation we use both notations $\nu(\omega, \cdot)$ and ν_ω for the probability measure associated by ν to ω.

A **standard Borel space** is, by definition, a measurable space which admits a measurable bijection to a Borel subset of some $\mathbb{R}^n$, whose inverse is also measurable (a **bimeasurable isomorphism**). All standard Borel spaces of the same cardinality are bimeasurably isomorphic, and so, each standard Borel space can be given a complete and separable (Polish) metric so that the induced measurable structure matches the original one. With this in mind, we talk of standard Borel spaces when only the measurable structure is relevant, and about Polish spaces when topological properties are required.

2. An abstract setting for the Dynamic Programming Principle (DPP)

Let the **time set** Time be either $[0, \infty)$ or $\mathbb{N}_0$. An overwhelming majority of applications will only use these two time sets, so we do not aim for greater generality. We do note that the results of this section will hold for

more general time structures (such as intersections with $[0,\infty)$ of Borel-measurable additive subgroups of $\mathbb{R}$).

2.1. **T-spaces (truncated spaces).** We start with the definition of T-spaces - a class of filtered probability spaces our analysis will be based on.

Definition 2.1 (T-spaces)**.** A filtered measurable space $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{t\in\mathsf{Time}})$ is called a **T-space** (or a **truncated space**) if

(1) $(\Omega, \mathcal{F})$ is a standard Borel space and $\mathcal{F} = \bigvee_{t\in\mathsf{Time}} \mathcal{F}_t$.
(2) there exists a family $\{T_t\}_{t\in\mathsf{Time}}$ of maps $T_t : \Omega \to \Omega$ - called a **truncation** - such that
   (a) $(t,\omega) \mapsto T_t(\omega)$ is (jointly) measurable,
   (b) $T_t \circ T_s = T_{s\wedge t}$ for all $s, t \in \mathsf{Time}$, and
   (c) $\mathcal{F}_t = \sigma(T_t)$ for each $t \in \mathsf{Time}$.

For notational reasons, we always add the identity map $T_\infty = \mathrm{Id}$ to any truncation. Moreover, we often use the alternative notation $\omega_{\leq t}$ for $T_t(\omega)$.

2.2. **First examples of T-spaces.** All T-spaces are necessarily countably generated, so not every filtered probability space can be endowed with the structure of a T-space. Nevertheless, as our examples below aim to show, many spaces used in stochastic analysis and optimal stochastic control are natural T-spaces. When it is necessary to make a distinction, we take $\mathsf{Time} = [0,\infty)$ and leave it to the reader to make the necessary minor adjustments needed for the case $\mathsf{Time} = \mathbb{N}_0$. Once we describe various natural constructions involving T-spaces in subsection 2.4 below, the reader will be able to produce many more examples.

2.2.1. *The path space D_E.* Let E be a Polish space, and let D_E denote the family of all càdlàg functions from Time to E. For $t \in \mathsf{Time}$, we define the truncation map $T_t : D_E \to D_E$ by

$$T_t(\omega)(s) = \omega(t \wedge s) \text{ for } s \in \mathsf{Time}, \tag{2.1}$$

so that (2b) of Definition 2.1 holds. It is well-known that D_E is a Polish space under the Skorokhod topology. The map T_t is Skorokhod-continuous, and therefore, measurable. Hence, as a Caratheodory function, $T : \mathsf{Time} \times \Omega \to \Omega$ is (jointly) measurable. The filtration $\mathcal{F}_t = \sigma(T_t), t \in \mathsf{Time}$ clearly coincides with the (raw) filtration generated by the coordinate maps $\omega \mapsto \omega(t)$.

2.2.2. *Path spaces G_E, C_E and $\mathrm{Lip}_{\mathbb{R}}^{L,x_0}$.* Analogous constructions can be performed on the space G_E of left-continuous and right limited paths from Time to E, or on the space C_E of continuous paths. Both of these are given the Skorokhod topology (and the induced Borel structure), which, in the case of C_E reverts to the usual topology of locally uniform convergence. Unless specified otherwise, these spaces (and their subspaces) will always be endowed with the standard truncation given by (2.1).

We will also have use for the space $\mathrm{Lip}_{\mathbb{R}}^{L,x_0}$ consisting of all functions $x : [0,\infty) \to \mathbb{R}$ such that $x(0) = x_0$ and $|x(t) - x(s)| \leq L\,|t-s|$ for all $s, t \in [0,\infty)$. It is easy to see that $\mathrm{Lip}_{\mathbb{R}}^{L,x_0}$ is also a T-space with the standard truncation.

2.2.3. *The space $\mathbb{L}_A^0$ and related spaces.* Let A be a standard Borel space, let λ be the Lebesgue measure (or any other Radon measure) on $[0,\infty)$, and let $\hat{\lambda}$ denote an equivalent probability measure on $[0,\infty)$ (e.g., $\hat{\lambda}(dt) = e^{-t}\,\lambda(dt)$, when λ is the Lebesgue measure). We define $\mathbb{L}_A^0$ as the set of all λ-a.e.-equivalence classes of Borel functions $\alpha : [0,\infty) \to A$. Given a bimeasurable isomorphism $\phi : A \to [-1,1]$ (which exists thanks to the standard Borel property of A) we metrize $\mathbb{L}_A^0$ by

$$d(\alpha,\beta) = ||\phi(\alpha) - \phi(\beta)||_{\mathbb{L}^1(\hat{\lambda})}.$$

This way, $\Omega = \mathbb{L}_A^0$ becomes a Polish space and a natural truncation on it is defined by

$$T_t(\alpha) = \begin{cases} \alpha_u, & u < t \\ \phi^{-1}(0), & u \geq t. \end{cases}$$

We note that the equivalence class of the right-hand side depends on α only through its equivalence class, and that, while d and the induced Polish topology depend on the choice of ϕ and $\hat{\lambda}$, the resulting standard Borel structure does not. The choice of this particular ϕ makes it easy to show that T_t is jointly measurable; indeed, it will be continuous under d in both of its arguments.

Once the space $\mathbb{L}_A^0$ is constructed, one can easily show that many of subsets (such as the $\mathbb{L}^p$ spaces when $A = \mathbb{R}$) are also T-spaces.

2.2.4. *Spaces of measures.* For a metrized Polish space U, let $\mathcal{M}^{\#}(U)$ be the family of all boundedly-finite Borel measures on U, i.e,. those measures μ such that $\mu(B) < \infty$, as soon as B is a bounded Borel set. There exists a metric on $\mathcal{M}^{\#}(U)$, whose topology coincides with the topology of weak convergence when restricted on measures supported by a fixed bounded set (see [DVJ03, Section A2.6, p. 402] for the proof of this and other statements about the space $\mathcal{M}^{\#}(U)$ we make below). Under the full topology induced by this metric, called the $w^{\#}$-topology, $\mathcal{M}^{\#}(U)$ becomes a Polish space. Moreover, a sequence $\{\mu_n\}_{n\in\mathbb{N}}$ in $\mathcal{M}^{\#}(U)$ converges if and only if $\int f\,d\mu_n \to \int f\mu$ for each bounded and continuous function $f : \Omega \to \mathbb{R}$ which vanishes outside a bounded set. The Borel σ-algebra on $\mathcal{M}^{\#}(U)$ is generated by the evaluation maps $\mu \mapsto \mu(A)$, where A ranges over a family of all bounded Borel subsets of U. The subsets $\mathcal{M}^{f}(U)$ and $\mathcal{M}^{p}(U) = \mathsf{Prob}(U)$ of $\mathcal{M}^{\#}(U)$, consisting only of finite or probability measures (respectively), are easily seen to be Borel subsets of $\mathcal{M}^{\#}(\Omega)$, and, therefore, standard Borel spaces themselves.

For a Polish space E, we set $\Omega = \mathcal{M}^{*}(U)$, where $U = [0,\infty) \times E$ and $* \in \{\#, f, p\}$. The truncation maps are given by

$$\mu_{\leq t}(A) = \mu\Big(([0,t] \times E) \cap A\Big), \text{ for } t \in [0,\infty), A \in [0,\infty) \times E.$$

With the filtration generated by the maps T_t, it is clear that $\vee_t \mathcal{F}_t$ is the Borel σ-algebra on Ω. The only remaining property from Definition 2.1 is (2a), for which it is sufficient to note that for any boundedly

supported function f we have $\int f\, d\mu_{\leq t} = \int f \mathbf{1}_{[0,t]\times E}\, d\mu$. Indeed, it follows that $(t,\mu) \mapsto \mu_{\leq t}$ is a Caratheodory function as it is right continuous in t and measurable in μ.

2.2.5. *Predictable truncations.* In many the examples above, it is possible to define several different truncations on the same underlying Polish space. For example, in the case of the canonical space D_E, we may set

$$T'_t(\omega)(s) = \begin{cases} \omega_s, & s < t \\ \omega_{t-}, & s \geq t \end{cases}.$$

It is easily checked that T'_t is indeed, a truncation on D_E; we call it the **predictable truncation**.

2.3. **Truncating at stopping times.** Given a T-space $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{t\in\mathsf{Time}})$, let the set of all stopping times be denoted by Stop. The index set for the family of truncation operators can be extended to Stop by setting

$$T_\tau(\omega) = T_{\tau(\omega)}(\omega) \text{ for } \tau \in \mathsf{Stop} \text{ and } \omega \in \Omega,$$

where the convention that T_∞ is the identity map is used. As is the case with deterministic times, the notation $T_\tau(\omega)$ will often be replaced by the less cumbersome (and more suggestive) $\omega_{\leq\tau}$.

Proposition 2.2. *For all $t \in \mathsf{Time}$, $\omega \in \Omega$, $\tau, \kappa \in \mathsf{Stop}$ and we have*

(1) T_τ and T_κ are measurable maps on Ω and $T_\tau \circ T_\kappa = T_{\tau\wedge\kappa}$
(2) $\sigma(T_\tau) = \{A \in \mathcal{F} : T_\tau^{-1}(A) = A\}$, and

$$\text{“}A \in \sigma(T_\tau)\text{”} \quad \textit{is equivalent to} \quad \text{“}\omega \in A \Leftrightarrow \omega_{\leq\tau} \in A\text{”}$$

(3) $\tau(\omega) = \tau(T_\tau(\omega))$, and hence τ is $\sigma(T_\tau)$-measurable
(4) $\sigma(T_\tau) = \mathcal{F}_\tau$, where $\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}_t, \forall t \in \mathsf{Time}\}$
(5) Let $(S, \mathcal{S})$ be a standard Borel space. An $(\mathcal{F}, \mathcal{S})$-measurable map $Z : \Omega \to S$ is $(\mathcal{F}_\tau, \mathcal{S})$-measurable if and only if $Z \circ T_\tau = Z$.

Proof.

(1) Measurability of T_τ follows directly from the measurability of stopping times and the joint measurability of $(t,\omega) \mapsto T_t(\omega)$ on $(\mathsf{Time} \cup \{\infty\}) \times \Omega$. Applying Definition 2.1, part (2b) pointwise for $t = \tau(\omega)$ and $s = \kappa(\omega)$ gives $T_\tau \circ T_\kappa = T_{\tau\wedge\kappa}$.
(2) By part (1) we have $T_\tau = T_\tau \circ T_\tau$ for each $\tau \in \mathsf{Stop}$, and so for any $A \in \mathcal{F}$, we have

$$A = T_\tau^{-1}(B) \text{ for some } B \in \mathcal{F} \quad \Leftrightarrow \quad A = T_\tau^{-1}(A).$$

Furthermore the condition $A = T_\tau^{-1}(A)$ is equivalent to:

$$\omega \in A \quad \Leftrightarrow \quad \omega_{\leq\tau} \in A$$

(3) Fix $\omega \in \Omega$, let $t = \tau(\omega)$, and let $A = \{\tau = t\}$. Since $\tau \in \mathsf{Stop}$, then $A \in \mathcal{F}_t = \sigma(T_t)$. Combining part (2) with the fact that $\omega \in A$ implies $T_t(\omega) \in A$. Therefore:

$$\tau(T_{\tau(\omega)}(\omega)) = \tau(T_t(\omega)) = t = \tau(\omega)$$

(4) For the forward inclusion, let $A \in \sigma(T_\tau)$. Thanks to (2) above, we have $A = T_\tau^{-1}(A)$. Therefore for all $t \in \mathsf{Time}$ we have:

$$\begin{aligned} A \cap \{\tau \leq t\} &= \{\omega \in \Omega \,:\, T_{\tau(\omega)}(\omega) \in A, \tau(\omega) \leq t\} \\ &= \{\omega \in \Omega \,:\, T_{\tau(\omega)\wedge t}(\omega) \in A, \tau(\omega) \leq t\} = T_{\tau\wedge t}^{-1}(A) \cap \{\tau \leq t\} \in \mathcal{F}_t, \end{aligned}$$

where we used the fact that $T_{\tau\wedge t} = T_{\tau\wedge t} \circ T_t$ is $\mathcal{F}_t$-measurable. Therefore $A \in \mathcal{F}_\tau$, and hence $\sigma(T_\tau) \subset \mathcal{F}_\tau$.

For the backward inclusion, let $A \in \mathcal{F}_\tau$. By part (2), it suffices to show:

$$\omega \in A \quad \Leftrightarrow \quad \omega_{\leq\tau} \in A$$

First suppose $\omega \in A$ and let $t = \tau(\omega)$. Since $A \in \mathcal{F}_\tau$, then $\omega \in A \cap \{\tau \leq t\} \in \mathcal{F}_t$. Applying (2) to $A \cap \{\tau \leq t\}$ gives $\omega_{\leq\tau} \in A \cap \{\tau \leq t\} \subset A$.

For the other direction, suppose $\omega_{\leq\tau} \in A$. By part (3) we have $\tau(\omega_{\leq\tau}) = \tau(\omega)$ and hence $\omega_{\leq\tau} \in A \cap \{\tau \leq t\} \in \mathcal{F}_t$. Applying (2) to $A \cap \{\tau \leq t\}$ gives $\omega \in A \cap \{\tau \leq t\} \subset A$.

(5) If $Z = Z \circ T_\tau$, then Z is $\mathcal{F}_\tau$-measurable as a measurable transformation of the $\mathcal{F}_\tau$-measurable map T_τ. Conversely, if Z is $\mathcal{F}_\tau$-measurable, the standard Borel property and the Doob-Dynkin lemma guarantee the existence of a measurable map $\zeta : \Omega \to S$ such that $Z = \zeta \circ T_\tau$. A composition with T_τ yields that

$$Z \circ T_\tau = \zeta \circ T_\tau \circ T_\tau = \zeta \circ T_\tau = Z. \qquad \square$$

2.4. **Constructions on T-spaces.**

Next, we describe several natural notions and constructions on T-spaces, as well as various operations that produce new T-spaces from the old ones. For the remainder of this subsection, let $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{t\in\mathsf{Time}})$ and $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}} = \{\tilde{\mathcal{F}}_t\}_{t\in\mathsf{Time}})$ be two T-spaces, with truncations $\{T_t\}_{t\in\mathsf{Time}}$ and $\{\tilde{T}_t\}_{t\in\mathsf{Time}}$, respectively.

2.4.1. *Structure-preserving maps.* A useful structure-preserving notion in the case of T-spaces turns out to be non-anticipation:

Definition 2.3. A measurable map $F : \Omega \to \tilde{\Omega}$ is said to be **non-anticipating** if it is $(\mathcal{F}_t, \tilde{\mathcal{F}}_t)$-measurable, i.e. $F^{-1}(\tilde{\mathcal{F}}_t) \subseteq \mathcal{F}_t$ for each $t \in \mathsf{Time}$.

We have the following characterization using the truncation maps:

Proposition 2.4. *A measurable map $F : (\Omega, \mathcal{F}) \to (\tilde{\Omega}, \tilde{\mathcal{F}})$ is non-anticipating if and only if*

$$\tilde{T}_t \circ F \circ T_t = \tilde{T}_t \circ F \text{ for all } t \in \mathsf{Time}.$$

Proof. By Proposition 2.2 part (2) we have $\tilde{\mathcal{F}}_t = \sigma(\tilde{T}_t) = \tilde{T}_t^{-1}(\tilde{\mathcal{F}})$, and by part (5) we have $\tilde{T}_t \circ F$ is $\mathcal{F}_t$-measurable if and only if $\tilde{T}_t \circ F \circ T_t = \tilde{T}_t \circ F$. Therefore for all $t \in \mathsf{Time}$:

$$\begin{aligned} F^{-1}(\tilde{\mathcal{F}}_t) \subset \mathcal{F}_t \quad &\Leftrightarrow \quad F^{-1}(\tilde{T}_t^{-1}(\tilde{\mathcal{F}})) \subset \mathcal{F}_t \\ &\Leftrightarrow \quad \tilde{T}_t \circ F \text{ is } \mathcal{F}_t\text{-measurable} \\ &\Leftrightarrow \quad \tilde{T}_t \circ F \circ T_t = \tilde{T}_t \circ F \end{aligned}$$ □

Remark 2.5. One could also consider an alternative notion of a structure-preserving map where we require that $\tilde{T}_t \circ F = F \circ T_t$ for all $t \in \mathsf{Time}$. Proposition 2.4 and the fact that $\tilde{T}_t \circ \tilde{T}_t = \tilde{T}_t$ imply that T-morphisms are non-anticipating, but the converse is not true.

2.4.2. *T-subspaces.* We say that a T-space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}} = \{\tilde{\mathcal{F}}_t\}_{t\in\mathsf{Time}})$ is a T**-subspace** of $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{t\in\mathsf{Time}})$ if $\tilde{\Omega} \subseteq \Omega$ and $\tilde{\mathcal{F}}_t \subseteq \mathcal{F}_t$, for all $t \in \mathsf{Time}$. As the following result show, subsets preserved by truncation inherit a structure of a T-space:

Proposition 2.6. *Let $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{t\in\mathsf{Time}})$ be a T-space, and let Ω' be a measurable subset of Ω with the property that $T_t(\Omega') \subseteq \Omega'$, for all $t \in \mathsf{Time}$. Then the family $\{T'_t\}_{t\in\mathsf{Time}}$ given by $T'_t = T_t|_{\Omega'}$, is a truncation, and the filtered space $(\Omega', \mathcal{F}', \{\mathcal{F}'_t\}_{t\in\mathsf{Time}})$, given by $\mathcal{F}' = \{B \in \mathcal{F} : B \subseteq \Omega'\}$, $\mathcal{F}_t = \sigma(T'_t)$, $t \in \mathsf{Time}$, is a T-space and a subspace of $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{t\in\mathsf{Time}})$.*

Proof. Clearly $(\Omega', \mathcal{F}')$ is a subspace of $(\Omega, \mathcal{F})$. To satisfy Definition 2.9 of T-spaces, note that part (1) follows from the construction of Ω' and $\mathcal{F}'$, and the properties of part (2) are passed down from T to T'. □

Example 2.7. Truncation operators on D_E leave invariant several important measurable subsets of D_E. Among the examples are

(1) C_E, the family of all everywhere continuous elements of D_E,
(2) $D_E^{E_0}$, the family of paths in D_E which start from a point in E_0, and
(3) D_{E^F}, the family of paths in D_E stopped once they hit the closed subset F of E.
(4) D^{fv} $(D^{\uparrow}, D^{\downarrow})$, the family of all paths in $D_{\mathbb{R}}$ all of whose components are of finite variation (nondecreasing, nonincreasing)
(5) $\mathrm{Lip}_{\mathbb{R}}^{L}$, the family of all Lipschitz continuous maps from $[0, \infty)$ to $\mathbb{R}$, with the Lipschitz constant at most L.

More examples can be produced by various intersections of the above sets.

2.4.3. *Products.* T-spaces behave well under products, too. Indeed, the standard Borel space $\hat{\Omega} = \Omega \times \tilde{\Omega}$ admits a natural truncation given by the family $\{\hat{T}_t\}_{t\in\mathsf{Time}}$ of maps on $\hat{\Omega}$ defined by

$$\hat{T}_t(\omega, \tilde{\omega}) = (T_t(\omega), \tilde{T}_t(\tilde{\omega})). \tag{2.2}$$

The resulting T-space $\hat{\Omega}$, together with the natural filtration generated by $\{\hat{T}_t\}_{t\in\mathsf{Time}}$, is called the **product** of the truncated spaces Ω and $\tilde{\Omega}$. It is not difficult to see that the same construction can be applied to countable products of truncated spaces.

2.4.4. *State maps.* A measurable map $X : \Omega \to E$, where E is a Polish space is called a **state map**. Such maps define a class of progressively measurable E-valued stochastic processes on Ω via

$$X_t(\omega) = X(T_t(\omega)), t \in \mathsf{Time} \cup \{\infty\}, \omega \in \Omega$$

(where the convention $T_\infty(\omega) = \omega$ is used). We can also write X_τ for $X \circ T_\tau$ when $\tau \in \mathsf{Stop}$.

Remark 2.8. Our notion of a state corresponds intuitively to that used in the theory of Markov processes, even though we insist upon assigning a state to each $\omega \in \Omega$. If one pictures $T_t(\omega)$ as trajectory ω stopped at t, then $X_t(\omega)$ is simply the "state" at which ω is stopped. When ω is not necessarily in the image of some T_t, we assign the state abstractly imagining it to be the "value of $\omega(\infty)$".

2.4.5. *Actions on measures and kernels.* For a probability measure $\mu \in \mathsf{Prob}(\Omega)$, and a stopping time $\tau \in \mathsf{Stop}$ we define the **truncated measure** $\mu_{\leq\tau}$ as the push-forward of μ via the truncation map T_τ.

Two analogous operations can be applied to kernels ν from Ω to Ω. We can truncate the second argument, leading to the **truncated kernel** $\nu_{\leq\tau}$, where, for each $\omega \in \Omega$, $\nu_{\leq\tau}(\omega, \cdot)$ is the truncation of the measure $\nu(\omega, \cdot)$, as above. On the other hand, we can define the **restricted kernel** $\nu^{\leq\tau}$ by truncating in the first argument, i.e., by setting

$$\nu^{\leq\tau}(\omega, B) = \nu(T_\tau(\omega), B).$$

That $\nu^{\leq\tau}$ is, indeed, a kernel follows from the fact that a Borel measurable function (like T_τ) between two Polish spaces remains measurable under the pair of universal σ-algebras (see [BS78, Proposition 7.44, p. 172]).

2.5. TC-spaces (truncation-concatenation spaces).

Definition 2.9. A **truncation-concatenation space** (or a **TC-space**) is a truncation space $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{t\in\mathsf{Time}})$ together with a measurable subset $\mathcal{C} \subseteq \Omega \times \mathsf{Time} \times \Omega$ - called the **compatibility set** - and a measurable map $* : \mathcal{C} \to \Omega$ - called the **concatenation operator**, such that the following conditions hold:

(1) for all $\omega, \omega' \in \Omega$ and $s, t \in \mathsf{Time}$ we have

$$(\omega, t, \omega') \in \mathcal{C} \Leftrightarrow (\omega_{\leq t}, t, \omega') \in \mathcal{C} \Leftrightarrow (\omega, t, \omega'_{\leq s}) \in \mathcal{C}. \tag{2.3}$$

(2) if $(\omega, t, \omega') \in \mathcal{C}$, then, for all $s \in \mathsf{Time}$ we have

$$\omega *_t \omega' = \omega_{\leq t} *_t \omega', \text{ as well as} \tag{2.4}$$

$$(\omega *_t \omega')_{\leq s} = \begin{cases} \omega_{\leq s}, & s \leq t \\ \omega *_t \omega_{\leq s-t}, & s > t \end{cases} \tag{2.5}$$

The action of the concatenation operator on the triplet $(\omega, t, \omega') \in \mathcal{C}$ is denoted by $\omega *_t \omega'$ and is usually interpreted as an element of Ω "obtained by following ω until time t, with ω' attached afterwards". The set $\mathcal{C}$ - the domain of $*$ - may encode a compatibility relation necessary for the concatenation to be possible. The

set of all $\omega' \in \Omega$ such that $(\omega, t, \omega') \in \mathcal{C}$ is denoted by $\mathcal{C}_{\omega,t}$, and we say that ω' **is compatible with** ω **at** t if $\omega' \in \mathcal{C}_{\omega,t}$.

In many examples compatibility is established via a state map (as defined in subsection 2.4.4 above):

Definition 2.10. Given a TC-space $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{t\in\mathsf{Time}})$ and a state map X, we say that the concatenation operator $*$

(1) **factors through** X if $\quad X_t(\omega) = X_0(\omega') \quad \Rightarrow (\omega, t, \omega') \in \mathcal{C}$, and
(2) **is a factor of** X if $\quad (\omega, t, \omega') \in \mathcal{C} \quad \Rightarrow X_t(\omega) = X_0(\omega')$.

When needed, we also define $\omega *_\infty \omega' = \omega$, declaring, implicitly, any two elements of Ω compatible at $t = \infty$, so that $\mathcal{C}_{\omega,\infty} = \Omega$. This way, as in the case of the truncation spaces, the time-set Time can be extended to the set of all stopping times by setting:

$$\omega *_\tau \omega' = \omega *_{\tau(\omega)} \omega' \text{ for } \omega' \in \mathcal{C}_{\omega,\tau(\omega)}. \tag{2.6}$$

By Proposition 2.2, part (3), $\tau(\omega_{\leq\tau}) = \tau(\omega)$, and, so, the stopping-time analogue of (2.4) holds in TC spaces:

$$\omega *_\tau \omega' = \omega *_{\tau(\omega)} \omega' = \omega_{\leq\tau(\omega)} *_{\tau(\omega)} \omega' = \omega_{\leq\tau} *_{\tau(\omega_{\leq\tau})} \omega' = \omega_{\leq\tau} *_\tau \omega'.$$

2.6. **Examples of TC-spaces.**

We go through the list of examples of T-spaces from subsection 2.2 and describe how a natural concatenation operator can be introduced.

2.6.1. *Strict concatenation on path spaces D_E and C_E.* We consider the space D_E with the truncation $\omega_{\leq t}(s) = \omega(s \wedge t)$. The **strict concatenation** operation $\bullet$ is given by

$$(\omega \bullet_t \omega')_s = \begin{cases} \omega(s), & s \leq t \\ \omega'(s-t), & s > t, \end{cases} \tag{2.7}$$

for $\omega, \omega' \in D_E$, where ω and ω' are considered t-compatible if and only if $\omega(t) = \omega'(0)$. To check that $\bullet$ is, indeed, a concatenation is straightforward, and we only remark that the joint measurability of $\bullet$ (in all three of its arguments) follows from the observation that, as a function of the inner argument t, it is right-continuous in the Skorokhod topology. When applied on its compatibility set $\mathcal{C}$, the operation $\bullet$ preserves continuity, so it can be used to define a concatenation operator on C_E, as well. Finally, it is straightforward that

$$X(\omega) := \liminf_{t\to\infty} \omega(t)$$

defines an $E = \bar{\mathbb{R}}$-valued state map with the property $X_t(\omega) = \omega(t)$ for $t \in \mathsf{Time}$ and such that the concatenation operator $\bullet$ factors through it.

Remark 2.11. Many subspaces of D_E, in addition to C_E, are closed under the strict concatenation. The reader will easily check that all the spaces in Example 2.7 have this property; it follows that they are TC-spaces themselves.

2.6.2. *Adjusted concatenation on D_E and C_E.* When E admits an additive structure, we can define another concatenation operator on it, namely the **adjusted concatenation** operator $\star$. It is given for $\omega, \omega' \in D_E$ by

$$(\omega \star_t \omega')_s = \begin{cases} \omega(s), & s \le t \\ \omega(t) + \omega'(s-t) - \omega'(0), & s > t, \end{cases} \tag{2.8}$$

with no restrictions on compatibility, i.e., with $\mathcal{C} = \Omega \times \mathsf{Time} \times \Omega$. It is clear that the strict and the adjusted concatenation operators agree on the compatibility set of $\bullet$, and that $*$ can be restricted to C_E without loosing any properties required of a concatenation.

2.6.3. *Spaces of measures.* We define the concatenation operator $*$ on the space $\Omega = \mathcal{M}^{\#}([0,\infty) \times E)$, described in subsection 2.2 as follows. For $\mu, \mu' \in \Omega$, we set

$$(\mu *_t \mu')(A) = \mu\Big(([0,t) \times E) \cap A\Big) + \mu'\Big(((([t,\infty) \times E) \cap A) - t\Big),$$

where $B - t = \{(x, s-t) \,:\, (x,s) \in B\}$, for $B \subseteq [t,\infty) \times E$. No compatibility restrictions are imposed. There should be no difficulty in checking that $*$ satisfies all defining properties of a concatenation. We also note that the same construction applies when $\mathcal{M}^{\#}$ is replaced by $\mathcal{M}^f$.

In the case when $\mathcal{M}^p$ is considered, the above operation does not preserve total mass. This cannot be fixed by restricting compatibility, but can be overcome by defining another concatenation operation as follows:

$$(\mu \,\tilde{*}_t\, \mu')(A) = \mu\Big(([0,t) \times E) \cap A\Big) + \Big(1 - \mu([0,t) \times E)\Big)\mu'\Big(((([t,\infty) \times E) \cap A) - t\Big),$$

2.6.4. *$\mathbb{L}^0_A$ spaces.* When the underlying measure λ is the Lebesgue measure, we usually concatenate $\mathbb{L}^0_A$ functions as follows:

$$(f *_t g)_u = \begin{cases} f_u, & u \le t \\ g_{u-t}, & u > t \end{cases},$$

with no compatibility restriction.

2.7. **Constructions and structure-preserving maps on TC spaces.** Like T-spaces, TC-spaces come with natural subspace and product constructions. Their properties extend those of naked T-spaces in a predictable way, so we skip any further discussion. The following notion of a structure-preserving map on TC spaces will play a major role in Section 3 below.

Definition 2.12. A measurable map $F : \Omega \to \tilde{\Omega}$ between two TC-spaces, with concatenation operators $*$ and $\tilde{*}$ (and compatibility sets $\mathcal{C}$ and $\tilde{\mathcal{C}}$) is called a **TC-morphism** if

(1) F is non-anticipating, and
(2) for all $t \in \mathsf{Time}$, and all $\omega, \omega' \in \Omega$ with $\omega' \in \mathcal{C}_{\omega,t}$ we have $F(\omega') \in \tilde{\mathcal{C}}_{F(\omega),t}$ and

$$F(\omega *_t \omega') = F(\omega) \,\tilde{*}_t\, F(\omega').$$

2.8. **Concatenation of measures in TC-spaces.**

The ability to concatenate elements of Ω extends to probability measures and kernels on Ω. We say that a measure $\mu \in \mathsf{Prob}(\Omega)$ and a kernel $\nu \in \mathsf{Kern}(\Omega)$ on a TC-space are **compatible** at the stopping time τ if

$$\nu_\omega^{\leq\tau}(\mathcal{C}_{\omega,\tau(\omega)}) = 1, \text{ for } \mu\text{-almost all } \omega.$$

When $*$ factors through a state map X, a sufficient condition for compatibility of $\mu \in \mathsf{Prob}(\Omega)$ and $\nu \in \mathsf{Kern}(\Omega)$ at τ is that

$$\nu_\omega^{\leq\tau}\Big(X_0 = X_\tau(\omega)\Big) = 1, \text{ for } \mu\text{-almost all } \omega \text{ with } \tau(\omega) < \infty. \tag{2.9}$$

Using the convention, as above, that $\Omega \times \{\infty\} \times \Omega' \subseteq \mathcal{C}$, we also note that, given a stopping time τ, the set $\mathcal{C}_\tau = \{(\omega, \omega') : (\omega, \tau(\omega), \omega') \in \mathcal{C}\}$ is a pullback of the Borel set $\mathcal{C}$ via the measurable map $(\omega, \omega') \mapsto (\omega, \tau(\omega), \omega')$, and, therefore, itself measurable.

For $\mu \in \mathsf{Prob}(\Omega)$ and a τ-compatible kernel $\nu \in \mathsf{Kern}(\Omega)$ let $\mu \otimes \nu^{\leq\tau} \in \mathsf{Prob}(\Omega \times \Omega)$ denote the product of μ and the τ-restriction of ν. The **concatenation** $\mu *_\tau \nu$ is then defined as the push-forward of this product via the measurable map $C_\tau \ni (\omega, \omega') \mapsto \omega *_{\tau(\omega)} \omega'$. We note that the compatibility relation introduced above implies that $\mu \otimes \nu^{\leq\tau}(C_\tau) = 1$, so that $\mu *_\tau \nu$ is, indeed, a probability measure. Moreover, we have

$$\int G(\omega)\,(\mu *_\tau \nu)(d\omega) = \int G(\omega *_\tau \omega')\,(\mu \otimes \nu^{\leq\tau})(d\omega, d\omega') = \iint G(\omega *_\tau \omega')\,\nu_\omega^{\leq\tau}(d\omega')\,\mu(d\omega),$$

for any sufficiently integrable random variable G on Ω. The compatibility condition (2.4) implies further that

$$\begin{aligned}\int G\,d(\mu *_\tau \nu) &= \iint G(\omega_{\leq\tau} *_\tau \omega')\,\nu_\omega^{\leq\tau}(d\omega')\mu(d\omega) \\ &= \iint G(\tilde{\omega} *_\tau \omega')\,\nu_\omega^{\leq\tau}(d\omega')\mu_{\leq\tau}(d\tilde{\omega}),\end{aligned} \tag{2.10}$$

where $\mu_{\leq\tau}$ is the push forward of μ via T_τ.

2.8.1. *Tail maps.* Tail maps on TC-spaces will play an important role in the dynamic programming principle and will model payoffs associated to controlled processes.

Definition 2.13. A measurable map G from a TC-space to a measurable space S is called a **tail map** if $G(\omega *_t \omega') = G(\omega')$ for all $t \in \mathsf{Time}$, all $\omega \in \Omega$ and all $\omega' \in \mathcal{C}_{\omega,t}$. When $S = \mathbb{R}$ ($S = \bar{\mathbb{R}}$), a tail map is called a **tail random variable** (**extended tail random variable**).

The tail property of random variables extends readily to stopping times in the following form:

$$G(\omega *_\tau \omega') = \begin{cases} G(\omega'), & \tau(\omega) < \infty \\ G(\omega), & \tau(\omega) = \infty, \end{cases}$$

as long as ω' is compatible with ω at τ. Combining this expression with (2.10) we obtain the following equality, valid for each stopping time τ, probability $\mu \in \mathsf{Prob}(\Omega)$, a τ-compatible kernel $\nu \in \mathsf{Kern}(\Omega)$, and a sufficiently integrable tail random variable G:

$$\int G\, d(\mu *_\tau \nu) = \int \tilde{G}(\omega_{\leq \tau})\, \mu(d\omega), \tag{2.11}$$

where

$$\tilde{G}(\omega) = G(\omega)\mathbf{1}_{\{\tau(\omega)=\infty\}} + \int G(\omega')\, \nu_\omega(d\omega')\mathbf{1}_{\{\tau(\omega)<\infty\}}.$$

2.9. **Control correspondences.**

A map $f : A \to 2^B$, where 2^B denotes the power-set of B is called a **correspondence** from A to B, and is also denoted by $f : A \twoheadrightarrow B$. Its **graph** $\Gamma(f) \subseteq A \times B$ is given by $\Gamma(f) = \{(a, b) \,:\, a \in A, b \in f(a)\}$, and its **image** by $\mathrm{Im}(f) = \cup_{a\in A} f(a)$. A correspondence is said to be **non-empty-valued** if $f(a) \neq \emptyset$ for all $a \in A$.

Definition 2.14. A non-empty-valued correspondence $\mathcal{P} : \Omega \twoheadrightarrow \mathsf{Prob}(\Omega)$, on a measurable space Ω is called a **control correspondence**.

Given a control correspondence $\mathcal{P}$, a universally measurable random variable G is said to be $\mathcal{P}$**-upper semi-integrable**, denoted by $G \in \mathcal{L}^{1-0}(\mathcal{P})$, if $G^+ \in \mathcal{L}^1(\mu)$ for each $\mu \in \mathrm{Im}\,\mathcal{P}$. To each control correspondence $\mathcal{P}$ and each $G \in \mathcal{L}^{1-0}(\mathcal{P})$ we associate the **value function** $v : \Omega \to [-\infty, \infty]$, given by

$$v(\omega) = \sup_{\mu\in\mathcal{P}(\omega)} \int G\, d\mu. \tag{2.12}$$

2.10. **Three key properties.**

There are three key properties that control correspondences must satisfy in order for our main results to apply. These properties appear in [Žit14] in a similar terminologial setting, but have been considered and understood in the literature in diffferent forms long before that (see [ET13a, NvH13] for two recent formulations). We recall that a **universally measurable** $\mathcal{P}$**-selector** (or, simply, a $\mathcal{P}$**-selector**) is a (universally measurable) kernel form Ω to $\mathsf{Prob}(\Omega)$ with the property that $\nu(\omega) \in \mathcal{P}(\omega)$, for each ω; the family of all $\mathcal{P}$-selectors is denoted by $\mathcal{S}(\mathcal{P})$.

Definition 2.15. A control correspondence $\mathcal{P}$ on standard Borel space Ω is called

(1) **analytic** if its graph $\Gamma(\mathcal{P})$ is an analytic subset of the (standard Borel) space $\Omega \times \mathsf{Prob}(\Omega)$.

A control correspondence $\mathcal{P}$ defined on a TC space $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{t\in\mathsf{Time}})$ is said to be

(2) **concatenable** if for each $\omega \in \Omega$, $\mu \in \mathcal{P}(\omega)$, $\nu \in \mathcal{S}(\mathcal{P})$, and each stopping time τ, ν is τ-compatible with μ and

$$\mu *_\tau \nu \in \mathcal{P}(\omega).$$

(3) **disintegrable** if for each $\omega \in \Omega$, $\mu \in \mathcal{P}(\omega)$ and a stopping time τ there exists $\nu \in \mathcal{S}(\mathcal{P})$ such that ν is μ-compatible at τ and

$$\mu = \mu *_\tau \nu.$$

Remark 2.16. It follows directly from the definitions of analyticity, concatenability and disintegrability that the following, useful, implications hold for any sequence of control correspondences $\{\mathcal{P}_n\}_{n\in\mathbb{N}}$ on the same Borel space Ω. Let $\cap_n \mathcal{P}_n$ and $\cup_n \mathcal{P}_n$ be the intersection and the union, defined pointwise, on $\{\mathcal{P}_n\}_{n\in\mathbb{N}}$.

(1) If each $\mathcal{P}_n$ is analytic, then so are $\cup_n \mathcal{P}_n$ and $\cap_n \mathcal{P}_n$.
(2) If each $\mathcal{P}_n$ is concatenable, then so is $\cap_n \mathcal{P}_n$.
(3) If each $\mathcal{P}_n$ is disintegrable, then so is $\cup_n \mathcal{P}_n$.

We state for completeness the following result which will be used in the sequel, and the proof of which follows almost verbatim the argument in [Žit14, Theorem 2.4, part 1., p. 1605], which, in turn, is a reformulation of the standard argument available, for example, in [BS78]. We remind the reader the convention $+\infty - \varepsilon = 1/\varepsilon$, for $\varepsilon > 0$.

Proposition 2.17 (Universal measurability of value functions)**.** *Suppose that Ω is a standard Borel space, $\mathcal{P}$ an analytic control correspondence, $G \in \mathcal{L}^{1-0}(\mathcal{P})$ and that v is the associated value function, given by* (2.12)*. Then v is universally measurable and for each $\varepsilon > 0$ there exists a (universally measurable) selector $\nu^\varepsilon \in \mathcal{S}(\mathcal{P})$ such that*

$$v(\omega) - \varepsilon \leq \int G \, d\nu^\varepsilon_\omega, \quad \textit{for all } \omega \in \Omega.$$

2.11. **An abstract version of the dynamic programming principle.**

We are ready to state the most abstract version of the DPP that holds in our setting. A more directly applicable - and more familiar-looking - version, based on the notion of a state map will be given below. The ideas in the proof are entirely standard. In fact, our setting is constructed as the most flexible one where this proof can be applied. We provide the details for the reader's convenience.

Theorem 2.18 (DPP)**.** *Let $\mathcal{P}$ be an analytic control correspondence on a TC space Ω, $G \in \mathcal{L}^{1-0}(\mathcal{P})$ a tail random variable, and v the associated value function, given by* (2.12)*. Then,*

(1) If $\mathcal{P}$ is concatenable, then for each $\omega \in \Omega$ and each stopping time τ we have

$$v(\omega) \geq \sup_{\mu\in\mathcal{P}(\omega)} \int \left(v \circ T_\tau \mathbf{1}_{\{\tau<\infty\}} + G\mathbf{1}_{\{\tau=\infty\}} \right) d\mu \tag{2.13}$$

(2) If $\mathcal{P}$ is disintegrable, then for each $\omega \in \Omega$ and each stopping time τ we have

$$v(\omega) \leq \sup_{\mu\in\mathcal{P}(\omega)} \int \left(v \circ T_\tau \mathbf{1}_{\{\tau<\infty\}} + G\mathbf{1}_{\{\tau=\infty\}} \right) d\mu \tag{2.14}$$

Proof. Suppose, first, that $\mathcal{P}$ is concatenable and pick $\omega \in \Omega$, $\mu \in \mathcal{P}(\omega)$ and a stopping time τ. Given $\varepsilon > 0$, Proposition 2.17 guarantees the existence of an ε-optimizing selector ν^ε, i.e., such that $v^\varepsilon(\omega) := \int G\, d\nu^\varepsilon_\omega \geq v(\omega) - \varepsilon$, for each $\omega \in \Omega$. We construct the measure μ' by concatenating μ and ν^ε at τ. The assumption of concatenability implies that they are compatible and that $\mu' \in \mathcal{P}(\omega)$. Therefore,

$$\begin{aligned} v(\omega) \geq \int G\, d\mu' &= \int G\, d(\mu *_\tau \nu^\varepsilon) = \iint G(\omega *_\tau \omega')\, (\nu^\varepsilon)^{\leq\tau}_\omega(d\omega')\, \mu(d\omega) \\ &= \iint \Big(G(\omega)\mathbf{1}_{\{\tau(\omega)=\infty\}} + G(\omega')\mathbf{1}_{\{\tau(\omega)<\infty\}} \Big)\, (\nu^\varepsilon)^{\leq\tau}_\omega(d\omega')\, \mu(d\omega) \\ &\geq \int \Big(G(\omega)\mathbf{1}_{\{\tau(\omega)=\infty\}} + (v(\omega_{\leq\tau}) - \varepsilon)\mathbf{1}_{\{\tau(\omega)<\infty\}} \Big)\, \mu(d\omega), \end{aligned}$$

which implies (2.13).

In the disintegrable case, we pick $\varepsilon > 0$, $\omega \in \Omega$, $\tau \in \mathsf{Stop}$ and choose $\mu^\varepsilon \in \mathcal{P}(\omega)$ such that $v(\omega) - \varepsilon \leq \int G\, d\mu^\varepsilon$. By disintegrability, we can write $\mu^\varepsilon = \mu^\varepsilon *_\tau \nu$ for some $\nu \in \mathcal{S}(\mathcal{P})$, and so

$$\begin{aligned} v(\omega) - \varepsilon \leq \int G\, d(\mu^\varepsilon *_\tau \nu) &= \int \Big(G(\omega)\mathbf{1}_{\{\tau=\infty\}} + \mathbf{1}_{\{\tau<\infty\}} \Big(\int G(\omega')\nu^{\leq\tau}_\omega(d\omega') \Big) \Big)\, \mu(d\omega) \\ &\leq \int \Big(G(\omega)\mathbf{1}_{\{\tau=\infty\}} + v(\omega_{\leq\tau})\mathbf{1}_{\{\tau<\infty\}} \Big)\, \mu(d\omega). \end{aligned}$$

□

2.11.1. *State maps and factoring.* We remind the reader that, as defined in subsection 2.4.4, a state map $X : \Omega \to E$ is simply a measurable map from a T-space to a Polish space E, and that X_τ is a shortcut for $X \circ T_\tau$, for $\tau \in \mathsf{Stop}$. Just like (concatenation) compatibility may factor through X, so can a control correspondence:

Definition 2.19. A control correspondence $\mathcal{P}$ on Ω is said to **factor through** a state map X if there exists a correspondence $\bar{\mathcal{P}} : E \twoheadrightarrow \mathsf{Prob}(\Omega)$ such that $\mathcal{P}(\omega) = \bar{\mathcal{P}}(X(\omega)) \subseteq \mathsf{Prob}(\Omega)$, i.e., the following diagram commutes:

$$\begin{array}{ccc} \Omega & \xrightarrow{\ X\ } & E \\ \Big\downarrow{\scriptstyle \mathcal{P}} & \swarrow{\scriptstyle \bar{\mathcal{P}}} & \\ \mathsf{Prob}(\Omega) & & \end{array} \tag{2.15}$$

A very simple, but important, consequence of the existence of a state map through which the control correspondence $\mathcal{P}$ factors is that in that case, v factors through it as well. Indeed, the function $\bar{v} : E \to [-\infty, \infty]$, given by $\bar{v}(x) = \sup_{\mu \in \bar{\mathcal{P}}(x)} \int G\, d\mu$, then has the property that $\bar{v}(X(\omega)) = v(\omega)$ and, under the conditions of Theorem 2.18, satisfies

$$\bar{v}(x) \leq (\geq) \sup_{\mu \in \mathcal{P}(x)} \int \Big(\bar{v}(X_\tau)\mathbf{1}_{\{\tau<\infty\}} + G\mathbf{1}_{\{\tau=\infty\}} \Big)\, d\mu$$

for all $x \in \operatorname{Im} X$, and all stopping times $\tau \in \mathsf{Stop}$.

3. MARTINGALE-GENERATED CONTROL CORRESPONDENCES

Our next task is so take the abstraction level down a notch and study a class of control correspondences defined via a family of martingale conditions. These correspondences generalize the standard martingale formulation in the theory of stochastic optimal control and are defined via a family of structure-preserving maps into the model space space $D^0_{\mathbb{R}}$ of $\mathbb{R}$-valued càdlàg paths $x : \mathsf{Time} \to \mathbb{R}$ with $x(0) = 0$.

3.1. **Canonical local martingale measures.**

With the T-space structure of $D_{\mathbb{R}}$ described in subsection 2.2, each non-anticipating map F from a T-space $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{t\in\mathsf{Time}})$ into $D_{\mathbb{R}}$ induces a sequence $\{F^n\}$ of non-anticipating maps

$$F^n_t = F_{\tau^F_n \wedge t} \text{ where } \tau^F_n(\omega) = \inf\{t \geq 0 : |F_t(\omega)| \geq n\} \wedge n. \tag{3.1}$$

When the choice of F is evident from context, we may drop the superscript and write $\tau_n = \tau^F_n$.

Definition 3.1. A probability measure $\mu \in \mathsf{Prob}(\Omega)$ is said to be a **canonical local-martingale probability for** F if the stochastic process $\{F^n_t(\cdot)\}_{t\in\mathsf{Time}}$ is a martingale under $(\mu, \mathbb{F})$ for each $n \in \mathbb{N}$. The set of all canonical local martingale probabilities for F is denoted by $\mathcal{M}^{F,loc}$.

Remark 3.2. The notion of a *canonical* local martingale differs from the standard notion of a local martingale in that it requires that the reducing sequence takes a particular form, namely that of the sequence of space-time exit times. This requirement is nontrivial, as it is known that there are local martingales that cannot be reduced by this particular sequence (see [Str77, Lemme 2.1., p. 57]). On the other hand, this notion suffices for many applications; indeed for continuous processes (or processes with jumps bounded from below) the notions of a canonical local martingale and that of a local martingale coincide.

With the notion of a canonical local martingale probability under our belt, we can define a large class of control correspondences. Housed on T-spaces, they need two ingredients to be specified: 1) a family of $\mathcal{D}$ of non-anticipating maps from $\Omega \to D_{\mathbb{R}}$, and 2) a state map X from Ω to a Polish space E. Once these are specified, for $x \in E$ we define

$$\bar{\mathcal{P}}(x) = \bigcap_{F\in\mathcal{D}} \mathcal{M}^{F,loc} \cap \Big\{\mu \in \mathsf{Prob}(\Omega) \,:\, X_0 = x,\ \mu\text{-a.s.}\Big\}, \tag{3.2}$$

where, as usual, X_0 is the shortcut for $X \circ T_0$. The $(\mathcal{D}, X)$-**generated control correspondence** $\mathcal{P} = \mathcal{P}(\mathcal{D}, X) : \Omega \twoheadrightarrow \mathsf{Prob}(\Omega)$ is then defined by

$$\mathcal{P}(\omega) = \bar{\mathcal{P}}(X(\omega)) \text{ for } \omega \in \Omega,$$

so that it naturally factors through X.

3.2. **Sufficient conditions for analyticity.**

The ubiquitous Polish-space structure woven into all the ingredients of our setup makes it possible to give widely met sufficient conditions on the family $\mathcal{D}$ such that the resulting $(\mathcal{D}, X)$-correspondence becomes analytic. The countability condition we impose on $\mathcal{D}$ is not the weakest possible, but since it holds in most relevant examples, we only comment on some possible routes towards establishing weaker versions in Remark 3.5 below.

Proposition 3.3. *Let $\mathcal{D}$ be a countable family of non-anticipating maps from a T-space Ω to $D_{\mathbb{R}}$ and let $X : \Omega \to E$ be a state map. Then the $(\mathcal{D}, X)$-generated control correspondence $\mathcal{P}$ is analytic.*

The proof is based on a modification of [Žit14, Lemma 3.6, p. 1611], where

$$\mathsf{QStop} = \left\{ q\mathbf{1}_A + r\mathbf{1}_{A^c} \,:\, q \leq r \in \mathsf{QTime}, A \in \Pi_q \right\}$$

with QTime denoting a countable dense set in Time, and $\{\Pi_q\}_{q\in\mathsf{QTime}}$ a collection of countable π-systems such that $\sigma(\Pi_q) = \mathcal{F}_q$ for all $q \in \mathsf{QTime}$. The exact choice of QTime or $\{\Pi_q\}_{q\in\mathsf{QTime}}$ is unimportant, as long as it is fixed throughout.

Lemma 3.4. *For each non-anticipating map F, we have*

$$\mathcal{M}^{F,loc} = \bigcap \left\{ \mu \in \mathsf{Prob}(\Omega) \,:\, F^n_q, F^n_r \in \mathbb{L}^1(\mu) \text{ and } \mathbb{E}^\mu[F^n_r \mathbf{1}_A] = \mathbb{E}^\mu[F^n_q \mathbf{1}_A] \right\} \tag{3.3}$$

where the intersection is taken over all $n \in \mathbb{N}$, $q < r \in \mathsf{QTime}$ and $A \in \Pi_q$.

Proof. The inclusion $\mathcal{M}^{F,loc} \subseteq \ldots$ is straightforward. Conversely, let $\mu \in \mathsf{Prob}(\Omega)$ be an element of the right-hand side of (3.3). We first show that $\mu \in \mathcal{M}^{F^n}_{\mathsf{QTime}}$, where $\mathcal{M}^{F^n}_{\mathsf{QTime}}$ denotes the set of all $\mu \in \mathsf{Prob}(\Omega)$ with the property that $\{F^n_t\}_{t\in\mathsf{QTime}}$ is a μ-martingale with respect to $\{\mathcal{F}_t\}_{t\in\mathsf{QTime}}$. That is an immediate consequence of the equalities of expectations under μ on the right-hand-side of (3.3). Considered over all $A \in \Pi_q$, with $q < r \in \mathsf{QTime}$, they amount to $\mathbb{E}^\mu[F^n_r|\mathcal{F}_q] = F^n_q$, a.s., by π-λ-theorem.

It remains to argue that F^n is a μ-martingale on entire Time. Assuming, without loss of generality, that $\mathsf{Time} = [0, \infty)$, we start by picking $s \in \mathsf{Time}\setminus\mathsf{QTime}$ and $r \in \mathsf{QTime}$ with $r > s$. The backward martingale convergence theorem implies that

$$\mathbb{E}^\mu[F^n_r|\mathcal{F}_{s+}] = F^n_s, \ \mu\text{-a.s.}$$

Since F^n is non-anticipating, F^n_s is $\mathcal{F}_s$-measurable and we may replace $\mathcal{F}_{s+}$ by $\mathcal{F}_s$ in the equality above. Finally, for $t \in \mathsf{Time}$ with $t > s$, we approximate F^n_t by a sequence $\{F^n_{r_m}\}_{m\in\mathbb{N}}$ with $r_m \searrow t$ and $r_m \in \mathsf{QTime}$, to conclude that F^n is, indeed, a martingale under μ. □

Proof of Proposition 3.3. For each $r \in \mathsf{Time}$, the coordinate maps are Borel measurable on $D_{\mathbb{R}}$ and, so, $\mu \mapsto \mathbb{E}^\mu[F_r\mathbf{1}_A]$ is Borel on Ω. It is easy to see that the family of probability measures under which a given real-valued Borel map is integrable is also a Borel set, so it follows that $\mathcal{M}^{F,loc}$ is Borel for each F. The

countability of $\mathcal{D}$ guarantees that $\cap_{F\in\mathcal{D}}\mathcal{M}^{F,loc}$, as well. Finally, the graph of $\mathcal{P}$ is analytic (in fact Borel) as it is given as an intersection of Borel sets

$$\Gamma(\mathcal{P}) = \Big\{(\omega,\mu) \,:\, \mu\Big(X_0 = X_0(\omega)\Big) = 1\Big\} \cap \left(\Omega \times \bigcap_{F\in\mathcal{D}} \mathcal{M}^F\right). \qquad \square$$

Remark 3.5. When $\mathcal{D}$ is not countable, the set $\cap_{F\in\mathcal{D}}\mathcal{M}^{F,loc}$ is not necessarily Borel measurable (or even analytic) in general. The situation is somewhat more pleasant when $\mathcal{D}$ admits a structure of a Borel space with the property that the maps

$$\mathcal{D} \ni F \mapsto \mathbb{E}^{\mu}[F_r], r \in \mathsf{Time},$$

are measurable for each probability measure $\mu \in \mathsf{Prob}(\Omega)$. In that case, the intersection $\cap_{F\in\mathcal{D}}\mathcal{M}^{F,loc}$ can be represented as a *co-projection*

$$\{\mu \in \mathsf{Prob}(\Omega) \,:\, \forall\, F \in \mathcal{D},\ (F,\mu) \in \mathcal{M}\}$$

of the Borel set $\mathcal{M} = \{(F,\mu) \in \mathcal{D} \times \mathsf{Prob}(\Omega) \,:\, \mu \in \mathcal{M}^{F,loc}\}$. Unlike projections, the images of co-projections are co-analytic, but not necessarily analytic sets. Not everything is lost, however, as we usually know a great deal more about the set $\mathcal{M}$, other than the fact that it is a Borel set. Indeed, the countable case of Proposition (3.3) corresponds to the measurable-selection theorem of Lusin for sets with countable sections (see [Sri98, Theorem 5.7.2, p. 205]). On the other side of the spectrum are measurable-selection theorems with large sections (see Section 5.8 in [Sri98]), which can be used for certain uncountable $\mathcal{D}$.

3.3. **Sufficient conditions for concatenability.** Having discussed analyticity, we turn to the second major assumption of our abstract DPP theorem, namely concatenability. It is not hard to see that without additional requirements on $\mathcal{D}$, no $(\mathcal{D}, X)$-generated control correspondence should be expected to be concatenable. A natural requirement, as we will see below, is that the maps F be TC-morphisms, introduced in Definition 2.12 above. Moreover, the target space for these TC-morphisms will be $D^0_{\mathbb{R}}$ - a model space for (the laws of) local martingales. We remind the reader (see section 2.6 above) that $D_{\mathbb{R}}$ comes with two different natural concatenations, namely the strict one ($\bullet$) and the adjusted one ($\star$). We will only work with the adjusted one in this section, but, in order to avoid any confusion, we will write $(D_{\mathbb{R}}, \star)$ and $(D^0_{\mathbb{R}}, \star)$ throughout.

Definition 3.6. A map $F : \Omega \to D_{\mathbb{R}}$ is said to be **canonically locally bounded** if there exists a sequence $\{M_n\}_{n\in\mathbb{N}}$ of positive constants so that

$$|F^n(\omega)_t| \le M_n \text{ for all } \omega \in \Omega, t \in \mathsf{Time}. \tag{3.4}$$

A simple sufficient condition for canonical local boundedness is that the jumps of F (when seen as a stochastic process on Ω) are uniformly bounded.

Proposition 3.7. *Let $\mathcal{D}$ be a family of canonically locally bounded TC-morphisms into $(D^0_{\mathbb{R}}, \star)$, and let X be a state map. Then the $(\mathcal{D}, X)$-generated control correspondence $\mathcal{P}$ is closed under concatenation.*

The proof is based on the several lemmas. We omit the straightforward proof of the first one.

Lemma 3.8. *Suppose that F is a TC-morphism into $(D_{\mathbb{R}}, \star)$. For all stopping times κ we have*

$$F_{\kappa+t}(\omega *_\kappa \omega') - F_{\kappa+s}(\omega *_\kappa \omega') = F_t(\omega') - F_s(\omega')$$

for all $\omega \in \Omega$ with $\kappa(\omega) < \infty$, $\omega' \in \mathcal{C}_{\omega,\kappa(\omega)}$ and all $s, t \in \mathsf{Time}$.

Our second lemma gives a convenient characterization of canonical local martingales. We use Stop, as in the case of T-spaces, to denote the set of all Time-valued (raw) stopping times. We also write $Y^n = Y^{\tau_n}$, where $\tau_n = \inf\{t \geq 0 \,:\, |Y_t| \geq n\} \wedge n$, and note that all sampled values of Y in the statement are well-defined thanks to the fact that each Y^n is constant after $t = n$.

Lemma 3.9. *Let $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{t\in\mathsf{Time}}, \mathbb{P})$ be a filtered probability space, $\{Y_t\}_{t\in\mathsf{Time}}$ a càdlàg and adapted process, and κ a stopping time with $Y^n_\kappa \in \mathbb{L}^1$ for each $n \in \mathbb{N}$. Then, the following two statements are equivalent*

(1) Y is a canonical local martingale.
(2) $G \in \mathbb{L}^1$ and $\mathbb{E}[G] = 0$ for all

$$G \in \bigcup_{n\in\mathbb{N}} \mathcal{X}_n^{\leq\kappa}(Y) \cup \mathcal{X}_n^{\geq\kappa}(Y),$$

where the countable sets $\mathcal{X}_n^{\leq\kappa}$ and $\mathcal{X}_n^{\geq\kappa}$ are given by

$$\mathcal{X}_n^{\leq\kappa}(Y) = \left\{Y^n_{\tau\wedge\kappa} - Y^n_\kappa \,:\, \tau \in \mathsf{QStop}\right\},$$
$$\mathcal{X}_n^{\geq\kappa}(Y) = \left\{Y^n_{\tau\vee\kappa} - Y^n_\kappa \,:\, \tau \in \mathsf{QStop}\right\}.$$

Proof. (1) $\Rightarrow$ (2) Assuming that Y is a canonical local martingale, each Y^n is martingale constant after $t = n$, and therefore a uniformly-integrable martingale. Stopping times in QStop are bounded, so, by the optional sampling theorem, (2) holds.

(2) $\Rightarrow$ (1) Suppose that (2) holds and that $n \in \mathbb{N}$ is fixed. We take the advantage of the fact that Y is càdlàg to conclude (as in the proof of Lemma 3.4) that it suffices to show that Y^n is a martingale on QTime. For that, in turn, we choose $\tau \in \mathsf{QStop}$, so that $\tau = p\mathbf{1}_A + q\mathbf{1}_{A^c}$ for some $p \leq q \in \mathsf{QTime}$ and $A \in \Pi_p$ and note that

$$Y^n_\tau - Y^n_\kappa = \left(Y^n_{\tau\wedge\kappa} - Y^n_\kappa\right) + \left(Y^n_{\tau\vee\kappa} - Y^n_\kappa\right).$$

Since $Y^n_{\tau\wedge\kappa} - Y^n_\kappa \in \mathcal{X}^{\leq\kappa}$ $Y^n_{\tau\vee\kappa} - Y^n_\kappa \in \mathcal{X}^{\geq\kappa}$ and $Y^n_\kappa \in \mathbb{L}^1$, we conclude that $Y^n_\tau \in \mathbb{L}^1$ and that $\mathbb{E}[Y^n_\tau] = \mathbb{E}[Y^n_\kappa]$. It follows that the value of $\mathbb{E}[Y^n_\tau]$ does not depend on the choice of τ, making Y^n into a martingale. □

Lemma 3.10. *Let Ω be a TC-space and $\kappa, \tau \in \mathsf{Stop}$ such that $\kappa \leq \tau$. For $\omega \in \Omega$ we define τ'_ω by*

$$\tau'_\omega(\omega') = \begin{cases} \tau(\omega *_\kappa \omega') - \kappa(\omega), & \kappa(\omega) < \infty \text{ and } \omega' \in \mathcal{C}_{\omega,\kappa(\omega)} \\ +\infty, & \text{otherwise,} \end{cases}$$

Then the map $(\omega, \omega') \mapsto \tau'_\omega(\omega')$ *is jointly measurable,* $\tau'_\omega \in \mathsf{Stop}$ *for any fixed* $\omega \in \Omega$*, and* $\tau(\omega *_\kappa \omega') = \kappa(\omega) + \tau'_\omega(\omega')$.

Proof. By construction, we clearly have $\tau(\omega *_\kappa \omega') = \kappa(\omega) + \tau'_\omega(\omega')$. With the convention that $\tau(\omega *_\kappa \omega') - \kappa(\omega) = \infty$ when $\kappa(\omega) = \infty$, we note that τ' can be expressed as:

$$\tau'_\omega(\omega') = (+\infty)\mathbf{1}_{\mathcal{C}^c}(\omega, \kappa(\omega), \omega') + (\tau(\omega *_\kappa \omega') - \kappa(\omega))\mathbf{1}_{\mathcal{C}}(\omega, \kappa(\omega), \omega')$$

and is hence jointly measurable. It remains to argue that τ'_ω is a stopping time. We fix $\omega \in \Omega$ with $k = \kappa(\omega) < \infty$, and for $s \in \mathsf{Time}$ define

$$A = \{\omega' \in \Omega \,:\, \tau'(\omega') \le s\} = \{\omega' \in \mathcal{C}_{\omega,k} \,:\, \tau(\omega *_k \omega') \le s + k\}.$$

By Proposition 2.2, part (1), it will suffice to show that $T_s^{-1}(A) = A$, i.e., for $\omega' \in \Omega$ we have $(a) \Leftrightarrow (b)$, where

(a) $\omega' \in \mathcal{C}_{\omega,k}$ and $\tau(\omega *_k \omega') \le s + k$, and
(b) $(\omega')_{\le s} \in \mathcal{C}_{\omega,k}$ and $\tau(\omega *_k (\omega'_{\le s})) \le s + k$.

The first, compatibility-related, parts of statements of (a) and (b) are equivalent to each other by the assumptions in (2.3) of Definition 2.9. To deal with the inequalities involving τ we use Proposition 2.2, part (2), as well as the assumption 2.5 of Definition 2.9 to conclude that

$$\tau\Big(\omega *_k (\omega'_{\le s})\Big) \le s + k \Leftrightarrow \tau\Big((\omega *_k (\omega'_{\le s}))_{\le s+k}\Big) \le s + k \Leftrightarrow \tau\Big((\omega *_k \omega')_{\le s+k}\Big) \le s + k$$
$$\Leftrightarrow \tau\Big(\omega *_k \omega'\Big) \le s + k. \qquad \square$$

Proof of Proposition 3.7. Let $\mathcal{P}$ be the $(\mathcal{D}, X)$-generated control correspondence as in the statement, and let $\omega_0 \in \Omega$, $\mu \in \mathcal{P}(\omega_0)$, a kernel $\nu \in \mathcal{S}(\mathcal{P})$ and a stopping time κ be given.

First, we argue that ν is κ-compatible with μ. By the definition of $\mathcal{P}$, we have $\nu_\omega(X_0 = X(\omega)) = 1$ for each $\omega \in \Omega$. After a composition with T_κ, we get $\nu_\omega^{\le\kappa}(X_0 = X_\kappa(\omega)) = 1$ for each $\omega \in \Omega$, which implies compatibility, according to the criterion of (2.9).

Next, we show that $\mu' = \mu *_\kappa \nu \in \mathcal{P}(\omega_0)$. Part (2) of Definition 2.9 makes it clear that for $x = X_0(\omega_0)$ we have $\mu'(X_0 = x) = 1$. Therefore, we need to argue that $\mu' \in \mathcal{M}^{F,loc}$, for each $F \in \mathcal{D}$. By Lemma 3.9, this is equivalent to checking $\int G\, d(\mu *_\kappa \nu) = 0$ for all $G \in \cup_{n\in\mathbb{N}} \mathcal{X}_n^{\le\kappa}(F) \cup \mathcal{X}_n^{\ge\kappa}(F)$. We fix $n \in \mathbb{N}$ and treat the two cases separately:

1. $G \in \mathcal{X}_n^{\le\kappa}(F)$: In this case there exists $\tau \in \mathsf{QStop}$, such that $G(\omega) = F^n_{(\tau\wedge\kappa)(\omega)}(\omega) - F^n_{\kappa(\omega)}(\omega)$. By Definition 2.9, part (2), we have $(\tau \wedge \kappa)(\omega *_\kappa \omega') = (\tau \wedge \kappa)(\omega)$ and $\kappa(\omega *_\kappa \omega') = \kappa(\omega)$, so that, by the non-anticipativity of F^n (which follows from the non-anticipativity of F), we have

$$G(\omega *_\kappa \omega') = F^n_{(\tau\wedge\kappa)(\omega)}(\omega *_\kappa \omega') - F^n_{\kappa(\omega)}(\omega *_\kappa \omega') = F^n_{(\tau\wedge\kappa)(\omega)}(\omega) - F^n_{\kappa(\omega)}(\omega) = G(\omega).$$

Since G is bounded (since so is F^n) we have

$$\int G\, d\mu' = \iint G(\omega *_\kappa \omega')\, \nu_\omega^{\leq \kappa}(d\omega')\, \mu(d\omega) = \int G(\omega)\, \mu(d\omega) = 0,$$

where the last equality follows from the fact that $\mu \in \mathcal{M}^{F,loc}$.

2. $G \in \mathcal{X}_n^{\geq \kappa}(F)$: Let $\tau \in \mathsf{QStop}$ be such that $G = F^n_{\tau\vee\kappa} - F^n_\kappa$. Then

$$\begin{aligned}\int F^n_{\tau\vee\kappa}(\omega) - F^n_\kappa(\omega)\, \mu'(d\omega) &= \int \mathbf{1}_{\{\tau_n > \kappa\}}(\omega)(F^n_{\tau\vee\kappa}(\omega) - F^n_\kappa(\omega))\, \mu'(d\omega) \\ &= \int \mathbf{1}_{\{\tau_n > \kappa\}}(\omega)(F_{(\tau\wedge\tau_n)\vee\kappa}(\omega) - F_\kappa(\omega))\, \mu'(d\omega)\end{aligned}$$

Note that $(\tau \wedge \tau_n) \vee \kappa \geq \kappa$, and let τ' be as in Lemma 3.10 (applied to $(\tau \wedge \tau_n) \vee \kappa$). Also note that by Proposition 2.2, $\{\tau_n > \kappa\} \in \mathcal{F}_\kappa = \sigma(T_\kappa)$. Therefore $\mathbf{1}_{\{\tau_n>\kappa\}}$ is $\sigma(T_\kappa)$-measurable and so $\mathbf{1}_{\{\tau_n>\kappa\}}(\omega *_\kappa \omega') = \mathbf{1}_{\{\tau_n>\kappa\}}(\omega_{\leq\kappa}) = \mathbf{1}_{\{\tau_n>\kappa\}}(\omega)$. Continuing with the equalities from above, we have

$$\begin{aligned}&\int F^n_{\tau\vee\kappa}(\omega) - F^n_\kappa(\omega)\, \mu'(d\omega) = \\ &= \iint \mathbf{1}_{\{\tau_n>\kappa\}}(\omega)(F_{\kappa(\omega)+\tau'_\omega(\omega')}(\omega *_\kappa \omega') - F_\kappa(\omega *_\kappa \omega'))\, \nu_\omega(d\omega')\mu(d\omega) \\ &= \iint \mathbf{1}_{\{\tau_n>\kappa\}}(\omega)(F_{\tau'_\omega}(\omega') - F_0(\omega'))\, \nu_\omega(d\omega')\mu(d\omega) = \iint \mathbf{1}_{\{\tau_n>\kappa\}}(\omega)F_{\tau'_\omega}(\omega')\, \nu_\omega(d\omega')\mu(d\omega).\end{aligned}$$

where the last equality used the TC-morphism assumption together with Lemma 3.8. With M_n given by (3.4), $|F|$ is bounded on $[0, \tau'_\omega]$ by $2M_{2n}$ when $\omega \in \{\kappa < \tau_n\}$, By the canonical local martingale property, we have $\int F_{\tau'_\omega}(\omega')\, \nu_\omega(d\omega')$ for each $\omega \in \{\kappa < \tau_n\}$. Thanks to boundedness, again, the integral $\int G\, d\mu'$ can be computed as an iterated integral $\int \mathbf{1}_{\{\tau_n>\kappa\}}(\omega) \int F_{\tau'_\omega}(\omega')\, \nu_\omega(d\omega')\mu(d\omega)$ and, so, $\int G\, d\mu' = 0$. □

3.4. **Sufficient conditions for disintegrability.**

3.4.1. *Shift operators.* The key to disintegrability for martingale-generated control correspondences is the existence of a shift operator, as described below. It plays the role of a partial inverse of the concatenation operator in the second argument.

Definition 3.11. A measurable map $\theta : \mathsf{Time} \times \Omega \to \Omega$ is said to be a **shift operator** if for all $\omega \in \Omega$, $t, s \in \mathsf{Time}$ and $\omega' \in \mathcal{C}_{\omega,t}$

(1) $\theta_t(\omega) \in \mathcal{C}_{\omega,t}$ and $\omega *_t \theta_t(\omega) = \omega$,
(2) $(\theta_t(\omega))_{\leq t+s} = (\theta_t(\omega_{\leq s}))_{\leq t+s}$

Remark 3.12. Since $\mathcal{F}_t = \sigma(T_t)$ on Ω, then part (2) of Definition 3.11 is equivalent to the $(\mathcal{F}_s, \mathcal{F}_{t+s})$-measurability of θ_t for all $t, s \in \mathsf{Time}$, i.e.,

$$\forall\, t, s \in \mathsf{Time}: \quad \theta_t^{-1}(\mathcal{F}_{t+s}) \subset \mathcal{F}_s$$

The stopping-time version of a shift operator θ is defined in the natural way

$$\theta_\tau(\omega) = \theta_{\tau(\omega)}(\omega),$$

where, for definiteness, we set $\theta_\infty(\omega) = \omega$, for all ω. This way, $\theta_\tau : \Omega \to \Omega$ is Borel measurable and retains the property that $\omega *_\tau \theta_\tau(\omega) = \omega$, for all $\omega \in \Omega$ and $\tau \in \mathsf{Stop}$.

Lemma 3.13. *For any $\kappa, \sigma \in \mathsf{Stop}$, the following is also a stopping time:*

$$\tau(\omega) := \kappa(\omega) + \sigma(\theta_\kappa(\omega))$$

Proof. Fix any $t \in \mathsf{Time}$ and $\omega \in \Omega$. In order to show $\{\tau \leq t\} \in \mathcal{F}_t$, it is enough to show that $\tau(\omega) \leq t$ if and only if $\tau(\omega_{\leq t}) \leq t$. Applying Proposition 2.2 to σ and using part (2) of the definition of θ gives the following equivalence:

$$\begin{aligned}
\tau(\omega) \leq t &\Leftrightarrow \sigma(\theta_{\kappa(\omega)}(\omega)) \leq t - \kappa(\omega) \\
&\Leftrightarrow \sigma((\theta_{\kappa(\omega)}(\omega))_{\leq t-\kappa(\omega)}) \leq t - \kappa(\omega) \\
&\Leftrightarrow \sigma((\theta_{\kappa(\omega)}(\omega_{\leq t}))_{\leq t-\kappa(\omega)}) \leq t - \kappa(\omega) \\
&\Leftrightarrow \sigma(\theta_{\kappa(\omega)}(\omega_{\leq t})) \leq t - \kappa(\omega)
\end{aligned}$$

First suppose $\tau(\omega) \leq t$. Since κ is a stopping time and $\kappa(\omega) \leq \tau(\omega) \leq t$, then $\kappa(\omega) = \kappa(\omega_{\leq t})$. Together with the above equivalence, this implies:

$$\begin{aligned}
\tau(\omega_{\leq t}) &= \kappa(\omega_{\leq t}) + \sigma(\theta_{\kappa(\omega_{\leq t})}(\omega_{\leq t})) \\
&= \kappa(\omega) + \sigma(\theta_{\kappa(\omega)}(\omega_{\leq t})) \leq t
\end{aligned}$$

For the other direction, suppose $\tau(\omega_{\leq t}) \leq t$. Since κ is a stopping time and $\kappa(\omega_{\leq t}) \leq \tau(\omega_{\leq t}) \leq t$, then $\kappa(\omega_{\leq t}) = \kappa(\omega)$. Therefore:

$$\kappa(\omega) + \sigma(\theta_{\kappa(\omega)}(\omega_{\leq t})) = \kappa(\omega_{\leq t}) + \sigma(\theta_{\kappa(\omega_{\leq t})}(\omega_{\leq t})) = \tau(\omega_{\leq t}) \leq t,$$

which implies $\tau(\omega) \leq t$ by the equivalence above. □

Proposition 3.14. *Let $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{t \in \mathsf{Time}})$ be a TC-space with concatenation operator $*$, on which a shift operator θ is defined. Suppose each $F \in \mathcal{D}$ is a canonically locally bounded TC-morphism into $(D^0_{\mathbb{R}}, \star)$, and that $*$ is a factor of X. Then, for each $\omega_0 \in \Omega$, $\mu \in \mathcal{P}(\mathcal{D}, X)(\omega_0)$, and $\kappa \in \mathsf{Stop}$ there exists a version $x \mapsto \bar{\nu}_x$ of the regular conditional probability $\mu(\theta_\kappa \in \cdot | X_\kappa = x)$ such that for $\nu = \bar{\nu} \circ X$ we have*

$$\nu \in \mathcal{S}(\mathcal{P}) \text{ and } \mu *_\kappa \nu = \mu.$$

In particular, $\mathcal{P}(\mathcal{D}, X)$ is disintegrable.

Proof. Having fixed a shift operator θ, we pick $\omega_0 \in \Omega, \mu \in \mathcal{P}(\omega_0)$ and $\kappa \in \mathsf{Stop}$. For a stopping time $\sigma \in \mathsf{QStop}$ and define

$$\sigma_n(\omega) = (\sigma \wedge \tau_n)(\omega)$$

$$\tau(\omega) = \kappa(\omega) + \sigma(\theta_\kappa(\omega))$$

so that τ is a stopping time by Lemma 3.13. Since F is a TC-morphism into $(D^0_{\mathbb{R}}, \star)$ Lemma 3.8 implies that

$$F_\tau(\omega) - F_\kappa(\omega) = F_{\kappa+\sigma_n(\theta_\kappa)}(\omega *_\kappa \theta_\kappa\omega) - F_\kappa(\omega) = F_{\sigma_n}(\theta_\kappa\omega) = F^n_\sigma(\theta_\kappa\omega).$$

The same Lemma implies that $|F|$ is bounded by $|F_\kappa| + M_n$ on the entire stochastic interval $[0, \tau]$. In particular, for $A_m = \{|F_\kappa| \le m\}$ we have

$$\mathbf{1}_{A_m} F_{\sigma_n} \circ \theta_\kappa = \mathbf{1}_{A_m}\Big(F_\tau - F_\kappa\Big) = \mathbf{1}_{A_m}\Big(F^{m+M_n}_\tau - F^{m+M_n}_\kappa\Big).$$

Since F^{m+M_n} is a bounded martingale under μ, for any bounded measurable function H on E we have $\int H(X(\omega_{\le\kappa}))\mathbf{1}_{A_m}(\omega)F^n_\sigma(\theta_\kappa\omega)\,\mu(d\omega) = 0$, and, given that F^n is bounded, we can pass to the limit $m \to \infty$ by the dominated convergence theorem to obtain

$$\int H(X(\omega_{\le\kappa}))F^n_\sigma(\theta_\kappa\omega)\,\mu(d\omega) = 0, \tag{3.5}$$

for all bounded and measurable H. With $\bar\nu_x$ denoting a version of the regular conditional distribution of θ_κ given $X_\kappa = x$, we then have

$$0 = \int H(X(\omega_{\le\kappa}))F^n_\sigma(\theta_\kappa\omega)\mu(d\omega) = \iint H(x)F^n_\sigma(\omega')\,\bar\nu_x(d\omega')\,\mu_{X_\kappa}(dx),$$

where μ_{X_κ} is the μ-distribution of X_κ. Since H is arbitrary, it follows that

$$\int F^n_\sigma\,d\bar\nu_x = 0 \text{ for } \mu_{X_\kappa}\text{-almost all } x \in E, \tag{3.6}$$

for all $\sigma \in \mathsf{QStop}$ and all $n \in \mathbb{N}$. Since QStop is countable, there exists a set $\mathcal{N}_1 \in \mathsf{Borel}(E)$ such that $\mu_{X_\kappa}(\mathcal{N}_1) = 0$, and the equality in (3.6) holds for all $x \in E \setminus \mathcal{N}_1$ and $\sigma \in \mathsf{QStop}$. Therefore $\bar\nu_x \in \mathcal{M}^{F,loc}$ for all $x \in E \setminus \mathcal{N}_1$.

Since $*$ is a factor of X, we have $X(T_\kappa(\omega)) = X_0(\theta_\kappa(\omega))$ for all ω, and so

$$1 = \int \mathbf{1}_{\{X_\kappa(\omega)=X_0(\omega_{\ge\kappa})\}}\,\mu(d\omega) = \iint \mathbf{1}_{\{x=X_0(\omega')\}}\bar\nu_x(d\omega')\mu_{X_\kappa}(dx),$$

This implies that there exists another zero set $\mathcal{N}_2 \in \mathsf{Borel}(E)$ such that $\mu_{X_\kappa}(\mathcal{N}_2) = 0$ and $X_0 = x$, $\bar\nu_x$-a.s. for all $x \in E \setminus \mathcal{N}_2$. Hence, $\bar\nu_x \in \bar{\mathcal{P}}(x)$ (where $\bar{\mathcal{P}}(x)$ is defined in (3.2)) for all $x \notin \mathcal{N}_1 \cup \mathcal{N}_2$. By picking a selector $\bar\nu'$ of $\bar{\mathcal{P}}$ (which is nonempty by Proposition 2.17) and using it to set the values of $\bar\nu_x$ on $\mathcal{N}_1 \cup \mathcal{N}_2$, we can arrange that $\bar\nu_x \in \bar{\mathcal{P}}(x)$, for all $x \in E$. □

3.5. **The main result for martingale-generated control correspondences.**

Theorem 3.15 (DPP for martingale-generated control correspondences)**.** *Let $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{t\in\mathsf{Time}})$ be a TC-space with concatenation operator $*$ and a shift operator θ. Suppose that X is a state map from Ω to*

a Polish space E such that $*$ is a factor of X, and that $\mathcal{D}$ is a countable collection of canonically locally bounded TC-morphisms from $(\Omega, *)$ into $(D^0_{\mathbb{R}}, \star)$. Let $\mathcal{P} = \mathcal{P}(\mathcal{D}, X)$, i.e.,

$$\bar{\mathcal{P}}(x) = \bigcap_{F\in\mathcal{D}} \mathcal{M}^{F,loc} \cap \left\{\mu \in \mathsf{Prob}(\Omega) \,:\, X_0 = x,\ \mu\text{-a.s.}\right\}$$
$$\mathcal{P}(\omega) = \bar{\mathcal{P}}(X(\omega)) \text{ for } \omega \in \Omega,$$

let $G \in \mathcal{L}^{1-0}(\mathcal{P})$ be a tail random variable, and let the value function $\bar{v}$ be given by

$$\bar{v}(x) = \sup_{\mu\in\bar{\mathcal{P}}(x)} \int G\, d\mu.$$

Then for all $\omega \in \Omega$, $x \in E$, and $\tau \in \mathsf{Stop}$ we have:

$$\bar{v}(x) = \sup_{\mu\in\bar{\mathcal{P}}(x)} \int \left(\bar{v}(X_\tau)\mathbf{1}_{\{\tau<\infty\}} + G\mathbf{1}_{\{\tau=\infty\}}\right) d\mu$$

Proof. Use Propositions 3.3, 3.7, and 3.14 to get the analyticity, concatenability, and disintegrability (respectively) of the control correspondence $(\mathcal{D}, X)$. Then apply Theorem 2.18. □

4. Application 1 - Controlled Diffusions in the Weak Formulation

4.1. Problem formulation and the main result.

Throughout this section we fix the following:

(1) a nonempty open set $\mathcal{O}$ in $\mathbb{R}^n$ and set $E = \operatorname{Cl}\mathcal{O}$ (the **state space**),
(2) a nonempty standard Borel space A, (the **control space**),
(3) Borel measurable functions $\beta : E \times A \to \mathbb{R}^n$ and $\sigma : E \times A \to \mathbb{R}^{n\times n}$ (the **coefficients**),
(4) a Borel measurable function $g : E \to [-\infty, \infty]$ (the **objective function**).

We remind the reader that $C_{E^{\partial\mathcal{O}}}$ denotes the set of all continuous trajectories with values in E that get absorbed once they hit the boundary $\partial\mathcal{O}$.

4.1.1. *Weak solutions to controlled SDEs.* With Einstein's convention of summation over repeated indices used throughout, we start by making precise what we mean by a controlled diffusion.

Definition 4.1 (Weak solutions to controlled SDEs)**.** A probability measure μ on $C_{E^{\partial\mathcal{O}}}$ is said to be a **weak solution of the controlled SDE**

$$d\xi^i_t = \beta^i(\xi_t, \alpha_t)\, dt + \sigma^i_k(\xi_t, \alpha_t)\, dW^k_t,\ \xi_0 = x, \tag{4.1}$$

with absorption in $\partial\mathcal{O}$ - denoted by $\mu \in \mathcal{L}^x(\beta, \sigma)$ - if there exists filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,\infty)}, \mathbb{P})$ on which three stochastic process $\{W_t\}_{t\in[0,\infty)}$, $\{\xi_t\}_{t\in[0,\infty)}$ and $\{\alpha_t\}_{t\in[0,\infty)}$ are defined, such that:

(1) W is an $\mathbb{R}^n$ valued $\{\mathcal{F}_t\}_{t\in[0,\infty)}$-Brownian motion,
(2) ξ is adapted and $\xi(\omega) \in C_{E^{\partial\mathcal{O}}}$ for all ω,

(3) α is A-valued and progressively measurable,
(4) $\int_0^t |\beta^i(\xi_u, \alpha_u)|\, du + \int_0^t (\sigma_k^i(\xi_u, \alpha_u))^2\, du < \infty$, a.s. for all i, k and $t \geq 0$,
(5) $\xi_t = x + \int_0^t \beta^i(\xi_u, \alpha_u)\, du + \int_0^t \sigma_k^i(\xi_u, \alpha_u)\, dW_t^k$, a.s., for all $t \in [0, \tau_{\partial\mathcal{O}}]$, where

$$\tau_{\partial\mathcal{O}} = \inf\{t \geq 0 \,:\, \xi_t \in \partial\mathcal{O}\}, \text{ and}$$

(6) μ is the law of $\xi_\cdot$ on $C_{E^{\partial\mathcal{O}}}$.

4.1.2. *The stochastic optimal control problem.* Given $x \in E$ and $\mu \in \mathcal{L}^x(\beta, \sigma)$, we set

$$J(\mu) = \mathbb{E}^\mu[G(\xi)] \text{ where } G(\xi) = \liminf_{t\to\infty} g(\xi_t), \tag{4.2}$$

with ξ denoting the coordinate map on $C_{E^{\partial\mathcal{O}}}$, where we assume that g is such that $\mathbb{E}^\mu[G^+(\xi)] < \infty$ for all $\mu \in \cup_{x\in E}\mathcal{L}^x(\beta, \sigma)$. The **value function** of the associated control problem is then given by

$$v(x) = \sup_{\mu\in\mathcal{L}^x(\beta,\sigma)} J(\mu),\ x \in E. \tag{4.3}$$

Remark 4.2. By choosing the state process ξ appropriately, this setup includes various common formulations of optimal stochastic control, including problems on a finite horizon (when $E = E_0 \times [0, T]$ and the last component plays the role of time) with terminal and/or running costs, discounted problems and stationary problems.

4.1.3. *DPP for controlled diffusions.*

Theorem 4.3 (A dynamic programming principle for controlled diffusions - the weak formulation)**.** *Suppose that,*

(1) there exist locally bounded real functions $\hat\beta : E \to \mathbb{R}$ and $\hat\sigma : E \to \mathbb{R}$ such that

$$|\beta^i(x, \alpha)| \leq \hat\beta(x) \text{ and } |\sigma_k^i(x, \alpha)| \leq \hat\sigma(x) \text{ for all } \alpha \in A,$$

(2) for each $x \in E$ we have $\mathcal{L}^x(\beta, \sigma) \neq \emptyset$, and
(3) $J(\mu) > -\infty$ for each $\mu \in \mathcal{L}^x(\beta, \sigma)$.

Then, the value function $v : E \to (-\infty, \infty]$ is universally measurable and satisfies the dynamic programming principle:

$$v(x) = \sup_{\mu\in\mathcal{L}^x(\beta,\sigma)} \mathbb{E}^\mu[v(\xi_\tau)\mathbf{1}_{\{\tau<\infty\}} + G(\xi_\cdot)\mathbf{1}_{\{\tau=\infty\}}], \text{ for all } x \in E,$$

for each (raw) stopping time τ on $C_{E^{\partial\mathcal{O}}}$.

Remark 4.4.

(1) Condition (1) in Theorem 4.3 is far from necessary. It is there to ensure existence and is placed mostly for convenience. It can be replaced by a different condition or relaxed by choosing a different control part Ω^α of the universal space $\Omega^{\alpha\xi}$ in the proof below.

(2) A very important feature of our control problem is that the law of the controlled process depends on the process α only through its Lebegue-a.e.-equivalence class (as a function of t), i.e., it is enough to think of α as an $\mathbb{L}^0_A$-random variable. This feature which is rarely stressed in the literature, allows us to construct a Polish setup for the problem, and consequently, prove the DPP.

4.2. **Proof of Theorem 4.3.**

Our proof of Theorem 4.3 consists of two steps. In the first one, we observe that the family $\mathcal{L}^x(\beta,\sigma)$ can be manufactured by varying admissible controls on a single, universal, filtered probability space, and that it admits a martingale characterization there. In the second one we show that this equivalent setup fits our abstract framework of Section 3 so that Theorem 3.15 can be applied.

4.2.1. *Construction of a universal setup.* Let $\Omega^\alpha = \mathbb{L}^0_A$ be the space of all Lebesgue-a.e equivalence classes of A-valued Borel functions from $[0,\infty)$ to A, and let Ω^ξ be the subspace $C_{E^{\partial\mathcal{O}}}$ of the canonical space $C_{\mathbb{R}^n}$. Both can be given the structure of a filtered measurable space, namely $(\Omega^\alpha, \mathcal{F}^\alpha, \mathbb{F}^\alpha = \{\mathcal{F}^\alpha_t\}_{t\in\mathsf{Time}})$, $(\Omega^\xi, \mathcal{F}^\xi, \mathbb{F}^\xi = \{\mathcal{F}^\xi_t\}_{t\in\mathsf{Time}})$, as described in more detail in subsection 2.2 and in Example 2.7. We define the (universal) filtered measurable space $(\Omega^{\alpha\xi}, \mathcal{F}^{\alpha\xi}, \mathbb{F}^{\alpha\xi} = \{\mathcal{F}^{\alpha\xi}_t\}_{t\in\mathsf{Time}})$ simply as their product. In particular $\mathcal{F}^{\alpha\xi}_t = \mathcal{F}^\alpha_t \otimes \mathcal{F}^\xi_t$. It will be used in the second step that $\Omega^{\alpha\xi}$ is, in fact, a T-space - the product of T-spaces Ω^α and Ω^ξ.

Let $\mathsf{Coord} = \{x_i, x_ix_j \,:\, 1\le i,j\le n\}$ be the family of coordinate functions and their products on $\mathbb{R}^n$, and let QCoord denote an arbitrary, but fixed throughout, countable family of bounded C^2-functions on $\mathbb{R}^n$ such that for each $f\in\mathsf{Coord}$ and each compact set $K\subseteq\mathbb{R}^n$ there exists $\tilde f\in\mathsf{QCoord}$ such that $f=\tilde f$ on K. Also, for $f\in C^2$ and $a\in A$ we define the $\mathcal{G}^a f$ by

$$(\mathcal{G}^a f)(x) = \beta^i(x,a)\partial_i f(x) + \tfrac{1}{2}\gamma^{ij}(a,x)\partial_{ij} f(x), \text{ with } \gamma^{ij} = \textstyle\sum_k \sigma^i_k\sigma^j_k,$$

Proposition 4.5 (A martingale characterization of weak solutions to controlled SDEs)**.** *The following two statements are equivalent for a probability measure μ on $C_{E^{\partial\mathcal{O}}}$:*

(1) μ is a weak solution to the controlled SDE (4.1) with absorption at $\partial\mathcal{O}$ starting at x, and
(2) there exists a probability measure $\bar\mu$ on $\Omega^{\alpha\xi}$ whose Ω^ξ-marginal is μ such that
  (a) $\xi_0 = x$, $\bar\mu$-a.s.,
  (b) $\int_0^t |\beta^i(\xi_u,\alpha_u)|\,du + \int_0^t(\sigma^i_k(\xi_u,\alpha_u))^2\,du < \infty$ for all i,k and $t\in[0,\tau_{\partial\mathcal{O}}]$, $\bar\mu$-a.s., and
  (c) for each $f\in\mathsf{QCoord}$, $f(\xi_t) - f(\xi_0) - \int_0^{t\wedge\tau_{\partial\mathcal{O}}} \mathcal{G}^{\alpha_u} f(\xi_u)\,du$ is an $(\{\mathcal{F}^{\alpha\xi}_t\}_{t\in[0,\infty)}, \bar\mu)$-local martingale.

If (1) holds, then (2c) is true for all $f\in C^2(E)$.

The proof follows, almost verbatim, the steps in the standard proof of the equivalence in the non-controlled case (see, e.g., Proposition 4.6, p. 315, [KS91]) so we omit the details. The only observation that needs to be made is that α is not a stochastic process in the classical sense. This difficulty can be circumvented by considering appropriate versions as in the following lemma. We remind the reader that an A-valued process

$\{\hat{\alpha}_t\}_{t\in[0,\infty)}$ is considered progressively measurable if $\{\phi(\hat{\alpha}_t)\}_{t\in[0,\infty)}$ is progressively measurable for each Borel measurable $\phi : A \to [-1, 1]$.

Lemma 4.6. *There exists an $\{\mathcal{F}^{\alpha\xi}_t\}_{t\in[0,\infty)}$-progressively measurable process $\{\hat{\alpha}_t\}_{t\in[0,\infty)}$ with values in A such that $\{\hat{\alpha}_t(\omega)\}_{t\geq 0}$ is a Leb-a.e.-representative of the coordinate map $\alpha(\omega)$ for each ω.*

Conversely, let (ξ, α) be a pair consisting of a continuous process ξ with values in $\mathbb{R}^n$ and an A-valued progressive process α defined on some filtered probability space $(\Omega, \{\mathcal{F}_t\}_{t\in[0,\infty)}, \mathcal{F}, \mathbb{P})$. Then (ξ, α) admits an $\Omega^{\alpha\xi}$-distribution, i.e., a probability measure $\bar{\mu}$ on $\Omega^{\alpha\xi}$ such that the $\mathbb{P}$-distribution of $\int_0^t \varphi(u, \xi_u, \alpha_u)\, du$ coincides with the $\bar{\mu}$-distribution of $\int_{[0,t]} \varphi(u, \alpha, \xi)\, d\lambda$, for each bounded and measurable φ and all $t \geq 0$.

Proof. Let ϕ be an isomorphism (a bimeasurable bijection) between A and the closed interval $[-1, 1]$. Given $\alpha(\omega) \in \mathbb{L}^0([0, \infty), A)$, we define $\hat{\alpha}$ by

$$\hat{\alpha}(t) = \phi^{-1}\Big(\liminf_{n\to\infty} \Phi^n_t(\omega)\Big) \text{ where } \Phi^n_t(\omega) = \tfrac{1}{n}\int_{(t-1/n)^+}^{t} \phi(\alpha_u(\omega))\, du.$$

It is straightforward to check that $\hat{\alpha}(\omega)$ is a representative of $\alpha(\omega)$ for each ω. Moreover $\phi(\hat{\alpha})$ (and, therefore, α) is a progressively-measurable process, as a pointwise limit of continuous adapted processes.

For the converse, and under the assumptions of the second part of the Lemma, let $\bar{\mu}$ be the pushforward of $\mathbb{P}$ via the map $\Phi : \Omega \to \Omega^{\alpha\xi}$ defined as follows:

$$\Phi(\omega) = \Big((\xi_t(\omega))_{t\geq 0}, \alpha(\omega)\Big),$$

where $\alpha(\omega)$ is the Leb-a.e.-equivalence class of $(\alpha_t(\omega))_{t\geq 0}$. (Progressive) measurability of α guarantees that Φ is a measurable map. The equality of the distributions of two integrals in the statement is then a simple consequence of the monotone-class theorem. □

4.2.2. *An application of the abstract DPP.* Proposition 4.5 allows us to reformulate our control problem so as to fit the setting of the first part of our paper. Indeed, it states that the value function $v(x)$ can be represented as

$$v(x) = \sup_{\bar{\mu}\in\bar{\mathcal{P}}(x)} \mathbb{E}^{\mu}[G(\xi)]$$

where $\bar{\mathcal{P}}(x)$ is the family of all probability measures on $\Omega^{\alpha\xi}$ such that (2a), (2b) and (2c) hold, and our job is to show that it is, in fact, a martingale generated control correspondence which satisfies all the requirements of the abstract Theorem 3.15.

Thanks to the discussion and examples in subsections 2.4 and 2.6, the space $\Omega^{\alpha\xi}$ admits a natural structure of a TC-space, with the strict concatenation used for the ξ component. The map $X : \Omega^{\alpha\xi} \to E$, given by $X(\xi, \alpha) = \liminf_{t\to\infty} \xi_t$ computed componentwise, and suitably measurably altered to take values in E and when the limits inferior take infinite values, so that $X_t(\xi, \alpha) = \xi_t$. Given that the concatenation operator in α requires no compatibility conditions, and the one in ξ is strict, the product concatenation operator $*$ factors through X (and is a factor of X). Also, there is a naturally-defined shift operator θ on $\Omega^{\alpha\xi}$.

Condition (1) of Theorem 4.3 takes care of the integrability condition (2b) of Proposition 4.5, so we can conclude that we are, indeed, dealing with a martingale-generated control correspondence with the state map X, generated by the family $\mathcal{D}$ which consists of (well-defined) maps of the form

$$F(\alpha,\xi)_t = f(\xi_t) - f(\xi_0) - \int_0^{t\wedge\tau_{\partial\mathcal{O}}} \mathcal{G}^{\alpha_u} f(\xi_u)\,du$$

with f ranging through the countable set QCoord. The last thing we need to check, before we can apply Theorem 3.15, is that each such F is a TC-morphism into $(D^0_{\mathbb{R}},\star)$. We fix $f \in \mathsf{QCoord}$, and note that the corresponding functional F clearly takes values in $D^0_{\mathbb{R}}$ and that it is non-anticipating. To establish the TC-morphism property let us fix $s,t \in \mathsf{Time}$ and $\omega,\omega' \in \Omega^{\alpha\xi}$ such that ω is compatible with ω' at t. The case of $s \le t$ is straightforward, so suppose $s > t$. Since the ξ component uses the strict concatenation operator, then $\xi_t(\omega) = \xi_0(\omega')$, and furthermore:

$$\tau_{\partial\mathcal{O}}(\omega) \le t \quad\Leftrightarrow\quad \xi_t(\omega) \in \partial\mathcal{O} \quad\Leftrightarrow\quad \xi_0(\omega') \in \partial\mathcal{O} \quad\Leftrightarrow\quad \tau_{\partial\mathcal{O}}(\omega') = 0$$

Combining this with the properties of concatenation gives:

$$\begin{aligned}
\int_{t\wedge\tau_{\partial\mathcal{O}}}^{s\wedge\tau_{\partial\mathcal{O}}} \mathcal{G}^{\alpha_u} f(\xi_u(\omega *_t \omega'))\,du &= \mathbf{1}_{\{\tau_{\partial\mathcal{O}}>t\}}(\omega) \int_{t\wedge\tau_{\partial\mathcal{O}}}^{s\wedge\tau_{\partial\mathcal{O}}} \mathcal{G}^{\alpha_u} f(\xi_u(\omega *_t \omega'))\,du \\
&= \mathbf{1}_{\{\tau_{\partial\mathcal{O}}>0\}}(\omega') \int_0^{(s-t)\wedge\tau_{\partial\mathcal{O}}} \mathcal{G}^{\alpha_u} f(\xi_u(\omega'))\,du \\
&= \int_0^{(s-t)\wedge\tau_{\partial\mathcal{O}}} \mathcal{G}^{\alpha_u} f(\xi_u(\omega'))\,du
\end{aligned}$$

Putting everything together gives:

$$\begin{aligned}
F(\omega *_t \omega')_s &= f(\xi_s(\omega *_t \omega')) - f(\xi_0(\omega *_t \omega')) - \int_0^{s\wedge\tau_{\partial\mathcal{O}}} \mathcal{G}^{\alpha_u} f(\xi_u(\omega *_t \omega'))\,du \\
&= \left(f(\xi_t(\omega *_t \omega')) - f(\xi_0(\omega *_t \omega')) - \int_0^{t\wedge\tau_{\partial\mathcal{O}}} \mathcal{G}^{\alpha_u} f(\xi_u(\omega *_t \omega'))\,du \right) \\
&\quad + \left(f(\xi_s(\omega *_t \omega')) - f(\xi_t(\omega *_t \omega')) - \int_{t\wedge\tau_{\partial\mathcal{O}}}^{s\wedge\tau_{\partial\mathcal{O}}} \mathcal{G}^{\alpha_u} f(\xi_u(\omega *_t \omega'))\,du \right) \\
&= \left(f(\xi_t(\omega)) - f(\xi_0(\omega)) - \int_0^{t\wedge\tau_{\partial\mathcal{O}}} \mathcal{G}^{\alpha_u} f(\xi_u(\omega))\,du \right) \\
&\quad + \left(f(\xi_{s-t}(\omega')) - f(\xi_0(\omega')) - \int_0^{(s-t)\wedge\tau_{\partial\mathcal{O}}} \mathcal{G}^{\alpha_u} f(\xi_u(\omega'))\,du \right) \\
&= F(\omega)_t + F(\omega')_{t-s} = (F(\omega) \star_t F(\omega'))_s
\end{aligned}$$

4.3. **Viscosity solutions.**

We conclude this example by showing how our result can be applied to show that value functions of stochastic control problems are viscosity solutions to the associated Hamilton-Jacobi-Bellman equations under weak conditions. In particular, we do not require that the equation itself admit an a-priori solution, or that any

solution is smooth or unique (i.e, that the comparison principle hold). Our results, in particular, imply some of the results in [BS13], [BT11] and the follow-up papers under weaker assumptions. We note that the lack of any strong ellipticity allow us keep assuming, without loss of generality, that the problem is time-independent; time can be incorporated as just another (space) variable with linear dynamics and the terminal condition imposed as part of the boundary condition.

For a C^2 function $\varphi : \mathcal{O} \to \mathbb{R}$ we define the **Hamiltonian** $H\varphi : \mathcal{O} \to (-\infty, \infty]$ by

$$H\varphi(x) = \sup_{a\in A} \mathcal{G}^a\varphi(x) = \sup_{a\in A} \Big(\beta^i(x,a)\partial_{x_i}\varphi(x) + \tfrac{1}{2}\gamma^{ij}(x,a)\partial_{x_ix_j}\varphi(x)\Big).$$

4.3.1. *The viscosity property of the value function.*

Definition 4.7. Let v be a real-valued function defined in a neighborhood $\mathcal{V}$ of a point $\bar{x} \in \mathcal{O}$, and let v_* and v^* denote its lower and upper semicontinuous envelopes, respectively. We say that v is a

(1) **viscosity supersolution** of the equation $Hv = 0$ at $\bar{x}$ if $H\varphi(\bar{x}) \le 0$ for each $\varphi \in C^2(\mathcal{V})$ with the property that $\varphi(\bar{x}) = v_*(\bar{x})$ and $\varphi(x) < v_*(x)$ for $x \in \mathcal{V} \setminus \{\bar{x}\}$, and
(2) **viscosity subsolution** of the equation $Hv = 0$ at $\bar{x}$ if $H\varphi(\bar{x}) \le 0$ for each $\varphi \in C^2(\mathcal{V})$ with the property that $\varphi(\bar{x}) = v^*(\bar{x})$ and $\varphi(x) > v^*(x)$ for $x \in \mathcal{V} \setminus \{\bar{x}\}$.

A function which is both a viscosity supersolution and a viscosity subsolution is called a **viscosity solution** to $Hv = 0$ at $\bar{x}$.

For $x \in \mathbb{R}^n$ and $r > 0$ we define

$$\tau^{r,x} = \inf\{t \ge 0 \,:\, d(x, \xi_t) \ge r\} \wedge r,$$

where d denotes the Euclidean distance on $\mathbb{R}^n$, so that $\tau^{r,x}$ is a raw stopping times on $\Omega^{\alpha\xi}$.

Theorem 4.8. *Given $\bar{x} \in \mathcal{O}$, suppose that there exists a neighborhood $\mathcal{V}$ of $\bar{x}$ in $\mathcal{O}$ such that*

*(1) **(availability of DPP)** the assumptions of Theorem 4.3 hold and v is finite on $\mathcal{V}$,*
*(2) **(continuity of coefficients)** $x \mapsto \beta^i(x,a)$ and $x \mapsto \sigma^i_k(x,a)$ are continuous functions on $\mathcal{V}$ for all $a \in A$,*
*(3) **(admissibility of locally constant controls)** there exists a constant $r > 0$ such that for each $x \in \mathcal{V}$ and $a \in A$ there exists a control process $\{\alpha_t\}_{t\in[0,\infty)}$ and an associated weak solution $\{\xi_t\}_{t\in[0,\infty)}$ of the controlled SDE (4.1) with $\xi_0 = x$ (defined on some filtered probability space) such that*

$$\alpha_t = a \ \text{ for } t \in [0,\tau] \text{ a.s., where } \tau = \inf\{t \ge 0 \,:\, d(\xi_t, \bar{x}) \ge r\} \wedge r.$$

Then the value function v is a viscosity solution to $Hv = 0$ at x_0.

Proof. We split the proof into two parts, in which we establish the supersolution and the subsolution property of v separately.

The supersolution property. We take $\varphi \in C^2$ which touches v_* at $\bar{x}$ from below, i.e. $v_*(\bar{x}) = \varphi(\bar{x})$ and $\varphi(x) < v_*(x)$ for $x \neq \bar{x}$ This implies that there exists a sequence $\{x_m\}_{m\in\mathbb{N}}$ such that

$$v(x_m) \leq \varphi(x_m) + \tfrac{1}{m} \text{ and } d(x_m, \bar{x}) \leq \tfrac{1}{m}. \tag{4.4}$$

Suppose, for contradiction, that $H\varphi(\bar{x}) > 0$. Then there exists $a \in A$ such that $(\mathcal{G}^a\varphi)(\bar{x}) > 0$. Since $\mathcal{G}^a\varphi$ is continuous in x, there exist constants $\varepsilon > 0$ and $r > 0$ such that $(\mathcal{G}^a\varphi)(x) \geq \varepsilon$ when $d(x, \bar{x}) \leq r$. Using the fact that $\varphi(x) < v_*(x)$ as soon as $x \neq \bar{x}$ and that the function $v_* - \varphi$ is lower semicontinuous, we find that

$$\delta = \min\{v_*(x) - \varphi(x) \,:\, d(x, \bar{x}) = r\} > 0.$$

For each $m \in \mathbb{N}$, let μ_m be the law of the weak solution $\{\xi_t\}_{t\in[0,\infty)}$ described in part 3 of the statement, where we assume, without loss of generality, that the same constant $r > 0$, as above, can be used. Proposition 4.5 and the local nonnegativity of $\mathcal{G}^a\varphi - \varepsilon$ imply that $\varphi(\xi_t) - \varepsilon t$ is a bounded μ_m-submartingale under μ_m on $[0, \tau^{r,\bar{x}}]$. Therefore, with $\tau = \tau^{r,\bar{x}}$ and for $m > 1/r$, we get

$$\begin{aligned}
\varphi(x_m) &\leq \mathbb{E}^{\mu_m}[\varphi(\xi_\tau) - \varepsilon\tau] \leq \mathbb{E}^{\mu_m}[\varphi(\xi_\tau)\mathbf{1}_{\{\tau<r\}}] + \mathbb{E}^{\mu_m}[(\varphi(\xi_\tau) - \varepsilon r)\mathbf{1}_{\{\tau=r\}}] \\
&\leq \mathbb{E}^{\mu_m}[(v_*(\xi_\tau) - \delta)\mathbf{1}_{\{\tau<r\}}] + \mathbb{E}^{\mu_m}[(v_*(\xi_\tau) - \varepsilon r)\mathbf{1}_{\{\tau=r\}}] \\
&\leq \mathbb{E}^{\mu_m}[v_*(\xi_\tau)] - \min(\delta, \varepsilon r).
\end{aligned}$$

Using the dynamic programming principle of Theorem 4.3 and the relation (4.4) above, we finally obtain

$$v(x_m) - \tfrac{1}{m} + \min(\delta, \varepsilon r) \leq \mathbb{E}^{\mu_m}[v_*(\xi_\tau)] \leq \mathbb{E}^{\mu_m}[v(\xi_\tau)] \leq \sup_{\mu\in\mathcal{L}^{x_m}(\beta,\sigma)} \mathbb{E}^{\mu}[v(\xi_\tau)] = v(x_m),$$

and reach a contradiction by taking m large enough.

The subsolution property. We pick $\varphi \in C^2$ which touches v^* at $\bar{x}$ from above, i.e. $v^*(\bar{x}) = \varphi(\bar{x})$ and $\varphi(x) > v_*(x)$ for $x \neq \bar{x}$. As in the first part of the proof, this implies that there exists a sequence $\{x_m\}_{m\in\mathbb{N}}$ such that

$$v(x_m) \geq \varphi(x_m) - \tfrac{1}{m} \text{ and } d(x_m, \bar{x}) \leq \tfrac{1}{m}. \tag{4.5}$$

Suppose, for contradiction, that $H\varphi(\bar{x}) < 0$. Being representable as a supremum of continuous functions, $H\varphi$ is upper semicontinuous, and so there exist constants $r > 0$ and $\varepsilon > 0$ such that $H\varphi(x) \leq -\varepsilon$ for all x with $d(x, \bar{x}) \leq r$. Using the fact that $\varphi(x) > v^*(x)$ as soon as $x \neq \bar{x}$ and that the function $\varphi - v^*$ is lower semicontinuous, we find, as above, that

$$\delta = \min\{\varphi(x) - v^*(x) \,:\, d(x, \bar{x}) = r\} > 0.$$

Let the laws $(\mu_m)_{m\in\mathbb{N}}$ be defined as in the first part of the proof, so that under each μ_m the process $\varphi(\xi_t) + \varepsilon t$ is supermartingale on $[0, \tau^{r,\bar{x}}]$. It follows that, with $\tau = \tau^{r,x}$, we have

$$\begin{aligned}
\varphi(x_m) &\geq \mathbb{E}^{\mu}[\varphi(\xi_\tau) + \varepsilon\tau] = \mathbb{E}^{\mu}[(\varphi(\xi_\tau) + \varepsilon\tau)\mathbf{1}_{\{\tau=r\}}] + \mathbb{E}^{\mu}[(\varphi(\xi_\tau) + \varepsilon\tau)\mathbf{1}_{\{\tau<r\}}] \\
&\geq \mathbb{E}^{\mu}[(v^*(\xi_\tau) + \delta)\mathbf{1}_{\{\tau=r\}}] + \mathbb{E}^{\mu}[(\varphi(\xi_\tau) + \varepsilon r)\mathbf{1}_{\{\tau<r\}}] \geq \mathbb{E}^{\mu}[v(\xi_\tau)] + \min(\delta, \varepsilon r)
\end{aligned}$$

We take a supremum over all $\mu \in \mathcal{P}^{x_m}$ on the right hand side and use the DPP to conclude that $\varphi(x_m) \geq v(x_m) + \min(\delta, \varepsilon r)$ for all m - a contradiction with (4.5). □

5. Application 2 - Singular Control Problems

5.1. **The Monotone-Follower Problem.** We show how singular control problems fit our framework on the example of the celebrated Monotone Follower Problem (first formulated by Bather and Chernoff [BC67], analyzed rigorously by Karatzas and Shreve in [KS84] and studied in many papers since). Formally, the Monotone Follower Problem asks for a minimal cost incurred while controlling a Brownian motion W by adding to it a non-decreasing left-continuous process α. The cost is typically given by

$$\mathbb{E}\left[\int_0^\tau f(t)d\alpha_t + g(W_T - \alpha_T) + \int_0^\tau h(t, W_t - \alpha_t)dt\right],$$

where g and h model the deviation of the controlled trajectory $W + \alpha$ from the desired optimal position and f plays the role of "fuel" cost.

5.2. **Formulation in our framework.** To make it easier to focus on the issues pertinent to the proof of the DPP, we generalize the problem to a degree. The continuous variables, such as time, running cost or the Brownian motion from the above description will be replaced by a general, multidimensional diffusion. This will not only allow us to reuse many of the conclusion of the previous section, but also to get a clearer understanding of the role different parts play as far as DPP is concerned.

5.2.1. *The space* Ω. Given $m, n \in \mathbb{N}$, let $\mathcal{O} \subseteq \mathbb{R}^{m+n}$ be a nonempty open set with closure $E = \operatorname{Cl}\mathcal{O}$, which will play the role of our state space. Let $C_\mathbb{R}$ and $G_\mathbb{R}$ denote the canonical spaces of all continuous and càglàd paths, respectively, with values in $\mathbb{R}$, and let $G^\uparrow_\mathbb{R}$ denote the subset of $G_\mathbb{R}$ consisting of nondecreasing paths. Let Ω^X denote the space of paths in $C^m_\mathbb{R} \times (G^\uparrow_\mathbb{R})^n$ with values in E, absorbed upon entry in $\partial\mathcal{O}$, i.e. stopped at the canonical stopping time

$$\tau_{\partial\mathcal{O}}(\omega) = \inf\{t \geq 0 \,:\, \omega(t) \in \partial\mathcal{O}\} \text{ for } \omega \in \Omega^X.$$

With the control component taking value in $\Omega^\alpha = G^\uparrow_\mathbb{R}$, the space Ω is defined as the subset of $\Omega^X \times \Omega^\alpha$ consisting of those paths (X, α) stopped once X hits $\partial\mathcal{O}$. Equivalently, Ω is the set of paths in $\Omega^X \times \Omega^\alpha$ that get absorbed once the coordinate map (X, α) enters the set $\partial\mathcal{O} \times \mathbb{R}$. We overload the notation $\tau_{\partial\mathcal{O}}$ to denote the hitting time of $\partial\mathcal{O} \times \mathbb{R}$, when considered as a stopping time on Ω.

The first n coordinate maps on Ω (corresponding to continuous paths) are denoted by Y, the next m (corresponding to left-continuous paths) by Z and the last one by α, so that $\omega(t) = (Y_t(\omega), Z_t(\omega), \alpha_t(\omega))$, for $\omega \in \Omega$ and $t \geq 0$.

5.2.2. *The T-space, TC-space structures.* We use the standard truncations on each of the components of Ω. To see that Ω carries a natural structure of a T-space, we simply need to combine the discussion in paragraph 2.2.2 in subsection 2.2 with the product construction of paragraph 2.4.3. It can be upgraded to a TC-space by equipping it with

(1) the strict concatenation operator $\bullet$, as defined in equation (2.7) in subsection 2.6.1, on Ω^X (i.e., for the first $m + n$ coordinates), and

(2) the adjusted concatenation $\star$, as defined by (2.8), on $\Omega^\alpha = G_{\mathbb{R}}^{\uparrow}$ (for the last coordinate).

The so-obtained concatenation on Ω will be denoted by $* = (\bullet, \star)$.

5.2.3. *The state X and the cost functional G.* Let

$$\underline{\lim} : \Omega^X \to E$$

be a "Banach limit", i.e., a map with the following properties:

(1) Its value on the trajectory ω coincides with the pointwise limit $\lim_{t\to\infty} \omega(t)$ whenever this limit exists; in particular, it equals the value at which ω is absorbed, when absorption happens.
(2) It returns a value in E in a Borel measurable way.
(3) It is invariant under the action of the shift operator.

A fairly general construction of such a map on spaces of right-continuous trajectories can be found in [Žit14, Lemma 3.12, p. 1614]. A closer inspection of the proof reveals that the right-continuity assumption can be replaced by the assumption of left continuity, and that the conclusion of the theorem applies to the present setting. Given such a map $\underline{\lim}$, we simply define

$$X(\omega) = \underline{\lim}\, \omega^X \text{ for } \omega = (\omega^X, \omega^\alpha) \in \Omega.$$

In agreement with the definition of the coordinate maps Y_t and Z_t above, we split the first n and the last m coordinates of X into Y and Z, i.e. $X(\omega) = (Y(\omega), Z(\omega))$. This way, since we are working with the standard truncation, we have

$$X_t(\omega) = X(\omega_{\le t}) = (Y(\omega_{\le t}), Z(\omega_{\le t})) = (Y_t(\omega), Z_t(\omega)).$$

With X defined, the cost function G is simply a Borel function of X:

$$G(\omega) = g(X(\omega)), \tag{5.1}$$

where we assume throughout that g is nonnegative so as not to need to pay attention to integrability conditions in the sequel. Much less restrictive assumptions are also possible.

5.2.4. *The control correspondence $\mathcal{P}$.* The control correspondence describing our monotone-follower problem will naturally factor through the state map X, so we define the family $(\overline{\mathcal{P}}^x)_{x\in E}$, and use it to construct the control correspondence in the usual way $\mathcal{P}(\omega) = \overline{\mathcal{P}}^{X(\omega)}$. Heuristically, the dynamics of the state $X_t = (Y_t, Z_t)$ under $\overline{\mathcal{P}}^x$ for $x \in E$ can be described as follows: Y is a diffusion on $\mathbb{R}^n$, with coefficients depending on X_t, absorbed once X_t hits $\partial\mathcal{O}$. The left-continuous component Z "moves" as follows

$$dZ_t = c(Y_t)\, d\alpha_t, \tag{5.2}$$

where c is a vector of m nonnegative and continuous functions.

To simplify the exposition, we express $\overline{\mathcal{P}}$ as an intersection of two control correspondences $\overline{\mathcal{P}}_c$ and $\overline{\mathcal{P}}_l$, where $\overline{\mathcal{P}}_c$ "constrains" the motion of continuous portion Y and $\overline{\mathcal{P}}_l$ the left-continuous portion Z, of the state process

X. To define $\overline{\mathcal{P}}_c$ we follow the approach of section 4 and consider a family $\mathcal{D}_c$ of maps from Ω to $C_{\mathbb{R}} \subseteq D_{\mathbb{R}}$ given by

$$F^f(\omega)_t = f(Y_t(\omega)) - f(Y_0(\omega)) - \int_0^{t\wedge\tau_{\partial\mathcal{O}}(\omega)} \mathcal{G}^{Z_u(\omega)} f(Y_u(\omega))\, du,$$

where f ranges through the set QCoord as in the second paragraph of section 4.2.1, and $\mathcal{G}^z$ is a differential operator of the form

$$(\mathcal{G}^z f)(y) = \beta^i(y,z)\partial_i f(y) + \tfrac{1}{2}\Sigma^{ij}(y,z)\partial_{ij} f(y), \text{ with } \Sigma^{ij} = \textstyle\sum_k \sigma^i_k \sigma^j_k,$$

with coefficients β and σ measurable, locally bounded and globally Lipschitz in y. These conditions are imposed to ensure that the control correspondence $\overline{\mathcal{P}}_c$ generated by $(\mathcal{D}, X)$ is well-defined and non-empty.

We note here that the dependence of any F on α is trivial; that means that even though we think of α as a control, its influence on F factors entirely through the left-continuous process Z and does not show up in $\overline{\mathcal{P}}_c$. To describe how Z depends on α, we need to introduce the control correspondence $\overline{\mathcal{P}}_l$. To describe it rigorously, we first need to agree on how to define the integral with respect to a left-continuous process in (5.2) above. Such a construction has been carried out already in [KS84, Remark 5.3., p. 873]; we simply exhibit parts of their discussion for the convenience of the reader. Given a nondecreasing càglàd function $\alpha : [0,\infty) \to \mathbb{R}$, we define the càdlàg function $\alpha^+ : [0,\infty) \to \mathbb{R}$ by setting $\alpha^+_t := \alpha_{t+} = \inf_{u>t} \alpha_u$. For a locally bounded Borel function $\gamma : [0,\infty) \to \mathbb{R}$, we define

$$\int_0^t \gamma(u)\, d\alpha_u := \begin{cases} 0, & t = 0, \\ \gamma(0)\Delta\alpha_0 + \int_{(0,t)} \gamma(u)\, d\alpha^+_u, & t > 0, \end{cases} \tag{5.3}$$

where $\Delta\alpha_0 = \alpha^+_0 - \alpha_0$ and the integral on the right-hand side is the Lebesgue-Stieltjes integral with respect to the measure induced by α^+ on $(0,t)$. We immediately observe that the function $\zeta_t = \int_0^t \gamma(u)\, d\alpha_u$ is càglàd and satisfies $\zeta^+_0 = \gamma(0)\Delta\alpha_0$. We also record, for later use, the following characterization:

Lemma 5.1. *Suppose that $\alpha \in G^{\uparrow}_{\mathbb{R}}$ and that $\gamma : [0,\infty) \to \mathbb{R}$ is continuous. For $\zeta \in G^{\uparrow}_{\mathbb{R}}$, the following two conditions are equivalent*

(1) $\zeta = \zeta_0 + \int_0^{\cdot} \gamma(u)\, d\alpha_u$, and
(2) $\Delta\zeta_{0+} = \gamma(0)\Delta\alpha_{0+}$ and for all rational $0 < r < s$ and each $n \in \mathbb{N}$ there exist rationals $p, q \in (r,s)$ such that

$$(\gamma(p) - \tfrac{1}{n})(\alpha^+_s - \alpha^+_r) \le \zeta^+_r - \zeta^+_s \le (\gamma(q) + \tfrac{1}{n})(\alpha^+_s - \alpha^+_r). \tag{5.4}$$

Proof. Thanks to right continuity of ζ^+, (1) above is equivalent to $\Delta\zeta_{0+} = \gamma(0)\Delta\alpha_{0+}$ and

$$\zeta^+ - \zeta^+_0 = \int_{(0,\cdot]} \gamma(u)\, d\alpha^+_u. \tag{5.5}$$

Using the right continuity of α^+ and the continuity of γ (which guarantees the equivalence between the Riemann-Stieltjes and the Lebesgue-Stieltjes integration in this case) we conclude that the equality in (5.5) is

equivalent to

$$\forall\, u < v \in (0,\infty),\ \left(\inf_{t\in[u,v]} \gamma(t)\right)(\alpha_u^+ - \alpha_v^+) \le \zeta_u^+ - \zeta_v^+ \le \left(\sup_{t\in[u,v]} \gamma(t)\right)(\alpha_u^+ - \alpha_v^+).$$

Thanks to the right continuity of α^+ and ζ^+, this is easily seen to be equivalent to (second statement in) (2) above. □

Given a continuous function $c : \mathbb{R}^m \to \mathbb{R}^n$, for $x \in E$ we define

$$\overline{\mathcal{P}}_l(x) = \{\mu \in \mathsf{Prob}(\Omega)\ :\ X_0 = x,\ Z = Z_0 + \int_0^{\cdot} c(Y_u)\, d\alpha_u, \mu - \text{a.s.}\}$$

where the left-continuous integral is interpreted component-wise. We set $\overline{\mathcal{P}}(x) = \overline{\mathcal{P}}_c(x) \cap \overline{\mathcal{P}}_l(x)$ and define the **value function** of the associated control problem by

$$v(x) = \inf_{\mu\in\overline{\mathcal{P}}(x)} \mathbb{E}^\mu[G],\ x \in E. \tag{5.6}$$

Remark 5.2. To see how the classical monotone-follower fits into this framework, we take $Y = (T, W, H)$ and $Z = (L, C)$, where, informally, the components have the following dynamics:

$$\begin{aligned}
dT_t &= -dt, && \text{time-to-go,}\\
dW_t &= dW_t, && \text{Brownian motion}\\
dH_t &= h(-T_t, W_t, L_t)\, dt, && \text{running cost}\\
dL_t &= d\alpha_t, && \text{position of the follower, and}\\
dC_t &= f(T_t)\, d\alpha_t, && \text{fuel cost,}
\end{aligned}$$

where f and h are nonnegative and continuous. The state space E is defined by $E = \operatorname{Cl}\mathcal{O}$, where

$$\mathcal{O} = (-\infty, 0) \times \mathbb{R} \times (0,\infty) \times (0,\infty) \times (0,\infty),$$

so as to keep the components H and C nonnegative. This will also make sure that the state process will exit E when (and only when) $T_t = 0$. A typical cost functional G will be of the form $G(X) = H + C + g(W, L)$, where g is a nonnegative Borel function.

5.3. **The Dynamic Programming Principle.** With all the components of our framework in place, we are ready to prove the following result:

Proposition 5.3 (DPP for the monotone-follower problem)**.** *Given the setting described above, the value function $v : E \to (-\infty, \infty]$ is universally measurable and satisfies the dynamic programming principle*

$$v(x) = \sup_{\mu\in\overline{\mathcal{P}}^x} \mathbb{E}^\mu[v(X_\tau)\mathbf{1}_{\{\tau<\infty\}} + G\mathbf{1}_{\{\tau=\infty\}}],\ \textit{for all } x \in E,$$

for each (raw) stopping time τ on $C_{E^{\partial\mathcal{O}}}$.

Proof. As in the previous section, we establish three key properties, namely, analyticity, concatenability and disintegrability, and use Theorem 2.18. The additional requirement that G be a tail random variable follows directly from the fact that it was defined in (5.1) using a "Banach limit", i.e., in a shift-invariant way. The membership in the class $\mathcal{L}^{0-1}(\overline{\mathcal{P}})$ of lower semi-integrable random variables is guaranteed by the assumption that the function g in (5.1) is bounded from below.

Analyticity: To establish the analyticity of $\overline{\mathcal{P}}$ it will be enough to show that both $\overline{\mathcal{P}}_c$ and $\overline{\mathcal{P}}_l$ are analytic (see Remark 2.16). All the maps in $\mathcal{D}_c$ are clearly non-anticipating and take values in $C_{\mathbb{R}} \subseteq D_{\mathbb{R}}$, so we can apply Proposition 3.3 to conclude that $\overline{\mathcal{P}}_c$ is analytic.

The analyticity of $\overline{\mathcal{P}}_l$, follows from Lemma 5.1. Indeed, it expresses $\overline{\mathcal{P}}_l$ as a result of a countable collection of Borel-preserving operations on cylinders.

Concatenability: Just like in the case of analyticity, Remark 2.16 allows us to prove concatenability of $\overline{\mathcal{P}}$ by proving it separately for $\overline{\mathcal{P}}_c$ and $\overline{\mathcal{P}}_l$. Starting with $\overline{\mathcal{P}}_c$, we simply note that the maps F^f in $\mathcal{D}_c$ are $C_{\mathbb{R}}$-valued and therefore canonically locally bounded. Their TC-morphism property is established exactly like in section 4.2.2 above, so we can use Proposition 3.7 to conclude that $\overline{\mathcal{P}}_c$ is closed under concatenation.

Next, we turn to the concatenability of $\overline{\mathcal{P}}_l$. Given $t \geq 0$ let $\omega, \omega' \in \Omega$ be such that

(1) $X_0(\omega') = X_t(\omega)$,
(2) $C(\omega) = \int_0^{\cdot} c(Y_u(\omega))\, d\alpha_u(\omega)$, and
(3) $C(\omega') = \int_0^{\cdot} c(Y_u(\omega'))\, d\alpha_u(\omega')$,

We note that these properties hold for (ω, ω') with probability 1, under $\mu \otimes_t \nu$. Using the fact that $*$ is strict in the first $m+n$ components and adjusted in α, we observe that for $s > t$ we have

$$C_s(\omega *_t \omega') - C_{t+}(\omega *_t \omega') = C_{s-t}(\omega') - C_{0+}(\omega') = \int_{(0,s-t)} c(Y_u(\omega'))\, d\alpha^+_u(\omega') = \\ = \int_{(t,s)} c(Y_{u-t}(\omega'))\, d\alpha^+_{u-t}(\omega') = \int_{(t,s)} c(Y_u(\omega *_t \omega'))\, d\alpha^+_u(\omega *_t \omega'),$$

as well as

$$C_{t+}(\omega *_t \omega') - C_t(\omega *_t \omega') = C_{0+}(\omega') - C_0(\omega') = c(Y_0(\omega'))(\alpha_{0+}(\omega') - \alpha_0(\omega')) \\ = c(Y_t(\omega *_t \omega'))(\alpha_{t+}(\omega *_t \omega') - \alpha_t(\omega *_t \omega')).$$

These two observations make it straightforward to complete the proof of the concatenability of $\overline{\mathcal{P}}_l$.

Disintegrability: While disintegrability cannot be established by showing it for $\overline{\mathcal{P}}_c$ and $\overline{\mathcal{P}}_l$ separately, we can use Proposition 3.14, whose conditions are easily shown to hold in the present setting, to perform most of the work for us. Indeed, given $\omega_0 \in \Omega$ and $\mu \in \overline{\mathcal{P}}_c(\omega_0)$ and $\kappa \in \mathsf{Stop}$, it states that there exists a version $x \mapsto \bar{\nu}_x$ of the regular conditional probability $\mu(\theta_\kappa \in \cdot | X_\kappa = x)$ with the following two properties: 1) $\nu \in \mathcal{S}(\overline{\mathcal{P}}_c)$ and 2) $\mu = \mu *_\kappa \nu$, where $\nu = \bar{\nu} \circ X$. In order to complete the proof, we need to show that a version of ν with $\nu \in \mathcal{S}(\overline{\mathcal{P}}_l)$, can be constructed. Let A denote the set of all $\omega \in \Omega$ such that

$Z(\omega) - Z_0(\omega) = \int_0^{\cdot} c(Y_u(\omega))\, d\alpha_u(\omega)$. For any $x \in E$ and any $\mu \in \bar{\mathcal{P}}(x)$ we have $\mu(A) = 1$. Therefore, by the concatenability property established above, we have

$$1 = \int \mathbf{1}_A(\omega)\, \mu(d\omega) = \int \int \mathbf{1}_A(\omega *_\kappa \omega') \bar{\nu}_{X_\kappa(\omega)}(d\omega') \mu(d\omega) \leq \int \int \mathbf{1}_A(\omega') \bar{\nu}_x(d\omega') \mu_{X_\kappa}(dx),$$

and, so, there exists a $\mathcal{N}_1 \in \mathsf{Borel}(E)$ with $\mu_{X_\kappa}(\mathcal{N}_1) = 0$ and such that for $x \in E \setminus \mathcal{N}_1$ we have $\bar{\nu}_x(A) = 1$. Similarly, $\bar{\nu}_x(X_0 = x) = 1$ for all $x \in E \setminus \mathcal{N}_2$, where $\mathcal{N}_2$ is a μ_{X_κ}-null set in $\mathsf{Borel}(E)$. It remains to redefine $\bar{\nu}$ on $\mathcal{N}_1 \cup \mathcal{N}_2$ so that $\bar{\nu} \in \mathcal{S}(\overline{\mathcal{P}}_c) \cap \mathcal{S}(\overline{\mathcal{P}}_l)$. This is easily achieved by picking an arbitrary selector $\bar{\nu}' \in \mathcal{S}(\bar{\overline{\mathcal{P}}}_c \cap \bar{\overline{\mathcal{P}}}_l)$ and setting setting $\bar{\nu}_x = \bar{\nu}'_x$ for all $x \in \mathcal{N}_1 \cup \mathcal{N}_2$. □

REFERENCES

[BC67] BATHER, J. AND CHERNOFF, H. Sequential decisions in the control of a spaceship. In *Proc. Fifth Berkeley Sympos. Mathematical Statistics and Probability (Berkeley, Calif., 1965/66), Vol. III: Physical Sciences*, pages 181–207. Univ. California Press, Berkeley, Calif., 1967.

[BN12] BOUCHARD, B. AND NUTZ, M. Weak dynamic programming for generalized state constraints. *SIAM J. Control Optim.*, 50(6):3344–3373, 2012.

[Bor89] BORKAR, V. S. *Optimal control of diffusion processes*, volume 203 of *Pitman Research Notes in Mathematics Series*. Longman Scientific & Technical, Harlow, 1989.

[BS78] BERTSEKAS, D. P. AND SHREVE, S. E. *Stochastic optimal control*, volume 139 of *Mathematics in Science and Engineering*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1978. The discrete time case.

[BS12] BAYRAKTAR, E. AND SÎRBU, M. Stochastic Perron's method and verification without smoothness using viscosity comparison: the linear case. *Proc. Amer. Math. Soc.*, 140(10):3645–3654, 2012.

[BS13] BAYRAKTAR, E. AND SÎRBU, M. Stochastic Perron's method for Hamilton-Jacobi-Bellman equations. *SIAM J. Control Optim.*, 51(6):4274–4294, 2013.

[BT11] BOUCHARD, B. AND TOUZI, N. Weak dynamic programming principle for viscosity solutions. *SIAM J. Control Optim.*, 49(3):948–962, 2011.

[BV10] BOUCHARD, B. AND VU, T. N. The obstacle version of the geometric dynamic programming principle: application to the pricing of American options under constraints. *Appl. Math. Optim.*, 61(2):235–265, 2010.

[DVJ03] DALEY, D. J. AND VERE-JONES, D. *An introduction to the theory of point processes. Vol. I*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2003. Elementary theory and methods.

[EK81] EL KAROUI, N. Les aspects probabilistes du contrôle stochastique. In *Ninth Saint Flour Probability Summer School—1979 (Saint Flour, 1979)*, volume 876 of *Lecture Notes in Math.*, pages 73–238. Springer, Berlin, 1981.

[ET13a] EL KAROUI, N. AND TAN, X. Capacities, Measurable Selection and Dynamic Programming Part I: Abstract Framework, 2013. arXiv:1310.3363.

[ET13b] EL KAROUI, N. AND TAN, X. Capacities, Measurable Selection and Dynamic Programming Part II: Application in Stochastic Control Problems, 2013. arXiv:1310.3364.

[FS93] FLEMING, W. H. AND SONER, H. M. *Controlled Markov processes and viscosity solutions*, volume 25 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1993. ISBN 0-387-97927-1.

[KS84] KARATZAS, I. AND SHREVE, S. E. Connections between optimal stopping and singular stochastic control. I. Monotone follower problems. *SIAM J. Control Optim.*, 22(6):856–877, 1984.

[KS91] KARATZAS, I. AND SHREVE, S. E. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[NvH13] NUTZ, M. AND VAN HANDEL, R. Constructing sublinear expectations on path space. *Stochastic Processes and Their Applications*, 123 (8):3100-3121, 2013.

[Sri98] SRIVASTAVA, S. M. *A course on Borel sets*, volume 180. Springer, New York, 1998.

[ST02a] SONER, H. M. AND TOUZI, N. Dynamic programming for stochastic target problems and geometric flows. *J. Eur. Math. Soc. (JEMS)*, 4(3):201–236, 2002.

[ST02b] SONER, H. M. AND TOUZI, N. Stochastic target problems, dynamic programming, and viscosity solutions. *SIAM J. Control Optim.*, 41(2):404–424, 2002.

[Str77] STRICKER, C. Quasimartingales, martingales locales, semimartingales et filtration naturelle. *Probability Theory and Related Fields*, 39:55–63, 1977.

[Žit14] ŽITKOVIĆ, G. Dynamic programming for controlled markov families: abstractly and over martingale measures. *SIAM Journal of Control and Optimization*, 52(3):1597–1621, 2014.

ROMAN FAYVISOVICH, DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF TEXAS AT AUSTIN

Email address: rfayvisovich@math.utexas.edu

GORDAN ŽITKOVIĆ, DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF TEXAS AT AUSTIN

Email address: gordanz@math.utexas.edu