# ON THE INFORMATIONAL STRUCTURE IN OPTIMAL DYNAMIC STOCHASTIC CONTROL

MATIJA VIDMAR AND SAUL JACKA

ABSTRACT. We formulate a massively general framework for optimal dynamic stochastic control problems which allows for a control-dependent informational structure. The issue of informational consistency is brought to light and investigated. Bellman's principle is formulated and proved. In a series of related results, we expound on the informational structure in the context of (completed) natural filtrations of (stochastic) processes.

## 1. INTRODUCTION: MOTIVATION – LITERATURE OVERVIEW – STRUCTURE OF PAPER & INFORMAL SUMMARY OF RESULTS

Optimal dynamic stochastic control, is (i) *stochastic*, in the sense that the output of the system is random; it is (ii) *optimal control*, to the extent that, with the goal of optimizing its expectation, said output is subject to exogenous influence; and it is (iii) *dynamic*, in that this control, at any one instant of time, is adapted to the current and past state of the system. In general, however, the controller in a dynamic stochastic control problem can observe some part only of all of the "universal" information which is being accumulated (e.g. he may only be able to observe the controlled process, or, worse, some (non one-to-one) function thereof). Moreover, this "observed information" may depend on the chosen control.

It is therefore only reasonable to insist *a priori* on all the admissible controls (as processes) to be adapted (or even predictable with respect) to what is thus a *control-dependent-informational-flow* that is being acquired by the controller. And while not doing so can emerge as having been (in some sense) immaterial *a posteriori* (e.g. since the optimal control turned out to have been only a function of the present (and past) state of an observable process), then this will, generally speaking, only *happen* to have been the case, rather than it *needing* to have been so, which, presumably, is the preferred of the two alternatives.

Some (informal) examples follow.

---

2010 *Mathematics Subject Classification.* Primary: 93E20; Secondary: 60G05, 60A10, 28A05.

*Key words and phrases.* Optimal dynamic stochastic control; informational structure; stopped (completed) natural filtrations and (completed) natural filtrations at stopping times; Bellman's principle.

This research was mainly conducted while MV was a PhD student in the Department of Statistics at the University of Warwick, and a recipient of a scholarship from the Slovene Human Resources Development and Scholarship Fund (contract number 11010-543/2011). MV acknowledges this latter support and thanks the hospitality of the former. Special mention of Jon Warren, for many helpful discussions on some of the material from Part 2 of this paper.

(•) *Job hunting.* Consider an academic trying to optimize his choice of current institution. The decision of whether or not to move, and where to move, will be based (in particular) on the knowledge of the quality of the research and of the amenities of the various universities/institutes that he can choose from. This knowledge itself will depend (at least partially) on the chosen sequence of institutions that he has already affiliated himself with (indeed, since the quality of the institutions changes over time, it will depend on it in a temporally local way – only once he has chosen to move and has spent some time at his new location, will he be able to properly judge the eventual ramifications of his choice). Costs are associated with moving, but also with staying for too long at an institution, which hinders his academic development; rewards are associated with the quality of the chosen institution; further, more or less resources can be devoted to determining the quality of institutions before a (potential) move. (See Section 5 for a toy model reminiscent of this situation.)

(•) Or consider *quality control* in a production line of light-bulbs. A fault with the machine may cause all the bulbs from some time onwards to be faulty. Once this has been detected, the machine can of course be fixed and the production of functioning bulbs restored. However, only the condition of those light-bulbs which are taken out of the production line, and tested, can actually be observed. A cost is associated with this process of testing (but, clearly, also with issuing faulty light-bulbs). Conversely, rewards accrue from producing functioning bulbs.

(•) From the theory of controlled SDEs, a folklore example of *loss of information* is Tanaka's SDE: let $X$ be a Brownian motion, $W_t := \int_0^t \mathrm{sgn}(X_s)dX_s$, $t \in [0, \infty)$. Then the completed natural filtration of $W$ is strictly included in the completed natural filtration of $X$ [11, p. 302].

(•) In economics the so-called class of *bandit* models is well-known (see the excellent literature overview in [8]). These are "sequential decision problems where, at each stage, a resource like time, effort or money has to be allocated strategically between several options, referred to as the arms of the bandit." And further: "The key idea in this class of models is that agents face a trade-off between experimentation (gathering information on the returns to each arm) and exploitation (choosing the arm with the highest expected value)." [8, p. 2].

(•) Miscellaneous other situations in which the control is non-trivially, and naturally, adapted, or even previsible, with respect to an informational flow, which it itself influences and helps engender, abound: controlling the movement of a probe in a stochastic field, only the local values of the field being observable, and costs/rewards associated with the speed of movement/intensity of the field (cf. Example 2.11); a similar situation for movement on a random graph, being only able to observe the values attached to the vertices visited; additively controlling a process, but observing the sum of the process itself and of the control etc.

In short, the phenomenon is ubiquitous; we suggest it is the norm, rather than the exception.

In particular, then, one should like a *general* framework for stochastic control (equipped with a suitable abstract version of Bellman's principle) which makes such a control-dependent informational flow explicit and inherent in its machinery. The (seeming) circularity of the controls being adapted to an informational structure which they themselves help engender, makes this a somewhat delicate point. Indeed, it would seem, this is the one aspect of general (abstract) dynamic stochastic control which has not yet received due attention in existing literature. Hitherto, only a single, non-control dependent, observable [5] informational flow [15, 14] appears to have been allowed. (This is of course not to say that the phenomenon has not entered *and* been studied in the literature in *specific situations/problems*, see e.g. [6, 8, 7, 16] [1, Chapter 8] and the references therein; the focus there having been (for the most part) on reducing (i.e. *a priori* proving a suitable equivalence of) the original control problem, which is based on *partial control-dependent observation*, to an associated '*separated*' problem, which is based on *complete observation*.)

In the present paper we attempt to fill this gap in the literature, by putting forward a general stochastic control framework which explicitly allows for a control-dependent informational flow – as it were 'embracing it', rather than trying to circumvent it in some or another manner. (Recognizing that it may not always be advantageous, or indeed possible, to work with an equivalent (but more complex) 'separated' formulation.) In particular, we provide a fully general (modulo the relevant (technical) condition, see Assumption 4.3) abstract version of Bellman's principle in such a setting. This is the content of Part 1. Specifically, Section 2 formally defines a system of stochastic control (in which observed information is an explicit function of control); Section 3 discusses its conditional payoff and 'Bellman' system; Section 4 formulates Bellman's principle – Theorem 4.6 is our main result. Lastly, Section 5 contains the solution to a formal (but artificial) example, illustrating some of the main ideas of this paper; several other (counter)examples are also given along the way.

Now, a crucial requirement for the above programme to be successful is that of informational consistency over controls (cf. Assumption 2.9): if two controls agree up to a certain time, then what we have observed up to that time should also agree. Especially at the level of random (stopping) times, this becomes a non-trivial statement – for example, when the observed information is that generated by a (controlled) process, which is often the case. We expound on this issue of informational consistency in the context of (completed) natural filtrations of processes in Part 2. Specifically, we consider there, amongst others relevant, the following natural and pertinent question, which is interesting in its own right: if $X$ and $Y$ are two processes, and $S$ a stopping time of one or both of their (possibly completed) natural filtrations, with the stopped processes agreeing, $X^S = Y^S$ (possibly only with probability one), must the two (completed) natural filtrations *at* the time $S$ agree also? To answer this question (with proofs) is non-trivial in the temporally non-discrete case, and several related findings are obtained along the way (see the introductory remarks to Part 2, on p. 21, for a more detailed account). In essence they are (consequences of/connected with) a generalization (Theorem 6.6) of (a part of) Galmarino's test, available in literature for coordinate processes on canonical spaces [4, p. 149, Theorem IV.100] [11, p. 320, Lemma 4.18], and extended here to a not necessarily canonical setting.

**Conventions**. Throughout this paper, for a probability measure $\mathsf{P}$ on $\Omega$ and $A \subset \Omega$, some property in $\omega \in A$ will be said to hold $\mathsf{P}$-a.s. on $A$, if the set of $\omega \in A$ for which the property does not hold is first measurable (i.e. belongs to the domain of $\mathsf{P}$), and second is of $\mathsf{P}$-measure zero. When $A = \Omega$, we shall of course just say that the property holds $\mathsf{P}$-a.s. Finally, for $\mathcal{L} \subset 2^\Omega$ and $A \subset \Omega$, $\mathcal{L}|_A := \{L \cap A : L \in \mathcal{L}\}$ is the trace of $\mathcal{L}$ on $A$.

## Part 1. Optimal dynamic stochastic control with control-dependent information

As announced in the Introduction, we provide and analyze in this part, a framework for optimal dynamic stochastic control, in which information is explicitly control-dependent. The informational flow itself is modeled using filtrations, and this can be done in one of the following two, essentially different, ways:

(1) Dealing with events 'with certainty', irrespective of the presence of probability.
(2) Dealing with events 'up to a.s. equality', insisting that the filtrations be complete relative to the underlying probability measure(s).

We develop the second 'probabilistic' approach – of complete filtrations – in parallel to the default first – for lack of a better word, 'measure-theoretic' – setting. Indeed, the formal differences between the two approaches are minor. For the most part one has merely to add, in the 'complete' setting, a number of a.s. qualifiers. We will put these, and any other eventual differences of the second approach as compared to the first, in {} braces. This will enforce a strict separation between the two settings, while still allowing us to repeat ourselves as little as possible.

## 2. Stochastic control systems

We begin by specifying the formal ingredients of a system of optimal dynamic stochastic control.

**Setting 2.1** (Stochastic control system)**.** A **stochastic control system** consists of:

(i) A set $T$ with a linear (antisymmetric, transitive & total) ordering $\leq$. We will assume (for simplicity) either $T = \mathbb{N}_0$, or else $T = [0, \infty)$, with the usual order. $T$ is the *time set.*
(ii) A set $\mathbf{C}$. The *set of admissible controls.* (These might be {equivalence classes of} processes or stopping times, or something different altogether.)
(iii) A set $\Omega$ endowed with a collection of $\sigma$-algebras $(\mathcal{F}^c)_{c \in \mathbf{C}}$. $\Omega$ is the *sample space* and $\mathcal{F}^c$ is all the *information accumulated* (but not necessarily acquired by the controller) by the "end of time" or, possibly, by a "terminal time", when $c$ is the chosen control. For example, in the case of optimal *stopping*, when there is given a process $X$, and it is *stopped*, the set of controls $\mathbf{C}$ would be the {equivalence classes of the} stopping times of the {completed}

natural filtration of $X$, and for any $S \in \mathbf{C}$, $\mathcal{F}^S = \sigma(X^S)$, the $\sigma$-field generated by the stopped process {or its completion}. We understand here optimal stopping in the *strict* sense: the exogenous act is that of *stopping*, not *sampling*; after the process has been stopped, it ceases to change.

(iv) $(\mathsf{P}^c)_{c\in\mathbf{C}}$, a collection of {complete} probability measures, each $\mathsf{P}^c$ having domain which includes the {$\mathsf{P}^c$-complete} $\sigma$-field $\mathcal{F}^c$ (for $c \in \mathbf{C}$). The controller chooses a probability measure from the collection $(\mathsf{P}^c)_{c\in\mathbf{C}}$. This allows for incorporation of the Girsanov approach to control, wherein the controller is seen as affecting the probability measure, rather than the random payoff. From the point of view of information, being concerned with laws rather than random elements, it is of course somewhat unnatural. Nevertheless, we will formally allow for it – it costs us nothing.

(v) A function $J : \mathbf{C} \to [-\infty, +\infty]^\Omega$, each $J(c)$ being $\mathcal{F}^c$ measurable (as $c$ runs over $\mathbf{C}$) {and defined up to $\mathsf{P}^c$-a.s. equality}. We further insist $\mathsf{E}^{\mathsf{P}^c} J(c)^- < \infty$ for all $c \in \mathbf{C}$. Given the control $c \in \mathbf{C}$, $J(c)$ is the random *payoff*. Hence, in general, we allow both the payoff, as well as the probability law, to vary.

(vi) A collection of filtrations[1] $(\mathcal{G}^c)_{c\in\mathbf{C}}$ on $\Omega$. It is assumed $\mathcal{G}^c_\infty := \vee_{t\in T}\mathcal{G}^c_t \subset \mathcal{F}^c$, and (for simplicity) that $\mathcal{G}^c_0$ is $\mathsf{P}^c$-trivial (for all $c \in \mathbf{C}$) {and contains all the $\mathsf{P}^c$-null sets}, while $\mathcal{G}^c_0 = \mathcal{G}^d_0$ {i.e. the null sets for $\mathsf{P}^c$ and $\mathsf{P}^d$ are the same} and $\mathsf{P}^c|_{\mathcal{G}^c_0} = \mathsf{P}^d|_{\mathcal{G}^d_0}$ for all $\{c, d\} \subset \mathbf{C}$. $\mathcal{G}^c_t$ is the *information acquired* by the controller by time $t \in T$, if the control chosen is $c \in \mathbf{C}$ (e.g. $\mathcal{G}^c$ may be the {completed} natural filtration of an observable process $X^c$ which depends on $c$). Perfect recollection is thus assumed.

**Definition 2.2** (Optimal expected payoff)**.** We define $v := \sup_{c\in\mathbf{C}} \mathsf{E}^{\mathsf{P}^c} J(c)$ ($\sup \emptyset := -\infty$), the **optimal expected payoff**. Next, $c \in \mathbf{C}$ is said to be **optimal** if $\mathsf{E}^{\mathsf{P}^c} J(c) = v$. Finally, a $\mathbf{C}$-valued net is said to be **optimizing** if its limit is $v$.

*Remark* 2.3.

(1) It is, in some sense, no restriction, to have assumed the integrability of the negative parts of $J$ in Setting 2.1(v). For, allowing any extra controls $c$ for which $\mathsf{E}^{\mathsf{P}^c} J(c)^- = \infty$, but for which $\mathsf{E}^{\mathsf{P}^c} J(c)$ would still be defined, would not change the value of $v$ (albeit it could change whether or not $\mathbf{C}$ is empty, but this is a trivial consideration).

(2) It is *not* natural *a priori* to insist on each $J(c)$ being $\mathcal{G}^c_\infty$-measurable (for $c \in \mathbf{C}$). The outcome of our controlled experiment need not be known to us (the controller) *at all* – not even by the end of time; all we are concerned with is the maximization of its expectation.

(3) In the case $\mathbf{C}$ is a collection of processes, the natural requirement is for each such process $c \in \mathbf{C}$ to be adapted (perhaps even previsible with respect) to $\mathcal{G}^c$. If it is a collection of

[1] All filtrations will be assumed to have the parameter set $T$.

random times, then each such $c \in \mathbf{C}$ should presumably be a (possibly predictable) stopping time of $\mathcal{G}^c$. But we do not formally insist on this.

We now introduce the concept of a controlled time, a (we would argue, natural) generalization of the notion of a stopping time to the setting of control-dependent filtrations.

**Definition 2.4** (Controlled times)**.** A collection of random times $\mathcal{S} = (\mathcal{S}^c)_{c\in\mathbf{C}}$ is called a **controlled time**, if $\mathcal{S}^c$ is a {defined up to $\mathsf{P}^c$-a.s. equality} stopping time of $\mathcal{G}^c$ for every $c \in \mathbf{C}$.

*Example* 2.5. A typical situation to have in mind is the following. What is observed is a process $X^c$, its values being contingent on the chosen control $c$ (this may, but need not, be the controlled process, e.g. it might be some non one-to-one function of it). Then $\mathcal{G}^c$ is the {completed} natural filtration of $X^c$. Letting, for example, for each $c \in \mathbf{C}$, $\mathcal{S}^c$ be the first entrance time of $X^c$ into some fixed set, the collection $(\mathcal{S}^c)_{c\in\mathbf{C}}$ would constitute a controlled time (as long as one can formally establish the stopping time property). ◇

**Definition 2.6** (Deterministic and control-constant times)**.** If there is some $a \in T \cup \{\infty\}$, such that $\mathcal{S}^c(\omega) = a$ for {$\mathsf{P}^c$-almost} all $\omega \in \Omega$, and every $c \in \mathbf{C}$, then $\mathcal{S}$ is called a **deterministic** time. More generally, if there is a random time $S$, which is a stopping time of $\mathcal{G}^c$ and $\mathcal{S}^c = S$ {$\mathsf{P}^c$-a.s} for each $c \in \mathbf{C}$, then $\mathcal{S}$ is called a **control-constant time**.

As yet, $\mathbf{C}$ is an entirely abstract set with no dynamic structure attached to it. The following establishes this structure. The reader should think of $\mathcal{D}(c,\mathcal{S})$ as being the controls "agreeing {a.s.} with $c$ up to time $\mathcal{S}$". (Example 2.11 and Section 5 contain definitions of the collections $\mathcal{D}(c,\mathcal{S})$ in the (specific) situations described there.)

**Setting 2.7** (Stochastic control system (cont'd) – control dynamics)**.** There is given *a* collection $\mathbf{G}$ of controlled times. Further, adjoined to the stochastic control system of Setting 2.1, is a family $(\mathcal{D}(c,\mathcal{S}))_{(c,\mathcal{S})\in\mathbf{C}\times\mathbf{G}}$ of subsets of $\mathbf{C}$ for which:

(1) $c \in \mathcal{D}(c,\mathcal{S})$ for all $(c,\mathcal{S}) \in \mathbf{C}\times\mathbf{G}$.
(2) For all $\mathcal{S} \in \mathbf{G}$ and $\{c,d\} \subset \mathbf{C}$, $d \in \mathcal{D}(c,\mathcal{S})$ implies $\mathcal{S}^c = \mathcal{S}^d$ {$\mathsf{P}^c$ & $\mathsf{P}^d$-a.s}.
(3) If $\{\mathcal{S},\mathcal{T}\} \subset \mathbf{G}$, $c \in \mathbf{C}$ and $\mathcal{S}^c = \mathcal{T}^c$ {$\mathsf{P}^c$-a.s}, then $\mathcal{D}(c,\mathcal{S}) = \mathcal{D}(c,\mathcal{T})$.
(4) If $\{\mathcal{S},\mathcal{T}\} \subset \mathbf{G}$ and $c \in \mathbf{C}$ for which $\mathcal{S}^d \leq \mathcal{T}^d$ {$\mathsf{P}^d$-a.s.} for $d \in \mathcal{D}(c,\mathcal{T})$, then $\mathcal{D}(c,\mathcal{T}) \subset \mathcal{D}(c,\mathcal{S})$.
(5) For each $\mathcal{S} \in \mathbf{G}$, $\{\mathcal{D}(c,\mathcal{S}) : c \in \mathbf{C}\}$ is a partition of $\mathbf{C}$.
(6) For all $(c,\mathcal{S}) \in \mathbf{C}\times\mathbf{G}$: $\mathcal{D}(c,\mathcal{S}) = \{c\}$ (resp. $\mathcal{D}(c,\mathcal{S}) = \mathbf{C}$), if $\mathcal{S}^c$ is identically {or $\mathsf{P}^c$-a.s.} equal to $\infty$ (resp. 0).[2]

**Definition 2.8.** Pursuant to Setting 2.7(5), we write $\sim_{\mathcal{S}}$ for the equivalence relation induced by the partition $\{\mathcal{D}(c,\mathcal{S}) : c \in \mathbf{C}\}$.

[2]This is not really a restriction; see Remark 4.4(ii).

Using this dynamical structure, a natural assumption on the temporal consistency of the filtrations $(\mathcal{G}^c)_{c\in\mathbf{C}}$ and the measures $(\mathsf{P}^c)_{c\in\mathbf{C}}$ — indeed a *key* condition on whose validity we shall insist throughout — is as follows:

**Assumption 2.9** (Temporal consistency)**.** *For all $\{c,d\}\subset\mathbf{C}$ and $\mathcal{S}\in\mathbf{G}$ satisfying $c\sim_{\mathcal{S}} d$, we have $\mathcal{G}^c_{\mathcal{S}^c}=\mathcal{G}^d_{\mathcal{S}^d}$ and $\mathsf{P}^c|_{\mathcal{G}^c_{\mathcal{S}^c}}=\mathsf{P}^d|_{\mathcal{G}^d_{\mathcal{S}^d}}$.*

Several remarks are now in order.

($\bullet$) First, when $\mathcal{S}$ is a non-control-constant time (e.g. when $\mathcal{S}^c$ is the first entrance time into some fixed set of an observed controlled process $X^c$, as $c$ runs over $\mathbf{C}$), then already the provisions of Setting 2.7 (leaving aside, for the moment, Assumption 2.9) are far from being entirely innocuous, viz. condition Setting 2.7(2) (which, e.g. would then be saying that controls agreeing with $c$ up to the first entrance time $\mathcal{S}^c$ of the observed controlled process $X^c$, will actually leave the latter invariant). They are thus as much a restriction/consistency requirement on the family $\mathcal{D}$, as they are on which controlled times we can put into the collection $\mathbf{G}$. Put differently, $\mathbf{G}$ is not (necessarily) a completely arbitrary, if non-specified, collection of controlled times. For, a controlled time is just *any* family of $\mathcal{G}^c$-stopping times, as $c$ runs over the control set $\mathbf{C}$. The members of $\mathbf{G}$, however, enjoy the further property of "agreeing between two controls, if the latter coincide prior to them". This is of course trivially satisfied for deterministic times (and, more generally, control-constant stopping times), but may hold of other controlled times as well.

($\bullet$) Second, the choice of the family $\mathbf{G}$ is guided by the specific problem at hand: not all controlled times are of interest. For example, sometimes the deterministic times may be relevant, the others not. On the other hand, it may be possible to effect the act of "controlling" only at some collection of (possibly non-control-constant) stopping times – then *these* times may be particularly worthy of study. The following example illustrates this point already in the control-independent informational setting (anticipating somewhat certain concepts, like the Bellman system, and conditional optimality, which we have not yet formally introduced; the reader might return to it once he has studied Sections 3 and 4).

*Example* 2.10. Given: a probability space $(\Omega,\mathcal{F},\mathsf{P})$; on it an $\mathbb{N}_0$-valued càdlàg Poisson process $N$ of unit intensity with arrival times $(S_n)_{n\in\mathbb{N}_0}$, $S_0:=0$, $S_n<\infty$ for all $n\in\mathbb{N}$; an independent independency of random signs $(R_n)_{n\in\mathbb{N}_0}$ with values in $\{-1,+1\}$, $\mathsf{P}(R_n=+1)=1-\mathsf{P}(R_n=-1)=2/3$.

The "observed process" is

$$W:=N+\int_0^\cdot\sum_{n\in\mathbb{N}_0}R_n\mathbb{1}_{[S_n,S_{n+1})}d\mathrm{Leb}$$

(so to $N$ is added a drift of $R_n$ during the random time interval $[S_n,S_{n+1})$, $n\geq 0$). Let $\mathcal{G}$ be the *natural* filtration of $W$. Remark the arrival times of $N$ are stopping times of $\mathcal{G}$.

The set of controls $\mathbf{C}$, on the other hand, consists of real-valued, measurable processes, starting at 0, which are adapted to the natural filtration of the bivariate process $(W\mathbb{1}_{\{\Delta N\neq 0\}}, N)$ (where $\Delta N$ is the jump process of $N$; intuitively, we must decide on the strategy for the whole of $[S_n, S_{n+1})$ based on the information available at time $S_n$ already, $n \geq 0$). For $X \in \mathbf{C}$ consider the *penalty* functional

$$J(X) := \int_{[0,\infty)} e^{-\alpha t}\mathbb{1}_{(0,\infty)} \circ |X_t - W_t| dt$$

(continuous penalization with discounting rate $\alpha \in (0, +\infty)$ of any deviation from the process $W$ by the control $X$). Let $v := \inf_{X\in\mathbf{C}} \mathsf{E}J(X)$ be the optimal expected penalty; clearly an optimal control is the process $\hat{X}$ which takes the value of $W$ at the instances which are the arrival times of $N$ and assumes a drift of $+1$ in between those instances, so that $v = 1/(3\alpha)$. Next, for $X \in \mathbf{C}$, let

$$V_S^X := \mathsf{P}\text{-essinf}_{Y\in\mathbf{C}, Y^S = X^S} \mathsf{E}[J(Y)|\mathcal{G}_S], \quad S \text{ a stopping time of } \mathcal{G},$$

be the Bellman system. We shall say $Y \in \mathbf{C}$ is conditionally admissible at time $S$ for the control $X$ (resp. conditionally optimal at time $S$), if $Y^S = X^S$ (resp. $V_S^Y = \mathsf{E}[J(Y)|\mathcal{G}_S]$ $\mathsf{P}$-a.s.). Denote $V := V^{\hat{X}}$ for short.

(1) We maintain first that the process $(V_t)_{t\in[0,\infty)}$ (the Bellman process (i.e. system at the deterministic times) for the optimal control), is not mean nondecreasing (in particular, is not a submartingale, let alone a martingale with respect to $\mathcal{G}$) and admits no a.s. right-continuous version.

For, $V_0 = v$; while for $t \in (0,\infty)$, the following control, denoted $X^\star$, is, apart from $\hat{X}$, also conditionally admissible at time $t$ for $\hat{X}$: It assumes the value of $W$ at the instances of the arrival times of $N$, and a drift of $+1$ in between those intervals, *until* before (inclusive of) time $t$; strictly after time $t$ and until strictly before the first arrival time of $N$ which is $\geq t$, denoted $S_t$, it takes the values of the process which starts at the value of $W$ at the last arrival time of $N$ strictly before $t$ and a drift of $-1$ thereafter; and after (and inclusive of) the instance $S_t$, it resumes to assume the values of $W$ at the arrival times of $N$ and a drift of $+1$ in between those times. Notice also that $R_t\mathbb{1}(t \text{ is not an arrival time of } N) \in \mathcal{G}_t$, where $R_t = \sum_{n\in\mathbb{N}_0} R_n\mathbb{1}_{[S_n,S_{n+1})}(t)$, i.e. $R_t\mathbb{1}(t \text{ is not an arrival time of } N)$ is the drift at time $t$, on the (almost certain) event that $t$ is not an arrival time of $N$, zero otherwise. It follows that, since $\hat{X}$ is conditionally admissible for $\hat{X}$ at time $t$:

$$V_t \leq \mathsf{E}[J(\hat{X})|\mathcal{G}_t],$$

so $\mathsf{E}V_t\mathbb{1}_{\{R_t=+1\}} \leq \mathsf{E}J(\hat{X})\mathbb{1}_{\{R_t=+1\}}$; whereas since $X^\star$ is also conditionally admissible at time $t$ for $\hat{X}$:

$$V_t \leq \mathsf{E}[J(X^\star)|\mathcal{G}_t],$$

so $\mathsf{E}V_t\mathbb{1}_{\{R_t=-1\}} \leq \mathsf{E}J(X^\star)\mathbb{1}_{\{R_t=-1\}} = \mathsf{E}J(\hat{X})\mathbb{1}_{\{R_t=-1\}} - \mathsf{E}\int_{(t,S_t)} e^{-\alpha t}dt\mathbb{1}_{\{R_t=-1\}} = \mathsf{E}J(\hat{X})\mathbb{1}_{\{R_t=-1\}} - \frac{1}{\alpha}e^{-\alpha t}(1 - \frac{1}{1+\alpha})\frac{1}{3}$ (properties of marked Poisson processes). Summing the two inequalities we obtain

$$\mathsf{E}V_t \leq v - \frac{1}{3(1+\alpha)}e^{-\alpha t},$$

implying the desired conclusion (for the nonexistence of a right continuous version, assume the converse, reach a contradiction via uniform integrability).

(2) We maintain second that the process $(V^X_{S_n})_{n\in\mathbb{N}_0}$, however, is a discrete-time submartingale (and martingale with $X = \hat{X}$) with respect to $(\mathcal{G}_{S_n})_{n\in\mathbb{N}_0}$, for all $X \in \mathbf{C}$.

For $X = \hat{X}$, this follows at once from the obvious observation that $\hat{X}$ is conditionally optimal at each of the arrival instances of $N$. On the other hand, for arbitrary $X \in \mathbf{C}$, $n \in \mathbb{N}_0$, $G \in \mathcal{G}_{S_n}$, and $\{Y, Z\} \subset \mathbf{C}$ with $Y^{S_n} = X^{S_n} = Z^{S_n}$, the control which coincides with $Y$ (hence $Z$) on $[0, S_n]$ and then with $Y$ (resp. $Z$) on (resp. the complement of) $G$ strictly after $S_n$, is conditionally admissible at time $S_n$ for $X$. The desired conclusion then follows through a general argument, see Theorem 4.6. Specifically, one finds that the family $\{\mathsf{E}[J(Y)|\mathcal{G}_{S_n}] : Y \in \mathbf{C}, Y^{S_n} = X^{S_n}\}$ is directed downwards for each $n \in \mathbb{N}_0$ (cf. proof of Proposition 4.2), hence can apply Lemma A.3 (cf. proofs of Proposition 4.5 and Theorem 4.6). ◇

In light of this example it is important to note that it will not matter to our general analysis, which controlled times are actually put into $\mathbf{G}$: *as long as* the explicit provisions that we (will, viz. Assumption 4.3) have made, are in fact met. This generality allows to work with/choose, in a given specific situation, such a family $\mathbf{G}$, as can be/is most informative of the problem.

(•) Finally, as already remarked, a typical example of an observed filtration is that of an observed process, i.e. for $c \in \mathbf{C}$, $X^c$ is a process whose values (in some measurable space) we can observe, and $\mathcal{G}^c_t := \sigma(X^c|_{[0,t]}) = \sigma(X^c_s : s \in [0,t])$, $t \in T$, is the natural filtration of $X^c$ {or possibly its $\mathsf{P}^c$-completion}. Let $c$ and $d$ be two controls, agreeing up to a controlled time $\mathcal{S}$, $c \sim_{\mathcal{S}} d$. Then, presumably, $X^c$ and $X^d$ do, also, i.e. $(X^c)^{\mathcal{S}^c} = (X^d)^{\mathcal{S}^d}$ {$\mathsf{P}^c$-a.s. and $\mathsf{P}^d$-a.s.} (where, *a priori*, $\mathcal{S}^c = \mathcal{S}^d$ {$\mathsf{P}^c$-a.s. and $\mathsf{P}^d$-a.s.}), and hence we should like to have (viz. Assumption 2.9) $\mathcal{G}^c_{\mathcal{S}^c} = \mathcal{G}^d_{\mathcal{S}^d}$. In other words, abstracting only slightly, and formulated without the unnecessary stochastic control-picture in the background, the following is a natural, and an extremely important, question. Suppose $X$ and $Y$ are two processes, defined on the same sample, and with values in the same measurable, space; $S$ a stopping time of both (or possibly just one) of their {completed} natural filtrations. Suppose furthermore the stopped processes agree, $X^S = Y^S$ {with probability one}. Must we have $\mathcal{F}^X_S = \mathcal{F}^Y_S$ $\{\overline{\mathcal{F}^X}_S = \overline{\mathcal{F}^Y}_S\}$ for the {completed} natural filtrations $\mathcal{F}^X$ and $\mathcal{F}^Y$ $\{\overline{\mathcal{F}^X}$ and $\overline{\mathcal{F}^Y}\}$ of $X$ and $Y$? Intuitively: yes, of course (at least when there are no completions in play). Formally, in the non-discrete case, it is not so straightforward. We obtain partial answers in Part 2.

We conclude this section with a rather general example illustrating the concepts introduced thusfar, focusing on the control-dependent informational flows, and with explicit references made to Settings 2.1 and 2.7.

*Example* 2.11. The time set is $[0,\infty)$ (Setting 2.1(i); $T=[0,\infty)$). Given are: a probability space $(\Omega,\mathcal{F},\mathsf{P})$ (Settings 2.1(iii) and 2.1(iv); $\mathcal{F}^c=\mathcal{F}$ and $\mathsf{P}^c=\mathsf{P}$ for each $c$, $\Omega$ is itself); an open subset $O\subset\mathbb{R}^d$ of Euclidean space; and a random (time-dependent) field $(R^o)_{o\in O}$ – each $R^o_t$ being an $\mathcal{F}$-measurable random variable, and the random map $((o,t)\mapsto R^o_t)$ being assumed continuous from $O\times[0,\infty)$ into $\mathbb{R}$ /so that the map $((\omega,o,t)\mapsto R^o_t(\omega))$ is automatically $\mathcal{F}\otimes\mathcal{B}(O)\otimes\mathcal{B}([0,\infty))/\mathcal{B}(\mathbb{R})$-measurable[3]/. Think of, for example, the local times of a Markov process, the Brownian sheet, solutions to SPDEs [3] etc.

Now, the idea is to control the movement in such a random field, observing only the values of the field at the current space-time point determined by the control $c$. Rewards accrue as a function of the value of the field at the location of the control, penalized is the speed of movement.

To make this formal, fix a discount factor $\alpha\in(0,\infty)$, an initial point $o_0\in O$, a measurable reward function $f:O\to\overline{\mathbb{R}}$ and a nondecreasing penalty function $g:[0,\infty)\to\overline{\mathbb{R}}$.

The controls (members of $\mathbf{C}$ of Setting 2.1(ii)) are then specified as being precisely all the $O$-valued, $\mathcal{F}\otimes\mathcal{B}([0,\infty))/\mathcal{B}(O)$-measurable, differentiable (from the right at zero), processes $c$, starting at $o_0$ (i.e. with $c_0=o_0$), adapted to the natural filtration (denoted $\mathcal{G}^c$; as in Setting 2.1(vi)) of the process $R^c$, and satisfying (cf. definition of $J$ in the paragraph following) $\mathsf{E}\int_0^\infty e^{-\alpha t}\left[f^-\circ R^{c_t}_t+g^+\circ|\dot{c}_t|\right]dt<\infty$. (Clearly the observed information $\mathcal{G}^c$ depends in a highly non-trivial way on the chosen control.)

Next, the payoff functional $J$ from Setting 2.1(v) is given as:

$$J(c):=\int_0^\infty e^{-\alpha t}\left[f\circ R^{c_t}_t-g\circ|\dot{c}_t|\right]dt,\quad c\in\mathbf{C}.$$

Finally, with regard to Setting 2.7, define for any $c\in\mathbf{C}$ and controlled time $\mathcal{S}$,

$$\mathcal{D}(c,\mathcal{S}):=\{d\in\mathbf{C}:d^{\mathcal{S}^c}=c^{\mathcal{S}^c}\},$$

and then let

$$\mathbf{G}:=\{\text{controlled times }\mathcal{S}\text{ such that }\forall c\forall d(d\in\mathcal{D}(c,\mathcal{S})\Rightarrow\mathcal{S}^c=\mathcal{S}^d\,\&\,\mathcal{G}^c_{\mathcal{S}^c}=\mathcal{G}^d_{\mathcal{S}^d})\}.$$

We will indeed see (Corollary 6.7) that $\mathbf{G}=\{$controlled times $\mathcal{S}$ such that $\forall c\forall d(d^{\mathcal{S}^c}=c^{\mathcal{S}^c}\Rightarrow\mathcal{S}^c=\mathcal{S}^d)\}$, as long as $(\Omega,\mathcal{F})$ is Blackwell (which can typically be taken to be the case). Regardless of whether or not $(\Omega,\mathcal{F})$ is in fact Blackwell, however, all the provisions of Settings 2.1 and 2.7, as well as those of Assumption 2.9, are in fact met. ◇

## 3. The conditional payoff and the Bellman system

**Definition 3.1** (Conditional payoff & Bellman system)**.** We define for $c\in\mathbf{C}$ and $\mathcal{S}\in\mathbf{G}$:

$$J(c,\mathcal{S}):=\mathsf{E}^{\mathsf{P}^c}[J(c)|\mathcal{G}^c_{\mathcal{S}^c}],\ \text{and then}\ V(c,\mathcal{S}):=\mathsf{P}^c|_{\mathcal{G}^c_{\mathcal{S}^c}}\text{-esssup}_{d\in\mathcal{D}(c,\mathcal{S})}J(d,\mathcal{S});$$

[3]For simplicity (so as not to be preoccupied with technical issues) we make all the processes in this example continuous.

and say $c \in \mathbf{C}$ is **conditionally optimal** at $\mathcal{S} \in \mathbf{G}$, if $V(c,\mathcal{S}) = J(c,\mathcal{S})$ $\mathsf{P}^c$-a.s. $(J(c,\mathcal{S}))_{(c,\mathcal{S})\in\mathbf{C}\times\mathbf{G}}$ is called the **conditional payoff system** and $(V(c,\mathcal{S}))_{(c,\mathcal{S})\in\mathbf{C}\times\mathbf{G}}$ the **Bellman system**.

*Remark* 3.2.

(i) Thanks to Assumption 2.9, the essential suprema appearing in the definition of the conditional payoff system are well-defined (up to the relevant a.s. equalities).
(ii) Also, thanks to Setting 2.7(3), $V(c,\mathcal{S})$ only depends on $\mathcal{S}$ through $\mathcal{S}^c$, in the sense that $V(c,\mathcal{S}) = V(c,\mathcal{T})$ as soon as $\mathcal{S}^c = \mathcal{T}^c$ $\{\mathsf{P}^c$-a.s.$\}$. Clearly the same holds true (trivially) of the system $J$.

Some further properties of the systems $V$ and $J$ follow. First,

**Proposition 3.3.** *$V(c,\mathcal{S})$ is $\mathcal{G}^c_{\mathcal{S}^c}$-measurable and its negative part is $\mathsf{P}^c$-integrable for each $(c,\mathcal{S}) \in \mathbf{C}\times\mathbf{G}$. Moreover if $c \sim_{\mathcal{S}} d$, then $V(c,\mathcal{S}) = V(d,\mathcal{S})$ $\mathsf{P}^c$-a.s. and $\mathsf{P}^d$-a.s.*

*Proof.* The appropriate measurability of $V(c,\mathcal{S})$ follows from its definition. Moreover, since each $\mathcal{D}(c,\mathcal{S})$ is non-empty, the integrability condition on the negative parts of $V$ is also immediate (from the assumed integrability of the negative parts of $J$). Finally, the last claim follows from the fact that $\mathcal{D}(c,\mathcal{S}) = \mathcal{D}(d,\mathcal{S})$ (partitioning property) and $\mathsf{P}^c|_{\mathcal{G}^c_{\mathcal{S}^c}} = \mathsf{P}^d|_{\mathcal{G}^d_{\mathcal{S}^d}}$ (consistency), when $c \sim_{\mathcal{S}} d$. □

Second, Proposition 3.9, will (i) establish that in fact $(J(c,\mathcal{S}))_{(c,\mathcal{S})\in\mathbf{C}\times\mathbf{G}}$ is a $(\mathbf{C},\mathbf{G})$-system in the sense of the definition which follows, and (ii) will also give sufficient conditions for the $\mathsf{P}^c$-a.s. equality $J(c,\mathcal{S}) = J(d,\mathcal{S})$ to obtain on an event $A \in \mathcal{G}^c_{\mathcal{S}^c}$, when $c \sim_{\mathcal{S}} d$ (addressing the situation when the two controls $c$ and $d$ agree “for all times” on $A$). Some auxiliary definitions and results are needed to this end; they precede Proposition 3.9.

**Definition 3.4** ($(\mathbf{C},\mathbf{G})$-system)**.** A collection $X = (X(c,\mathcal{T}))_{(c,\mathcal{T})\in\mathbf{C}\times\mathbf{G}}$ of functions from $[-\infty,+\infty]^\Omega$ is a **$(\mathbf{C},\mathbf{G})$-system**, if (i) $X(c,\mathcal{T})$ is $\mathcal{G}^c_{\mathcal{T}^c}$-measurable for all $(c,\mathcal{T}) \in \mathbf{C}\times\mathbf{G}$ and (ii) $X(c,\mathcal{S}) = X(c,\mathcal{T})$ $\mathsf{P}^c$-a.s. on the event $\{\mathcal{S}^c = \mathcal{T}^c\}$, for all $c \in \mathbf{C}$ and $\{\mathcal{S},\mathcal{T}\} \subset \mathbf{G}$.

**Definition 3.5** (Times accessing infinity)**.** For a sequence $(t_n)_{n\in\mathbb{N}}$ of elements of $T$, we say it **accesses infinity**, if for all $t \in T$, there exists an $n \in \mathbb{N}$ with $t \leq t_n$.

**Lemma 3.6.** *Suppose $\mathcal{H}$ is a $\{\mathsf{P}$-complete$\}$ filtration on $\Omega$ $\{\mathsf{P}$ being a complete probability measure$\}$ and $(S_n)_{n\geq 1}$ a sequence of its stopping times $\{$each defined up to $\mathsf{P}$-a.s. equality$\}$, which accesses infinity pointwise $\{$or $\mathsf{P}$-a.s.$\}$ on $A \subset \Omega$, i.e. $(S_n(\omega))_{n\in\mathbb{N}}$ accesses infinity for $\{\mathsf{P}$-almost$\}$ every $\omega \in A$. Then $\mathcal{H}_\infty|_A = \vee_{n\in\mathbb{N}}\mathcal{H}_{S_n}|_A$.*

*Proof.* The inclusion $\supset$ is manifest. For the reverse inclusion, let $t \in T$ and $B \in \mathcal{H}_t$ (noting that for any $\mathcal{L} \subset 2^\Omega$ and $A \subset \Omega$, $\sigma_\Omega(\mathcal{L})|_A = \sigma_A(\mathcal{L}|_A)$). Then $\{\mathsf{P}$-a.s.$\}$ $B \cap A = \cup_{n=1}^\infty (B \cap \{S_n \geq t\}) \cap A$ with $B \cap \{S_n \geq t\} \in \mathcal{H}_{S_n}$. □

**Lemma 3.7.** *Let $\{c,d\}\subset \mathbf{C}$ and $A\subset\Omega$. Let furthermore $(\mathcal{S}_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbf{G}$ accessing infinity $\{$a.s.$\}$ on $A$ for the controls $c$ and $d$ (i.e. $(\mathcal{S}_n^h(\omega))_{n\in\mathbb{N}}$ accesses infinity for $\{\mathsf{P}^h$-almost$\}$ every $\omega\in A$, each $h\in\{c,d\}$), and for which $c\sim_{\mathcal{S}_n} d$ for each $n\in\mathbb{N}$. Then $\mathcal{G}^c_\infty|_A=\mathcal{G}^d_\infty|_A$.*

*If further, $A\in\mathcal{G}^c_{\mathcal{S}^c_n}$ for all $n\in\mathbb{N}$, and the sequence $(\mathcal{S}_n)_{n\in\mathbb{N}}$ is $\{$a.s.$\}$ nondecreasing on $A$ and for the controls $c$ and $d$ (i.e. $(\mathcal{S}_n^h(\omega))_{n\in\mathbb{N}}$ is nondecreasing for $\{\mathsf{P}^h$-almost$\}$ every $\omega\in A$, each $h\in\{c,d\}$), then $\mathsf{P}^c|_{\mathcal{G}^c_\infty}$ and $\mathsf{P}^d|_{\mathcal{G}^d_\infty}$ agree when traced on $A$.*

*Remark* 3.8. We mean to address here abstractly the situation when the two controls $c$ and $d$ agree for all times on $A$.

*Proof.* By the consistency properties, certainly $\mathsf{P}^c|_{\mathcal{G}^c_{\mathcal{S}^c_n}}$ agrees with $\mathsf{P}^d|_{\mathcal{G}^d_{\mathcal{S}^d_n}}$ for each $n\in\mathbb{N}$, while $(\mathcal{S}^c_n=\mathcal{S}^d_n)_{n\in\mathbb{N}}$ accesses infinity $\{\mathsf{P}^c$-a.s. and $\mathsf{P}^d$-a.s.$\}$ on $A$. Then apply Lemma 3.6 to obtain $\mathcal{G}^c_\infty|_A=\sigma_A(\cup_{n\in\mathbb{N}}\mathcal{G}^c_{\mathcal{S}^c_n}|_A)=\sigma_A(\cup_{n\in\mathbb{N}}\mathcal{G}^d_{\mathcal{S}^d_n}|_A)=\mathcal{G}^d_\infty|_A$. If, moreover $A\in\mathcal{G}^c_{\mathcal{S}^c_n}$ for all $n\in\mathbb{N}$, then the traces of $\mathsf{P}^c$ and $\mathsf{P}^d$ on $A$ agree on $\cup_{n\in\mathbb{N}}\mathcal{G}^c_{\mathcal{S}^c_n}|_A$. Provided in addition $(\mathcal{S}^c_n)_{n\in\mathbb{N}}$ is $\{\mathsf{P}^c$-a.s.$\}$ nondecreasing on $A$, the latter union is a $\pi$-system (as a nondecreasing union of $\sigma$-fields, so even an algebra) on $A$. This, coupled with the fact that two finite measures of the same mass, which agree on a generating $\pi$-system, agree (by a monotone class argument), yields the second claim. □

**Proposition 3.9.** *$(J(c,\mathcal{S}))_{(c,\mathcal{S})\in\mathbf{C}\times\mathbf{G}}$ is a $(\mathbf{C},\mathbf{G})$-system. Moreover, if*

*(i) $c\sim_{\mathcal{S}} d$, $A\in\mathcal{G}^c_{\mathcal{S}^c}$;*
*(ii) there exists a sequence $(\mathcal{S}_n)_{n\in\mathbb{N}}$ from $\mathcal{G}$ $\{$a.s.$\}$ nondecreasing and accessing infinity on $A$ for the controls $c$ and $d$, and for which $c\sim_{\mathcal{S}_n} d$ and $A\in\mathcal{G}^c_{\mathcal{S}^c_n}$ for each $n\in\mathbb{N}$;*
*(iii) $\mathsf{E}^{\mathsf{P}^c}[J(c)|\mathcal{G}^c_\infty]=\mathsf{E}^{\mathsf{P}^d}[J(d)|\mathcal{G}^d_\infty]$ $\mathsf{P}^c$-a.s. and $\mathsf{P}^d$-a.s. on $A$;*

*then $J(c,\mathcal{S})=J(d,\mathcal{S})$ $\mathsf{P}^c$-a.s. and $\mathsf{P}^d$-a.s. on $A$.*

*Proof.* By definition, $J(c,\mathcal{T})$ is $\mathcal{G}^c_{\mathcal{T}^c}$-measurable. Next, if $c\in\mathbf{C}$, then $\mathcal{G}^c_{\mathcal{S}^c}=\mathcal{G}^c_{\mathcal{T}^c}$ when traced on $\{\mathcal{T}^c=\mathcal{S}^c\}\in\mathcal{G}^c_{\mathcal{S}^c}\cap\mathcal{G}^c_{\mathcal{T}^c}$, whence $J(c,\mathcal{T})=J(c,\mathcal{S})$ $\mathsf{P}^c$-a.s. thereon (applying Lemma A.1). Finally, to show that $J(c,\mathcal{S})=J(d,\mathcal{S})$ $\mathsf{P}^c$-a.s. (or $\mathsf{P}^d$-a.s., it is the same) on $A$ under the indicated conditions, we need only establish that:

$$\begin{aligned}
\mathsf{P}^c\text{-a.s. } \mathbb{1}_A\mathsf{E}^{\mathsf{P}^c}[J(c)|\mathcal{G}^c_{\mathcal{S}^c}] &= \mathbb{1}_A\mathsf{E}^{\mathsf{P}^d}[J(d)|\mathcal{G}^d_{\mathcal{S}^d}] \Leftrightarrow (\text{since } \mathcal{G}^d_{\mathcal{S}^d}=\mathcal{G}^c_{\mathcal{S}^c},\ A\in\mathcal{G}^c_{\mathcal{S}^c})\\
\forall B\in\mathcal{G}^c_{\mathcal{S}^c}\quad \mathsf{E}^{\mathsf{P}^c}[J(c)\mathbb{1}_A\mathbb{1}_B] &= \mathsf{E}^{\mathsf{P}^c}[\mathsf{E}^{\mathsf{P}^d}[J(d)\mathbb{1}_A\mathbb{1}_B|\mathcal{G}^d_{\mathcal{S}^d}]] \Leftrightarrow (\text{since } \mathsf{P}^c|_{\mathcal{G}^c_{\mathcal{S}^c}}=\mathsf{P}^d|_{\mathcal{G}^d_{\mathcal{S}^d}})\\
\forall B\in\mathcal{G}^c_{\mathcal{S}^c}\quad \mathsf{E}^{\mathsf{P}^c}[J(c)\mathbb{1}_A\mathbb{1}_B] &= \mathsf{E}^{\mathsf{P}^d}[J(d)\mathbb{1}_A\mathbb{1}_B] \Leftrightarrow (\text{conditioning})\\
\forall B\in\mathcal{G}^c_{\mathcal{S}^c}\quad \mathsf{E}^{\mathsf{P}^c}[\mathsf{E}^{\mathsf{P}^c}[J(c)|\mathcal{G}^c_\infty]\mathbb{1}_A\mathbb{1}_B] &= \mathsf{E}^{\mathsf{P}^d}[\mathsf{E}^{\mathsf{P}^d}[J(d)|\mathcal{G}^d_\infty]\mathbb{1}_A\mathbb{1}_B] \Leftrightarrow\\
&\quad (\text{since } \mathsf{E}^{\mathsf{P}^c}[J(c)|\mathcal{G}^c_\infty]=\mathsf{E}^{\mathsf{P}^d}[J(d)|\mathcal{G}^d_\infty]\ \mathsf{P}^c\text{-a.s. on } A)\\
\forall B\in\mathcal{G}^c_{\mathcal{S}^c}\quad \mathsf{E}^{\mathsf{P}^c}[\mathsf{E}^{\mathsf{P}^d}[J(d)|\mathcal{G}^d_\infty]\mathbb{1}_A\mathbb{1}_B] &= \mathsf{E}^{\mathsf{P}^d}[\mathsf{E}^{\mathsf{P}^d}[J(d)|\mathcal{G}^d_\infty]\mathbb{1}_A\mathbb{1}_B],
\end{aligned}$$

where finally one can apply Lemma 3.7. □

## 4. Bellman's principle

**Definition 4.1** ($(\mathbf{C},\mathbf{G})$-super/-/sub-martingale systems)**.** A collection $X = (X(c,\mathcal{S}))_{(c,\mathcal{S})\in(\mathbf{C},\mathbf{G})}$ of functions from $[-\infty,+\infty]^\Omega$ is a $(\mathbf{C},\mathbf{G})$- (resp. **super-**, **sub-**) **martingale system**, if for each $(c,\mathcal{S}) \in \mathbf{C}\times\mathbf{G}$ (i) $X(c,\mathcal{S})$ is $\mathcal{G}^c_{\mathcal{S}^c}$-measurable, (ii) $X(c,\mathcal{S}) = X(d,\mathcal{S})$ $\mathsf{P}^c$-a.s. and $\mathsf{P}^d$-a.s., whenever $c \sim_{\mathcal{S}} d$, (iii) (resp. the negative, positive part of) $X(c,\mathcal{S})$ is integrable and (iv) for all $\{\mathcal{S},\mathcal{T}\} \subset \mathbf{G}$ and $c \in \mathbf{C}$ with $\mathcal{S}^d \leq \mathcal{T}^d$ $\{\mathsf{P}^d$-a.s.$\}$ for $d \in \mathcal{D}(c,\mathcal{T})$,

$$\mathsf{E}^{\mathsf{P}^c}[X(c,\mathcal{T})|\mathcal{G}^c_{\mathcal{S}^c}] = X(c,\mathcal{S})$$

(resp. $\mathsf{E}^{\mathsf{P}^c}[X(c,\mathcal{T})|\mathcal{G}^c_{\mathcal{S}^c}] \leq X(c,\mathcal{S})$, $\mathsf{E}^{\mathsf{P}^c}[X(c,\mathcal{T})|\mathcal{G}^c_{\mathcal{S}^c}] \geq X(c,\mathcal{S})$) $\mathsf{P}^c$-a.s.

In order to be able to conclude the supermartingale property of the Bellman system (Bellman's principle), we shall need to make a further assumption (see Assumption 4.3 below; cf. Lemma A.3). The following proposition gives some guidance as to when it may be valid.

**Proposition 4.2.** *Let $c \in \mathbf{C}$, $\mathcal{S} \in \mathbf{G}$ and $\epsilon \in [0,\infty)$, $M \in (0,\infty]$. Then (1)⇒(2)⇒(3).*

*(1) (i) For all $d \in \mathcal{D}(c,\mathcal{S})$, $\mathsf{P}^d = \mathsf{P}^c$ . AND*
*(ii) For all $\{d,d'\} \subset \mathcal{D}(c,\mathcal{S})$ and $G \in \mathcal{G}^c_{\mathcal{S}^c}$, there is a $d'' \in \mathcal{D}(c,\mathcal{S})$ such that $J(d'') \geq M \wedge [\mathbb{1}_G J(d) + \mathbb{1}_{\Omega\backslash G} J(d')] - \epsilon$ $\mathsf{P}^c$-a.s.*

*(2) For all $\{d,d'\} \subset \mathcal{D}(c,\mathcal{S})$ and $G \in \mathcal{G}^c_{\mathcal{S}^c}$, there is a $d'' \in \mathcal{D}(c,\mathcal{S})$ such that $J(d'',\mathcal{S}) \geq M \wedge [\mathbb{1}_G J(d,\mathcal{S}) + \mathbb{1}_{\Omega\backslash G} J(d',\mathcal{S})] - \epsilon$ $\mathsf{P}^c$-a.s.*

*(3) $(J(d,\mathcal{S}))_{d\in\mathcal{D}(c,\mathcal{S})}$ has the "$(\epsilon,M)$-upwards lattice property":*
*For all $\{d,d'\} \subset \mathcal{D}(c,\mathcal{S})$ there exists a $d'' \in \mathcal{D}(c,\mathcal{S})$ such that*

$$J(d'',\mathcal{S}) \geq (M \wedge J(d,\mathcal{S})) \vee (M \wedge J(d',\mathcal{S})) - \epsilon$$

*$\mathsf{P}^c$-a.s.*

*Proof.* Implication (1)⇒(2) follows by conditioning on $\mathcal{G}^c_{\mathcal{S}^c}$ under $\mathsf{P}^c$. Implication (2)⇒(3) follows by taking $G = \{J(d,\mathcal{S}) > J(d',\mathcal{S})\} \in \mathcal{G}^c_{\mathcal{S}^c}$. □

**Assumption 4.3** (Upwards lattice property)**.** *For all $c \in \mathbf{C}$, $\mathcal{S} \in \mathbf{G}$ and $\{\epsilon, M\} \subset (0,\infty)$, $(J(d,\mathcal{S}))_{d\in\mathcal{D}(c,\mathcal{S})}$ enjoys the $(\epsilon,M)$-upwards lattice property (as in Proposition 4.2, Property (3)).*

(We shall make it explicit in the sequel when this assumption will be in effect.)

*Remark* 4.4.

(i) The upwards lattice property of Assumption 4.3 represents a direct connection between the set of controls $\mathbf{C}$ on the one hand and the collection of observable filtrations $(\mathcal{G}^c)_{c\in\mathbf{C}}$ and set of controlled times $\mathbf{G}$ on the other. It is weaker than insisting that every system $(J(c,\mathcal{S}))_{c\in\mathbf{C}}$ be upwards-directed (Proposition 4.2, Property (3) with $\epsilon = 0$, $M = \infty$), but still sufficient to allow one to conclude Bellman's (super)martingale principle (Theorem 4.6). A more precise investigation into the relationship between the validity of Bellman's principle, and the linking between $\mathbf{C}$, $\mathbf{G}$ and the collection $(\mathcal{G}^c)_{c\in\mathbf{C}}$ remains open to future research.

(ii) It may be assumed without loss of generality (in the precise sense which follows) that $\{0,\infty\}\subset\mathbf{G}$. Specifically, we can always simply extend the family $\mathcal{D}$, by defining $\mathcal{D}(c,\infty):=\{c\}$ and $\mathcal{D}(c,0):=\mathbf{C}$ for each $c\in\mathbf{C}$ – none of the provisions of Section 2.1 (Setting 2.1 and 2.7, Assumption 2.9), nor indeed the validity or non-validity of Assumption 4.3, being thus affected.

(iii) Assumption 4.3 is of course trivially verified when the filtrations $\mathcal{G}$ all consist of (probabilistically) trivial $\sigma$-fields alone.

*Example 2.11 continued.* We verify that in Example 2.11, under the assumption that in fact the base $(\Omega,\mathcal{F})$ therein is Blackwell, Property (1) from Proposition 4.2 obtains with $M=\infty$, $\epsilon=0$.

Let $G\in\mathcal{G}^c_{\mathcal{S}^c}$, $\{d,d'\}\subset\mathcal{D}(c,\mathcal{S})$. It will suffice to show that $d'':=d\mathbb{1}_G+d'\mathbb{1}_{\Omega\backslash G}\in\mathbf{C}$ (for, then, in fact, we will have $J(d'')=J(d)\mathbb{1}_G+J(d')\mathbb{1}_{\Omega\backslash G}$). Now, $\mathcal{S}^d=\mathcal{S}^c=\mathcal{S}^{d'}$ are all stopping times of $\mathcal{G}^c$, and $d^{\mathcal{S}^c}=c^{\mathcal{S}^c}=d'^{\mathcal{S}^c}$; by Theorem 6.6, Proposition 6.5, and Proposition 6.9 to follow in Part 2, all the events $\{S^c>t\}$, $\{\mathcal{S}^c\leq t\}\cap G$ and $\{\mathcal{S}^c\leq t\}\cap(\Omega\backslash G)$ belong to $\sigma((R^c)^{\mathcal{S}^c\wedge t})=\sigma((R^{d''})^{\mathcal{S}^c\wedge t})\subset\sigma((R^{d''})^t)=\mathcal{G}^{d''}_t$. Next, for sure, $d''$ is a $\mathcal{F}\otimes\mathcal{B}([0,\infty))/\mathcal{B}(O)$-measurable, differentiable $O$-valued process with initial value $o_0$, satisfying the requisite integrability condition on $f^-$ and $g^+$. So it remains to check $d''$ is $\mathcal{G}^{d''}$-adapted; let $t\in[0,\infty)$. Then $d''_t\mathbb{1}_{\{\mathcal{S}^c>t\}}=c_t\mathbb{1}_{\{\mathcal{S}^c>t\}}\in\mathcal{G}^c_t$, hence $d''_t\mathbb{1}_{\{\mathcal{S}^c>t\}}\in\mathcal{G}^{d''}_t$, since $\mathcal{G}^{d''}_t=\mathcal{G}^c_t$ on $\{\mathcal{S}^c>t\}$. On the other hand, $d''_t\mathbb{1}_{\{\mathcal{S}^c\leq t\}\cap G}=d_t\mathbb{1}_{\{\mathcal{S}^c\leq t\}\cap G}\in\mathcal{G}^d_t$, hence $d''_t\mathbb{1}_{\{\mathcal{S}^c\leq t\}\cap G}\in\mathcal{G}^{d''}_t$, since $\mathcal{G}^{d''}_t=\mathcal{G}^d_t$ on $\{\mathcal{S}^c\leq t\}\cap G$; similarly for $\Omega\backslash G$ in place of $G$. ⋄

**Proposition 4.5.** *[Cf.* [5, p. 94, Lemma 1.14]*.] Under Assumption 4.3, for any $c\in\mathbf{C}$, $\mathcal{T}\in\mathbf{G}$ and any sub-$\sigma$-field $\mathcal{A}$ of $\mathcal{G}^c_{\mathcal{T}^c}$, $\mathsf{P}^c$-a.s.:*

$$\mathsf{E}^{\mathsf{P}^c}[V(c,\mathcal{T})|\mathcal{A}]=\mathsf{P}^c|_{\mathcal{A}}\text{-esssup}_{d\in\mathcal{D}(c,\mathcal{T})}\mathsf{E}^{\mathsf{P}^d}[J(d)|\mathcal{A}].$$

*In particular, $\mathsf{E}^{\mathsf{P}^c}V(c,\mathcal{T})=\sup_{d\in\mathcal{D}(c,\mathcal{T})}\mathsf{E}^{\mathsf{P}^d}J(d)$.*

*Proof.* By Lemma A.3, we have, $\mathsf{P}^c$-a.s.:

$$\begin{aligned}\mathsf{E}^{\mathsf{P}^c}[V(c,\mathcal{T})|\mathcal{A}]&=\mathsf{P}^c|_{\mathcal{A}}\text{-esssup}_{d\in\mathcal{D}(c,\mathcal{T})}\mathsf{E}^{\mathsf{P}^c}[\mathsf{E}^{\mathsf{P}^d}[J(d)|\mathcal{G}^d_{\mathcal{T}^d}]|\mathcal{A}]\\&=\mathsf{P}^c|_{\mathcal{A}}\text{-esssup}_{d\in\mathcal{D}(c,\mathcal{T})}\mathsf{E}^{\mathsf{P}^d}[\mathsf{E}^{\mathsf{P}^d}[J(d)|\mathcal{G}^c_{\mathcal{T}^c}]|\mathcal{A}],\ \text{since } \mathcal{G}^c_{\mathcal{T}^c}=\mathcal{G}^d_{\mathcal{T}^d}\ \&\ \mathsf{P}^c|_{\mathcal{G}^c_{\mathcal{T}^c}}=\mathsf{P}^d|_{\mathcal{G}^d_{\mathcal{T}^d}},\end{aligned}$$

for $d\sim_{\mathcal{T}}c$, where from the claim follows at once. □

**Theorem 4.6** (Bellman's principle)**.** *We work under the provisions of Assumption 4.3 and insist $\{0,\infty\}\subset\mathbf{G}$ (recall Remark 4.4(ii)).*

*$(V(c,\mathcal{S}))_{(c,\mathcal{S})\in\mathbf{C}\times\mathbf{G}}$ is a $(\mathbf{C},\mathbf{G})$-supermartingale system. Moreover, if $c^*\in\mathbf{C}$ is optimal, then $(V(c^*,\mathcal{T}))_{\mathcal{T}\in\mathbf{G}}$ has a constant $\mathsf{P}^{c^*}$-expectation (equal to the optimal value $v=\mathsf{E}^{\mathsf{P}^{c^*}}J(c^*)$). If further $\mathsf{E}^{\mathsf{P}^{c^*}}J(c^*)<\infty$, then $(V(c^*,\mathcal{T}))_{\mathcal{T}\in\mathbf{G}}$ is a $\mathbf{G}$-martingale in the sense that (i) for each $\mathcal{T}\in\mathbf{G}$,*

*$V(c^*,\mathcal{T})$ is $\mathcal{G}^{c^*}_{\mathcal{T}^{c^*}}$-measurable and $\mathsf{P}^{c^*}$-integrable and (ii) for any $\{\mathcal{S},\mathcal{T}\}\subset\mathbf{G}$ with $\mathcal{S}^d\leq\mathcal{T}^d$ $\{\mathsf{P}^d$-a.s.$\}$ for $d\in\mathcal{D}(c^*,\mathcal{T})$, $\mathsf{P}^{c^*}$-a.s.,*

$$\mathsf{E}^{\mathsf{P}^{c^*}}[V(c^*,\mathcal{T})|\mathcal{G}^{c^*}_{\mathcal{S}^{c^*}}]=V(c^*,\mathcal{S}).$$

*Furthermore, if $c^*\in\mathbf{C}$ is conditionally optimal at $\mathcal{S}\in\mathbf{G}$ and $\mathsf{E}^{\mathsf{P}^{c^*}}J(c^*)<\infty$, then $c^*$ is conditionally optimal at $\mathcal{T}$ for any $\mathcal{T}\in\mathbf{G}$ satisfying $\mathcal{T}^d\geq\mathcal{S}^d$ $\{\mathsf{P}^d$-a.s.$\}$ for $d\in\mathcal{D}(c^*,\mathcal{T})$. In particular, if $c^*$ is optimal, then it is conditionally optimal at $0$, so that if further $\mathsf{E}^{\mathsf{P}^{c^*}}J(c^*)<\infty$, then $c^*$ must be conditionally optimal at any $\mathcal{S}\in\mathbf{G}$.*

*Conversely, and regardless of whether Assumption 4.3 holds true, if $\mathbf{G}$ includes a sequence $(\mathcal{S}_n)_{n\in\mathbb{N}_0}$ for which (i) $\mathcal{S}_0=0$, (ii) the family $(V(c^*,\mathcal{S}_n))_{n\geq0}$ has a constant $\mathsf{P}^{c^*}$-expectation and is uniformly integrable, and (iii) $V(c^*,\mathcal{S}_n)\to V(c^*,\infty)$, $\mathsf{P}^{c^*}$-a.s. (or even just in $\mathsf{P}^{c^*}$-probability), as $n\to\infty$, then $c^*$ is optimal.*

*Proof.* Let $\{\mathcal{S},\mathcal{T}\}\subset\mathbf{G}$ and $c\in\mathbf{C}$ with $\mathcal{S}^d\leq\mathcal{T}^d$ $\{\mathsf{P}^d$-a.s.$\}$ for $d\in\mathcal{D}(c,\mathcal{T})$. Then, since $\mathcal{S}^c\leq\mathcal{T}^c$ $\{\mathsf{P}^c$-a.s.$\}$, $\mathcal{G}^c_{\mathcal{S}^c}\subset\mathcal{G}^c_{\mathcal{T}^c}$, and $\mathcal{D}(c,\mathcal{T})\subset\mathcal{D}(c,\mathcal{S})$, so that we obtain via Proposition 4.5, $\mathsf{P}^c$-a.s.,

$$\begin{aligned}\mathsf{E}^{\mathsf{P}^c}[V(c,\mathcal{T})|\mathcal{G}^c_{\mathcal{S}^c}] &= \mathsf{P}^c|_{\mathcal{G}^c_{\mathcal{S}^c}}\text{-esssup}_{d\in\mathcal{D}(c,\mathcal{T})}\mathsf{E}^{\mathsf{P}^d}[J(d)|\mathcal{G}^c_{\mathcal{S}^c}]\\ &\leq \mathsf{P}^c|_{\mathcal{G}^c_{\mathcal{S}^c}}\text{-esssup}_{d\in\mathcal{D}(c,\mathcal{S})}\mathsf{E}^{\mathsf{P}^d}[J(d)|\mathcal{G}^c_{\mathcal{S}^c}]=V(c,\mathcal{S}),\end{aligned}$$

since $\mathcal{G}^c_{\mathcal{S}^c}=\mathcal{G}^d_{\mathcal{S}^d}$ for $s\sim_{\mathcal{S}}d$, which establishes the first claim. The second follows at once from the final conclusion of Proposition 4.5. Then let $c^*$ be optimal and $\{\mathcal{S},\mathcal{T}\}\subset\mathbf{G}$ with $\mathcal{S}^d\leq\mathcal{T}^d$ $\{\mathsf{P}^d$-a.s.$\}$ for $d\in\mathcal{D}(c^*,\mathcal{T})$. Note that by the supermartingale property, $v=\mathsf{E}^{\mathsf{P}^{c^*}}\mathsf{E}^{\mathsf{P}^{c^*}}[V(c^*,\mathcal{T})|\mathcal{G}^{c^*}_{\mathcal{S}^{c^*}}]\leq\mathsf{E}^{\mathsf{P}^{c^*}}V(c^*,\mathcal{S})=v$. So, if furthermore $v<\infty$ (remark that then $v\in\mathbb{R}$), we conclude that $V(c^*,\mathcal{T})$ is $\mathsf{P}^{c^*}$-integrable, and the martingale property also follows.

Next, if $c^*$ is conditionally optimal at $\mathcal{S}$, $\mathsf{E}^{\mathsf{P}^{c^*}}J(c^*)<\infty$, and $\mathcal{S}^d\leq\mathcal{T}^d$ $\{\mathsf{P}^d$-a.s.$\}$ for $d\in\mathcal{D}(c^*,\mathcal{T})$, then since $V$ is a $(\mathbf{C},\mathbf{G})$-supermartingale system, $\mathsf{E}^{\mathsf{P}^{c^*}}J(c^*)=\mathsf{E}^{\mathsf{P}^{c^*}}J(c^*,\mathcal{S})=\mathsf{E}^{\mathsf{P}^{c^*}}V(c^*,\mathcal{S})\geq\mathsf{E}^{\mathsf{P}^{c^*}}V(c^*,\mathcal{T})$. On the other hand, for sure, $V(c^*,\mathcal{T})\geq J(c^*,\mathcal{T})$, $\mathsf{P}^{c^*}$-a.s., so $\mathsf{E}^{\mathsf{P}^{c^*}}V(c^*,\mathcal{T})\geq\mathsf{E}^{\mathsf{P}^{c^*}}J(c^*,\mathcal{T})=\mathsf{E}^{\mathsf{P}^{c^*}}J(c^*)$ hence we must have $V(c^*,\mathcal{T})=J(c^*,\mathcal{T})$, $\mathsf{P}^{c^*}$-a.s., i.e. $c^*$ is conditionally optimal at $\mathcal{T}$. The penultimate claim is then also evident.

For the final claim notice that $V(c^*,\mathcal{S}_n)\to V(c^*,\infty)$ in $L^1(\mathsf{P}^{c^*})$, as $n\to\infty$, and so $v=\sup_{c\in\mathbf{C}}\mathsf{E}^{\mathsf{P}^c}J(c)=\mathsf{E}^{\mathsf{P}^{c^*}}V(c^*,0)=\mathsf{E}^{\mathsf{P}^{c^*}}V(c^*,\mathcal{S}_n)\to\mathsf{E}^{\mathsf{P}^{c^*}}V(c^*,\infty)=\mathsf{E}^{\mathsf{P}^{c^*}}J(c^*)$, as $n\to\infty$. □

**Proposition 4.7.** *Under Assumption 4.3 and insisting that $\infty\in\mathbf{G}$, $V$ is the minimal $(\mathbf{C},\mathbf{G})$-supermartingale system $W$ satisfying the terminal condition*

$$W(c,\infty)\geq\mathsf{E}^{\mathsf{P}^c}[J(c)|\mathcal{G}^c_\infty],\quad \mathsf{P}^c\text{-a.s. for each } c\in\mathbf{C}.$$

*Proof.* That $V$ is a $(\mathbf{C},\mathbf{G})$-supermartingale system satisfying the indicated terminal condition is clear from the definition of $V$ and Theorem 4.6. Next, let $W$ be a $(\mathbf{C},\mathbf{G})$-supermartingale system satisfying said terminal condition. Then for all $(c,\mathcal{T})\in\mathbf{C}\times\mathbf{G}$ and $d\in\mathcal{D}(c,\mathcal{T})$, $\mathsf{P}^c$-a.s. *and* $\mathsf{P}^d$-a.s.

$W(c,\mathcal{T}) = W(d,\mathcal{T}) \geq \mathsf{E}^{\mathsf{P}^d}[W(d,\infty)|\mathcal{G}^d_{\mathcal{T}^d}] \geq \mathsf{E}^{\mathsf{P}^d}[\mathsf{E}^{\mathsf{P}^d}[J(d)|\mathcal{G}^d_\infty]|\mathcal{G}^d_{\mathcal{T}^d}] = J(d,\mathcal{T})$. Thus $W(c,\mathcal{T}) \geq V(c,\mathcal{T})$, $\mathsf{P}^c$-a.s. $\square$

## 5. A solved formal example

Recall the notation of Section 2. The time set will be $[0,\infty)$ (Setting 2.1(i); $T = [0,\infty)$).

Fix next a discount factor $\alpha \in (0,\infty)$, let $(\Omega,\mathcal{H},\mathsf{P})$ be a probability space supporting two independent, sample-path-continuous, Brownian motions $B^0 = (B^0_t)_{t\in[0,\infty)}$ and $B^1 = (B^1_t)_{t\in[0,\infty)}$, starting at 0 and $-x \in \mathbb{R}$, respectively (Setting 2.1(iii) and 2.1(iv); $\mathcal{F}^c = \mathcal{H}$, $\mathsf{P}^c = \mathsf{P}$ for all $c$, $\Omega$ is itself). We may assume $(\Omega,\mathcal{H})$ is Blackwell. By $\mathcal{F}$ denote the natural filtration of the bivariate process $(B^0,B^1)$. Then for each càdlàg, $\mathcal{F}$-adapted, $\{0,1\}$-valued process $c$, let $\mathcal{G}^c$ be the natural filtration of $B^c$, the observed process (Setting 2.1(vi); the $\mathcal{G}^c$s are themselves); let $(J^c_k)_{k=0}^\infty$ be the jump times of $c$ (with $J^c_0 := 0$; $J^c_k = \infty$, if $c$ has less than $k$ jumps); and define:

$$\begin{aligned}\mathbf{C} \;:=\; & \bigcup_{\epsilon>0}\Big\{\mathcal{F}\text{-adapted, càdlàg, }\{0,1\}\text{-valued, processes } c, \text{ with } c_0 = 0, \\ & \quad \text{that are } \mathcal{G}^c\text{-predictable and such that } J^c_{k+1} - J^c_k \geq \epsilon \text{ on } \{J^c_k < \infty\} \text{ for all } k \in \mathbb{N}\Big\}\end{aligned}$$

(Setting 2.1(ii); $\mathbf{C}$ is itself). The insistence on the "$\epsilon$-separation" of the jumps of the controls appears artificial – our intention is to emphasize the salient features of the control-dependent informational flow, not to be preoccupied with the technical details.

For $c \in \mathbf{C}$, define next:

- for each $t \in [0,\infty)$ (with the convention $\sup\emptyset := 0$) $\sigma^c_t := \sup\{s \in [0,t] : c_s \neq c_t\}$ and $\tau^c_t := t - \sigma^c_t$, the last jump time of $c$ before time $t$ and the lag since then, respectively;
- $Z^c := B^c - B^{1-c}_{\sigma^c}$, the current distance of the observed Brownian motion to the last recorded value of the unobserved Brownian motion;
- $J(c) := \int_0^\infty e^{-\alpha t} Z^c_t dt - \int_{(0,\infty)} e^{-\alpha t} K(Z^c_{t-},\tau^c_{t-})|dc_t|$, where $K : \mathbb{R}\times[0,\infty) \to \mathbb{R}$ is a measurable function with polynomial growth, to be specified later (Setting 2.1(v); $J$ is itself; remark $|Z^c| \leq \overline{B^0} + \overline{B^1}$ (where a line over a process denotes its running supremum), so there are no integrability issues (due to the '$\epsilon$'-separation of the jumps of $c$)).

Notice that $\mathsf{E}\left[\int_0^\infty e^{-\alpha t} Z^c_t dt\right] = \mathsf{E}\left[\int_0^\infty e^{-\alpha t}\left(B^c_t - B^{1-c}_t\right)dt\right]$. Define $V(x) := \sup_{c\in\mathbf{C}} \mathsf{E}J(c)$.

Finally, with regard to Setting 2.7, introduce for every $c \in \mathbf{C}$ and controlled time $\mathcal{S}$,

$$\mathcal{D}(c,\mathcal{S}) := \{d \in \mathbf{C} : d^{\mathcal{S}^c} = c^{\mathcal{S}^c}\}.$$

Then let:

$$\mathbf{G} := \{\text{controlled times } \mathcal{S} \text{ such that } \forall c \forall d (d \in \mathcal{D}(c,\mathcal{S}) \Rightarrow \mathcal{S}^c = \mathcal{S}^d \,\&\, \mathcal{G}^c_{\mathcal{S}^c} = \mathcal{G}^d_{\mathcal{S}^d})\};$$

all the provisions of Section 2 (specifically, those of Setting 2.1, Setting 2.7, as well as Assumption 2.9) being thus satisfied. Thanks to Corollary 6.7, $\mathbf{G} = \{\text{controlled times } \mathcal{S} \text{ such that } \forall c \forall d (d^{\mathcal{S}^c} = c^{\mathcal{S}^c} \Rightarrow \mathcal{S}^c = \mathcal{S}^d)\}$.

Moreover, Assumption 4.3 can also be verified, as follows (here, the interplay between the nature of the controls, and the observed information, will be crucial to the argument). Let $\{c_1, c_2\} \subset \mathbf{C}$, $\mathcal{S} \in \mathbf{G}$, assume $c_1 \sim_{\mathcal{S}} c_2$, define $P := \mathcal{S}^{c_1} = \mathcal{S}^{c_2}$, take $A \in \mathcal{G}_P^{c_1} = \mathcal{G}_P^{c_2}$. Consider the process $c := c_1 \mathbb{1}_A + c_2 \mathbb{1}_{\Omega\backslash A}$. We claim $c \in \mathbf{C}$ (then, since $J(c) = J(c_1)\mathbb{1}_A + J(c_2)\mathbb{1}_{\Omega\backslash A}$, the condition of Proposition 4.2(1) will follow). For sure, $c$ is a $\{0,1\}$-valued, càdlàg process, vanishing at zero. Furthermore, if the jumps of $c_1$ and $c_2$ are temporally separated by $\epsilon_1 > 0$ and $\epsilon_2 > 0$, respectively, then the jumps of $c$ are temporally separated by $\epsilon := \epsilon_1 \wedge \epsilon_2 > 0$. Finally, $c$ is $\mathcal{G}^c$-predictable. To see this, note first that for all $t \in [0,\infty)$, $\mathcal{G}_t^c|_{\{t\leq P\}} = \mathcal{G}_t^{c_1}|_{\{t\leq P\}} = \mathcal{G}_t^{c_2}|_{\{t\leq P\}}$, whilst $\mathcal{G}_t^c|_{\{P<t\}\cap A} = \mathcal{G}_t^{c_1}|_{\{P<t\}\cap A}$ and $\mathcal{G}_t^c|_{\{P<t\}\cap(\Omega\backslash A)} = \mathcal{G}_t^{c_2}|_{\{P<t\}\cap(\Omega\backslash A)}$ (also, by Theorem 6.6, Proposition 6.5 and Proposition 6.9, from Part 2, $\{\{t\leq P\}, A\cap\{P<t\}, (\Omega\backslash A)\cap\{P<t\}\} \subset \sigma((B^{c_1})^{P\wedge t}) = \sigma((B^{c_2})^{P\wedge t}) = \sigma((B^c)^{P\wedge t}) \subset \mathcal{G}_t^c$). Then it will suffice to argue that $c\mathbb{1}_{[\![0,P]\!]} = c_1\mathbb{1}_{[\![0,P]\!]} = c_2\mathbb{1}_{[\![0,P]\!]}$, $c\mathbb{1}_{(\!(P,\infty)\!)}\mathbb{1}_A = c_1\mathbb{1}_{(\!(P,\infty)\!)}\mathbb{1}_A$ and $c\mathbb{1}_{(\!(P,\infty)\!)}\mathbb{1}_{\Omega\backslash A} = c_2\mathbb{1}_{(\!(P,\infty)\!)}\mathbb{1}_{\Omega\backslash A}$ are all $\mathcal{G}^c$-predictable. To see this, one need only consider the class of $\mathcal{G}^{c_1}$ or $\mathcal{G}^{c_2}$, respectively $\mathcal{G}^{c_1}$, $\mathcal{G}^{c_2}$, predictable processes $V$, for which $V\mathbb{1}_{[\![0,P]\!]}$, respectively, $V\mathbb{1}_{(\!(P,\infty)\!)}\mathbb{1}_A$, $V\mathbb{1}_{(\!(P,\infty)\!)}\mathbb{1}_{\Omega\backslash A}$, are $\mathcal{G}^c$-predictable. Then one establishes that this is a monotone class, containing the multiplicative class of all the left continuous, $\mathcal{G}^{c_1}$ or $\mathcal{G}^{c_2}$, respectively $\mathcal{G}^{c_1}$, $\mathcal{G}^{c_2}$, adapted processes. The Functional Monotone Class Theorem allows to conclude.

Now, we shall:

(a) Identify an instance of $K$ for which any control in $\mathbf{C}$ is optimal. It will emerge that $K(z,t) = -2z/\alpha$, $(z,t) \in \mathbb{R}\times[0,\infty)$, fits this bill, and then $V(x) = x/\alpha$.

(b) Provide a class of functions $K$ for which $V(x)$ is a symmetric function of the parameter $x$. Here $K$ will have the form:

$$K(z,t) = \int_{\mathbb{R}} \left( \frac{|\sqrt{t}u - z| - |z|}{\alpha} + \frac{e^{-\gamma|\sqrt{t}u-z|} - e^{-\gamma|z|}}{\alpha\gamma} \right) N(0,1)(du) + \mathbb{1}_{(0,\infty)}(z)L(z,t),\ (z,t) \in \mathbb{R}\times[0,\infty),$$

with $L$ nonnegative, measurable, bounded in polynomial growth; $\gamma := \sqrt{2\alpha}$. We will see, letting $c^\epsilon$ be the control, which waits an $\epsilon \in (0,\infty)$ amount of time each time it has jumped (also at time zero), and thereafter jumps at the first entrance time of $Z^c$ into $(-\infty, 0]$ – we are witnessing here, finally, an example of a whole sequence of non-control-constant controlled times, members of $\mathbf{G}$ — that for any such $K$, $\mathsf{E}J(c^\epsilon) \to V(x) = \frac{\gamma|x| + e^{-\gamma|x|}}{\alpha\gamma}$, as $\epsilon \downarrow 0$.

Indeed, according to Bellman's principle (Theorem 4.6), and the strong Markov property, for each $c \in \mathbf{C}$, the following process (where $V$ is, by a slight abuse of notation, the would-be value function[4]):

$$S_t^c := \int_0^t e^{-\alpha s} Z_s^c ds - \int_{(0,t]} e^{-\alpha s} K(Z_{s-}^c, \tau_{s-}^c)|dc_s| + e^{-\alpha t} V(Z_t^c, \tau_t^c), \tag{5.1}$$

[4]More precisely, for $z \in \mathbb{R}$, $u \in [0,\infty)$, $V(z,u)$ is the optimal payoff of the related optimal control problem in which, *ceteris paribus*, $B^1 = z + H_{u+\cdot}$, for a Brownian motion $H$ independent of $B^0$.

should be a $(\mathcal{G}^c, \mathsf{P})$-supermartingale (in $t \in [0,\infty)$). Moreover, if an optimal strategy $c^*$ exists, then $S^{c^*}$ should be a $(\mathcal{G}^{c^*}, \mathsf{P})$-martingale (or, when dealing with a sequence/net of optimizing controls, it should, *in expectation*, 'be increasingly close to being one').

Guided by (5.1), let us now consider the semimartingale decomposition of $S^c$, assuming $K$ is such that in fact (i) for $z \in \mathbb{R}$, $s \in [0,\infty)$ (again with a slight abuse of notation) $V(z,s) = V(z)$ (i.e. no explicit lag dependence in the value function); (ii) $V$ is of class $C^1$, and also twice differentiable, with second derivative continuous, except possibly at finitely many points, wherein still the left and right derivatives exist and remain continuous from the left, respectively right; and (iii) $V$ and $V'$ are bounded in polynomial growth.

The semimartingale decomposition (for which we require, in principle, the "usual conditions") may then be effected relative to the completed measure $\overline{\mathsf{P}}$ and the usual augmentation $\overline{\mathcal{G}^c}_+$ of $\mathcal{G}^c$, with respect to which $Z^c$ is a semimartingale (indeed, its jump part is clearly of finite variation, whilst its continuous part is, in fact, a Brownian motion relative to the augmentation of the natural filtration of $(B^0, B^1)$).

We thus obtain, by the Itô-Tanaka-Meyer formula [12, p. 214, Theorem IV.70, p. 216, Corollary IV.1], $\overline{\mathsf{P}}$-a.s. for all $t \in [0,\infty)$:

$$
\begin{aligned}
S_t^c \;=\;& V(x) + \underbrace{\int_0^t e^{-\alpha s} Z_{s-}^c ds}_{=:C_1} + \underbrace{\int_{(0,t]} e^{-\alpha s}(-K(Z_{s-}^c, \tau_{s-}^c))|dc_s|}_{=:J_1} + \underbrace{\int_0^t e^{-\alpha s}(-\alpha)V(Z_{s-}^c)ds}_{=:C_2} \\
&+ \overbrace{\int_0^t e^{-\alpha s} V'(Z_{s-}^c) dZ_s^c}^{(*)} + \underbrace{\int_0^t e^{-\alpha s}\frac{1}{2}V''(Z_{s-}^c)d\overbrace{[Z^c]_s^{\mathrm{cts}}}^{s}}_{=:C_3} + \sum_{0<s\leq t} e^{-\alpha s}\left[\overbrace{\Delta V(Z_s^c)}^{(**)}\overbrace{-V'(Z_{s-}^c)\Delta Z_s^c}^{(*)}\right]
\end{aligned} \tag{5.2}
$$

Note that the starred parts combine into:

$$
\underbrace{\overbrace{\int_0^t e^{-\alpha s} V'(Z_{s-}^c) d(Z^c)_s^{\mathrm{cts}}}^{(*)}}_{=:M_1}, \tag{5.3}
$$

which is a $(\overline{\mathcal{G}^c}_+, \overline{\mathsf{P}})$-martingale in $t \in [0,\infty)$ (since $|Z^c| \leq \overline{B^0} + \overline{B^1}$). On the other hand, the compensator of the double-starred term is:

$$
\underbrace{\overbrace{\int_{(0,t]} |dc_s| e^{-\alpha s}\left[\int_{\mathbb{R}} N(0,1)(du)\left(V(\sqrt{\tau_{s-}^c}u - Z_{s-}^c) - V(Z_{s-}^c)\right)\right]}^{(**)}}_{=:J_2}, \tag{5.4}
$$

making:

$$\underbrace{\overbrace{\int_{(0,t]}|dc_s|e^{-\alpha s}\left[\Delta V(Z^c_s)-\int_{\mathbb{R}}N(0,1)(du)\left(V(\sqrt{\tau^c_{s-}}u-Z^c_{s-})-V(Z^c_{s-})\right)\right]}^{(**)}}_{=:M_2} \tag{5.5}$$

into a $(\mathcal{G}^c,\mathsf{P})$-martingale (in $t\in[0,\infty)$). For, if $\tau$ is a predictable stopping time with respect to some filtration (in continuous time) $\mathcal{Z}$, $U$ is a $\mathcal{Z}_\tau$-measurable random variable, and $\mathsf{Q}$ a probability measure with $\mathsf{Q}[|U|\mathbb{1}_{[0,t]}\circ\tau]<\infty$ for each $t\in[0,\infty)$, then the compensator of $U\mathbb{1}_{[\![\tau,\infty[\![}$ (relative to $(\mathcal{Z},\mathsf{Q})$) is $\mathsf{Q}[U|\mathcal{Z}_{\tau-}]\mathbb{1}_{[\![\tau,\infty[\![}$. This fact may be applied to each jump time of the $\mathcal{G}^c$-predictable process $c$ (since $|Z^c|\leq\overline{B^0}+\overline{B^1}$), whence linearity allows to conclude (due to the '$\epsilon$-separation' of the jumps of $c$).

Remark now that the properties of being a càdlàg (super)martingale [13, p. 173, Lemma II.67.10] or predictable process (of finite variation) are preserved when passing to the usual augmentation of a filtered probability space. Therefore it follows that, relative to $(\overline{\mathcal{G}^c}_+,\overline{\mathsf{P}})$, $M:=M_1+M_2$ is a martingale, whilst $J=J_1+J_2$ (respectively $C=C_1+C_2+C_3$) is a pure-jump (respectively continuous) predictable process of finite variation. On the other hand, the process $S^c$ is supposed to be a $(\overline{\mathcal{G}^c}_+,\overline{\mathsf{P}})$-supermartingale. But then, we have obtained in this manner nothing but the Doob-Meyer decomposition of $S^c=V(x)+M+J+C$, so that $J$ and $C$ are both, respectively continuous and pure-jump, nonincreasing processes of finite variation [9, p. 32, Corollary 3.16] [10, p. 412, Theorem 22.5].

Now assume furthermore that an optimizing (as $\epsilon\downarrow 0$) net of optimal controls is to wait for a period of $\epsilon$ after each jump of $c$ and also at the start, and then each time change the observed Brownian motion precisely at the first entrance time of $Z^c$ into the set $(-\infty,-l]$ /for some prespecified level $l\in[0,\infty)$/. Remark such a control is previsible with respect to $\mathcal{G}^c$.

Then we should like to have (the first two conditions follow from the supermartingale property, the last two from the 'near martingale'/'limiting martingale' condition; the presence of the a.e. qualifiers being a reflection of the Occupation Time Density Formula [12, p. 216, Corollary IV.1]):

$$\text{[from } C] \quad z-\alpha V(z)+\frac{1}{2}V''(z)\leq 0,\ \text{for a.e. } z\in\mathbb{R}; \tag{5.6}$$

$$\text{[from } J] \quad -K(z,t)+\int_{\mathbb{R}}[V(\sqrt{t}u-z)-V(z)]N(0,1)(du)\leq 0,\ \text{for all } z\in\mathbb{R},\, t\in[0,\infty); \tag{5.7}$$

$$\text{[from } C] \quad z-\alpha V(z)+\frac{1}{2}V''(z)=0,\ \text{for a.e. } z\geq -l; \tag{5.8}$$

$$\text{[from } J] \quad -K(z,t)+\int_{\mathbb{R}}[V(\sqrt{t}u-z)-V(z)]N(0,1)(du)=0,\ \text{for all } z\leq -l,\, t\in(0,\infty). \tag{5.9}$$

This concludes the first part of the analysis, deriving what ought to hold of $V$. In the second part we flip, as is usual, the argument upside-down. $V$ will be specified *a priori* (along with $K$); $S^c$ remains defined in terms of this prespecified $V$, viz. Eq. (5.1); and *then* it is shown that $V(x)$ is the optimal payoff, via the semimartingale decomposition of $S^c$ which the latter will continue to enjoy in the form (5.2)-(5.3)-(5.4)-(5.5).

Indeed, as regards (a), we may take $K(z,t) = -2z/\alpha$ and $V(z) = z/\alpha$, for $z \in \mathbb{R}$, in which case (5.6)-(5.7)-(5.8)-(5.9) are all satisfied with equality. Taking expectations in (5.1)-(5.2), and passing to the limit $t \to \infty$ via dominated convergence, we see that $V(x)$ *is* the optimal payoff, and *any* control from $\mathbf{C}$ realizes it.

For a less degenerate case, let us solve (5.8) on $z \in [-l, \infty)$ and in the general solution throw away the exponentially increasing part. Then $V(z) = \psi(z)$ for $z \geq -l$, where $\psi(u) := \frac{u}{\alpha} + Ae^{-\gamma u}$ ($u \in \mathbb{R}$, $A \in \mathbb{R}$, $\gamma := \sqrt{2\alpha}$). To obtain a symmetric function of $V$ it is natural to take $l = 0$ and then $V(z) = \psi(|z|)$, $z \in \mathbb{R}$. For such a $V$, (5.6) is in fact satisfied by a strict inequality on $z \in (-\infty, 0)$; and $V$ is $C^1$ for $A = 1/(\gamma\alpha)$ (and then it is even $C^2$). Then (5.9) and (5.7) essentially necessitate taking the form of $K$ as specified in (b) on p. 17.

Now, to see that $c^\epsilon$ (as described in (b)) is in fact an optimizing net of controls (as $\epsilon \downarrow 0$), first take expectations in (5.1)-(5.2), and pass to the limit as $t \to \infty$ (via dominated and monotone convergence), in order to see that $V(x) \geq \mathsf{E}J(c)$ for each $c \in \mathbf{C}$; second apply (5.1)-(5.2) to $c = c^\epsilon$, pass to the limit $t \to \infty$, and note that

$$\mathsf{E}\int_0^\infty e^{-\alpha s}\left(Z^{c^\epsilon}_{s-} - \alpha V(Z^{c^\epsilon}_{s-}) + \frac{1}{2}V''(Z^{c^\epsilon}_{s-})\right) ds = 2\mathsf{E}\int_0^\infty e^{-\alpha s} Z^{c^\epsilon}_s \mathbb{1}_{(-\infty,0)} \circ Z^{c^\epsilon}_s \, ds \to 0$$

as $\epsilon \downarrow 0$. To convince the reader of this, it will suffice to check:

$$\mathsf{E}\int_0^T e^{-\alpha s} Z^{c^\epsilon}_s \mathbb{1}_{(-\infty,0)} \circ Z^{c^\epsilon}_s \, ds \to 0, \text{ as } \epsilon \downarrow 0,$$

for each $T \in (0, \infty)$. Fix such a $T$. Further, it will be sufficient to argue that:

$$\mathsf{E}\int_0^T e^{-\alpha s} Z^{c^\epsilon}_s \mathbb{1}_{(-\infty,-a)} \circ Z^{c^\epsilon}_s \, ds \to 0, \text{ as } \epsilon \downarrow 0,$$

for each $a \in (0, \infty)$. Fix such an $a$. It will now be enough to demonstrate that $\mathsf{P}$-a.s. the Lebesgue measure of the set of times $A^\epsilon := \{s \in [0,T] : Z^{c^\epsilon}_s < -a\}$ converges to 0, as $\epsilon \downarrow 0$. Call the intervals of time $A_k := [J^{c^\epsilon}_k, J^{c^\epsilon}_k + \epsilon)$, $k \in \mathbb{N}_0$, holding periods for the control $c^\epsilon$. Remark the holding periods constitute a pairwise disjoint cover of $A^\epsilon$. Moreover, if $s \in A^\epsilon \cap A_k$ for $k \in \mathbb{N}$, then (denoting $t_0 := J^{c^\epsilon}_k$, $j_0 := c^\epsilon_{t_0}$ and $T_0 := J^{c^\epsilon}_{k-1}$) $-a > Z^{c^\epsilon}_s = B^{j_0}_s - B^{1-j_0}_{t_0}$, whilst $0 \geq Z^{c^\epsilon}_{t_0^-} = B^{1-j_0}_{t_0} - B^{j_0}_{T_0}$, hence $-a > B^{j_0}_s - B^{j_0}_{T_0}$. Thus, if $s \in A^\epsilon$ belongs to the $k$-th holding period for $c^\epsilon$ (and $k \geq 1$), then in the time interval between the start of the $(k-1)$-th holding period and the end of the $k$-th holding period, one of the Brownian motions $B^0$ and $B^1$ must have moved by more than $a$. However, thanks to the continuity of the sample paths of $B^0$ and $B^1$, the infimum over the amounts of time required for either $B^0$ or $B^1$ to move by more than $a$ (on the interval $[0,T]$) is strictly positive (albeit dependent on the sample point). It follows that the number of $k \geq 1$ for which there can be an $s \in A^\epsilon$ with $s \in A_k$ is bounded by some number, depending on the sample point, but not on $\epsilon$, and this establishes the claim.

# Part 2. Stopping times, stopped processes and natural filtrations at stopping times – informational consistency

Recall from p. 9 the contents of the third, final, bullet point remark following Assumption 2.9 – it is to the questions posed and motivated there, that we now turn our attention. Along the way, we shall (be forced to) investigate (i) the precise relationship between the sigma-fields of the stopped processes, on the one hand, and the natural filtrations of the processes at these stopping times, on the other and (ii) the nature of the stopping times of the processes and of the stopped processes, themselves. Here is an informal statement of the kind of results that we will /seek to/ formally establish ($\mathcal{F}^X$ denotes the natural filtration of a process $X$):

> If $X$ is a process, and $S$ a time, then $S$ is a stopping time of $\mathcal{F}^X$, if and only if it is a stopping time of $\mathcal{F}^{X^S}$. When so, then $\mathcal{F}^X_S = \sigma(X^S)$. In particular, if $X$ and $Y$ are two processes, and $S$ is a stopping time of either $\mathcal{F}^X$ or of $\mathcal{F}^Y$, with $X^S = Y^S$, then $S$ is a stopping time of $\mathcal{F}^X$ and $\mathcal{F}^Y$ both, moreover $\mathcal{F}^X_S = \sigma(X^S) = \sigma(Y^S) = \mathcal{F}^Y_S$. Further, if $U \leq V$ are two stopping times of $\mathcal{F}^X$, $X$ again being a process, then $\sigma(X^U) = \mathcal{F}^X_U \subset \mathcal{F}^X_V = \sigma(X^V)$.

We will perform this investigation into the nature of information generated by processes in the two 'obvious' settings: first the 'measure-theoretic' one, without reference to probability measures (Section 6) and then the 'probabilistic' one, involving a complete probability measure, under which all the filtrations and $\sigma$-fields are completed (Section 7). This will also dovetail nicely with the parallel development of the two frameworks for stochastic control – the 'measure-theoretic' and the 'probabilistic' one – from Part 1.

Now, the most important findings of this part are as follows:

- Lemma 6.2, Proposition 6.5, Theorem 6.6, Theorem 6.7 and Proposition 6.9, in the 'measure-theoretic' case;
- Corollaries 7.2 and 7.3 (in discrete time) and Proposition 7.6, Corollaries 7.7, 7.9 and 7.10 (in continuous time), for the case with completions.

(Indeed, we have already referenced many of these results in Part 1 – which fact further demonstrates their relevance to this study.) It emerges that everything that intuitively *ought* to hold, *does* hold, if either the time domain is discrete, or else the underlying space is Blackwell (and, when dealing with completions, the stopping times are predictable; but see the negative results of Examples 7.4 and 7.5). While we have not been able to drop the "Blackwell assumption", we believe many of the results should still hold true under weaker conditions – this remains open to future research.

Finally, we note that the whole of the remainder of this part is in fact independent from the rest of the paper (in particular, from Part 1).

## 6. The 'measure-theoretic' case

We begin by fixing quite a bit of notation and by enunciating a couple of well-known (and some less well-known) measure-theoretic facts along the way – we ask the reader to bear with us.

(•) $T = \mathbb{N}_0$ or $T = [0, \infty)$, with the usual linear order, $\Omega$ is some set and $(E, \mathcal{E})$ a measurable space. By a **process** (on $\Omega$, with time domain $T$ and values in $E$), we mean simply a collection $X = (X_t)_{t \in T}$ of functions from $\Omega$ into $E$. With $\mathcal{F}^X_t := \sigma(X_s : s \in [0, t])$ (for $t \in T$), $\mathcal{F}^X = (\mathcal{F}^X_t)_{t \in T}$ is then the **natural filtration** of $X$.[5] Remark that for every $t \in T$, $\mathcal{F}^X_t = \sigma(X|_{[0,t]})$, with $X|_{[0,t]}(\omega) = (X_s(\omega))_{s \in [0,t]}$ for $\omega \in \Omega$; $X|_{[0,t]}$ is an $\mathcal{F}^X_t/\mathcal{E}^{\otimes[0,t]}$-measurable mapping. The $\omega$-**sample path** of $X$, $X(\omega)$, is the mapping from $T$ into $E$, given by $(t \mapsto X_t(\omega))$, $\omega \in \Omega$. In this sense, $X$ may of course be viewed as an $\mathcal{F}^X_\infty/\mathcal{E}^{\otimes T}$-measurable mapping, indeed $\mathcal{F}^X_\infty = \sigma(X)$. Then $\mathrm{Im}X$ will denote the range (image) of the mapping $X : \Omega \to E^T$.

(•) If further $S : \Omega \to T \cup \{\infty\}$ is a time and $\mathcal{G}$ is a filtration on $\Omega$, then $\mathcal{G}_S := \{A \in \mathcal{G}_\infty : A \cap \{T \le t\} \in \mathcal{G}_t \text{ for all } t \in T\}$ is the **filtration** $\mathcal{G}$ **at (the time)** $S$, whilst the **stopped process** $X^S$ (of a process $X$) is defined via $X^S_t(\omega) := X_{S(\omega)\wedge t}(\omega)$, $(\omega, t) \in \Omega \times T$. Note, that if $T = \mathbb{N}_0$, $X$ is a $\mathcal{G}$-adapted process and $S$ is a $\mathcal{G}$-stopping time, then $X^S$ is automatically adapted to the stopped filtration $(\mathcal{G}_{n\wedge S})_{n \in \mathbb{N}_0}$. For, if $n \in \mathbb{N}_0$, $Z \in \mathcal{E}$, then $(X^S_n)^{-1}(Z) = \left(\cup_{\mathbb{N}_0 \ni m \le n} X^{-1}_m(Z) \cap \{S = m\}\right) \cup \left(X^{-1}_n(Z) \cap \{n < S\}\right) \in \mathcal{G}_{S\wedge n}$. On the other hand, in continuous-time, when $T = [0, \infty)$, if $X$ is $\mathcal{G}$-progressively measurable and $S$ is a $\mathcal{G}$-stopping time, then $X^S$ is also adapted to the stopped filtration $(\mathcal{G}_{t\wedge S})_{t \in [0,\infty)}$ (and is $\mathcal{G}$-progressively measurable) [11, p. 9, Proposition 2.18]. Remark also that every right- or left-continuous Euclidean space-valued $\mathcal{G}$-adapted process is automatically $\mathcal{G}$-progressively measurable.

(•) Next, for a $\sigma$-field $\mathcal{F}$ on $\Omega$, the measurable space $(\Omega, \mathcal{F})$ is said to be (i) **separable** or **countably generated**, when it admits a countable generating set; (ii) **Hausdorff**, or **separated**, when its atoms[6] are the singletons of the members of $\Omega$ [4, p. 10]; and finally (iii) **Blackwell** when its associated Hausdorff space ($(\Omega, \mathcal{F})$ quotiented out by $\sim$ of Footnote 6) $(\dot{\Omega}, \dot{\mathcal{F}})$ is Souslin [4, p. 50, III.24]. Furthermore, a Souslin space is a measurable space, which is Borel isomorphic to a Souslin topological space. The latter in turn is a Hausdorff topological space, which is also a continuous image of a Polish space (i.e. of a completely metrizable separable topological space). Every Souslin measurable space is necessarily separable and separated. [4, p. 46, III.16; p. 76, III.67] For a measurable space, clearly being Souslin is equivalent to being simultaneously Blackwell and Hausdorff. The key result for us, however, will be Blackwell's Theorem [4, p. 51 Theorem III.26] (repeated here for the reader's convenience – we shall use it time and again):

> **Blackwell's Theorem**. Let $(\Omega, \mathcal{F})$ be a Blackwell space, $\mathcal{G}$ a sub-$\sigma$-field of $\mathcal{F}$ and $\mathcal{S}$ a separable sub-$\sigma$-field of $\mathcal{F}$. Then $\mathcal{G} \subset \mathcal{S}$, if and only if every atom of $\mathcal{G}$ is a

[5] $[0, t]$ is to be understood throughout as the set $\{0, \dots, t\}$ when $t \in T = \mathbb{N}_0$.

[6] Equivalence classes for the equivalence relation $\sim$ on $\Omega$, given by $(\omega \sim \omega') \Leftrightarrow (\text{for all } A \in \mathcal{F}, \mathbb{1}_A(\omega) = \mathbb{1}_A(\omega'))$, $\{\omega, \omega'\} \subset \Omega$.

union of atoms of $\mathcal{S}$. In particular, a $\mathcal{F}$-measurable real function $g$ is $\mathcal{S}$-measurable, if and only if $g$ is constant on every atom of $\mathcal{S}$.

($\bullet$) Remark finally that if $Y$ is a mapping from $A$ into some Hausdorff (resp. separable) measurable space $(B,\mathcal{B})$, then $Y$ is constant on the atoms of $\sigma(Y)$ (resp. $\sigma(Y)$ is separable). For, if $\{\omega_1,\omega_2\}\subset A$, $A$ being an atom of $\sigma(Y)$, with *per absurdum* (under the hypothesis that $(B,\mathcal{B})$ is Hausdorff) $Y(\omega_1)\neq Y(\omega_2)$, then there is a $W\in\mathcal{B}$ with $\mathbb{1}_W(Y(\omega_1))\neq\mathbb{1}_W(Y(\omega_2))$, hence $\mathbb{1}_{Y^{-1}(W)}(\omega_1)\neq\mathbb{1}_{Y^{-1}(W)}(\omega_2)$, a contradiction. Conversely, if $Y$ is a surjective mapping from $A$ onto some measurable space $(B,\mathcal{B})$, constant on the atoms of $\sigma(Y)$, then $(B,\mathcal{B})$ is Hausdorff. For, if $\{b,b'\}\subset B$ and $\mathbb{1}_W(b)=\mathbb{1}_W(b')$ for all $W\in\mathcal{B}$, then if $\{a,a'\}\subset A$ are such that $Y(a)=b$ and $Y(a')=b'$, $\mathbb{1}_Z(a)=\mathbb{1}_Z(a')$ for all $Z\in\sigma(Y)$, so that $a$ and $a'$ belong to the same atom of $(A,\sigma(Y))$ and consequently $b=b'$. Furthermore, any measurable subspace (with the trace $\sigma$-field) of a separable (resp. Hausdorff) space is separable (resp. Hausdorff). Lastly, if $f:A\to(B,\mathcal{B})$ is any map into a measurable space, then the atoms of $\sigma(f)$ always 'respect' the equivalence relation induced by $f$, i.e., for $\{\omega,\omega'\}\subset A$, if $f(\omega)=f(\omega')$, then $\omega$ and $\omega'$ belong to the same atom of $\sigma(f)$: for all $\Sigma\in\mathcal{B}$, $\mathbb{1}_{f^{-1}(\Sigma)}(\omega)=\mathbb{1}_\Sigma(f(\omega))=\mathbb{1}_\Sigma(f(\omega'))=\mathbb{1}_{f^{-1}(\Sigma)}(\omega')$.

Now, a key result in this section will establish that, for a process $X$ and a stopping time $S$ thereof, $\sigma(X^S)=\mathcal{F}^X_S$, i.e. that the initial structure (with respect to $\mathcal{E}^{\otimes T}$) of the stopped process coincides with the filtration of the process at the stopping time – under suitable conditions.

Indeed, our first lemma towards this end tells us that elements of $\mathcal{F}^X_S$ are *functions* (albeit not (as yet) necessarily *measurable* functions) of the stopped process $X^S$.

**Lemma 6.1** (Key lemma)**.** *Let $X$ be a process (on $\Omega$, with time domain $T$ and values in $E$), $S$ an $\mathcal{F}^X$-stopping time, $A\in\mathcal{F}^X_S$. Then the following holds for every $\{\omega,\omega'\}\subset\Omega$: If $X_t(\omega)=X_t(\omega')$ for all $t\in T$ with $t\leq S(\omega)\wedge S(\omega')$, then $S(\omega)=S(\omega')$, $X^S(\omega)=X^S(\omega')$ and $\mathbb{1}_A(\omega)=\mathbb{1}_A(\omega')$.*

*Proof.* Define $t:=S(\omega)\wedge S(\omega')$. If $t=\infty$, for sure $S(\omega)=S(\omega')$. If not, then $\{S\leq t\}\in\mathcal{F}^X_t$, so that there is a $U\in\mathcal{E}^{\otimes[0,t]}$ with $\{S\leq t\}=X|_{[0,t]}^{-1}(U)$. Then at least one of $\omega$ and $\omega'$ must belong to $\{S\leq t\}$, hence to $X|_{[0,t]}^{-1}(U)$. Consequently, since by assumption $X|_{[0,t]}(\omega)=X|_{[0,t]}(\omega')$, both do. It follows that $S(\omega)=S(\omega')$. In particular, $X^S(\omega)=X^S(\omega')$.

Similarly, since $A\in\mathcal{F}^X_S$, $A\cap\{S\leq t\}\in\mathcal{F}^X_t$, so that there is a $U\in\mathcal{E}^{\otimes[0,t]}$ (resp. $U\in\mathcal{E}^{\otimes T}$), with $A\cap\{S\leq t\}=X|_{[0,t]}^{-1}(U)$ (resp. $A\cap\{S\leq t\}=X^{-1}(U)$), when $t<\infty$ (resp. $t=\infty$). Then $\mathbb{1}_A(\omega)=\mathbb{1}_{A\cap\{S\leq t\}}(\omega)=\mathbb{1}_U(X|_{[0,t]}(\omega))=\mathbb{1}_U(X|_{[0,t]}(\omega'))=\mathbb{1}_{A\cap\{S\leq t\}}(\omega')=\mathbb{1}_A(\omega')$ (resp. $\mathbb{1}_A(\omega)=\mathbb{1}_{A\cap\{S\leq t\}}(\omega)=\mathbb{1}_U(X(\omega))=\mathbb{1}_U(X(\omega'))=\mathbb{1}_{A\cap\{S\leq t\}}(\omega')=\mathbb{1}_A(\omega')$). □

Our second lemma will allow to handle the discrete case.

**Lemma 6.2** (Stopping times)**.** *Let $X$ be a process (on $\Omega$, with time domain $\mathbb{N}_0$ and values in $E$). For a time $S:\Omega\to\mathbb{N}_0\cup\{\infty\}$ the following are equivalent:*

*(1) $S$ is an $\mathcal{F}^X$-stopping time.*
*(2) $S$ is an $\mathcal{F}^{X^S}$-stopping time.*

*Proof.* Suppose first $S$ is an $\mathcal{F}^{X^S}$-stopping time. Let $n \in \mathbb{N}_0$. Then for each $m \in [0,n]$, $\{S \leq m\} \in \mathcal{F}^{X^S}_m$, so there is an $E_m \in \mathcal{E}^{\otimes[0,m]}$ with $\{S \leq m\} = (X^S|_{[0,m]})^{-1}(E_m)$. Then $\{S = m\} \subset X|_{[0,m]}^{-1}(E_m) \subset \{S \leq m\}$. Consequently $\{S \leq n\} = \cup_{m\in[0,n]} X|_{[0,m]}^{-1}(E_m) \in \mathcal{F}^X_n$.

Conversely, suppose $S$ is an $\mathcal{F}^X$-stopping time. Let $n \in \mathbb{N}_0$. For each $m \in [0,n]$, $\{S \leq m\} \in \mathcal{F}^X_m$, hence there is an $E_m \in \mathcal{E}^{\otimes[0,m]}$ with $\{S \leq m\} = X|_{[0,m]}^{-1}(E_m)$. Then $\{S = m\} \subset (X^S|_{[0,m]})^{-1}(E_m) \subset \{S \leq m\}$. Consequently $\{S \leq n\} = \cup_{m\in[0,n]}(X^S|_{[0,m]})^{-1}(E_m) \in \mathcal{F}^{X^S}_n$. □

The next step establishes that members of $\mathcal{F}^X_S$ are, in fact, *measurable* functions of the stopped process $X^S$ – at least under certain conditions (but always in the discrete case).

**Proposition 6.3.** *Let $X$ be a process, $S$ an $\mathcal{F}^X$-stopping time. If any one of the conditions (1)-(2)-(3) below is fulfilled, then $\mathcal{F}^X_S \subset \sigma(X^S)$ (where $X^S$ is viewed as assuming values in $(E^T, \mathcal{E}^{\otimes T})$).*

1. *$T = \mathbb{N}_0$.*
2. *$\mathrm{Im}X^S \subset \mathrm{Im}X$.*
3. *(a) $(\Omega, \mathcal{G})$ is Blackwell for some $\sigma$-field $\mathcal{G} \supset \mathcal{F}^X_S \vee \sigma(X^S)$.*
   *(b) $\sigma(X^S)$ is separable (in particular, this obtains if $(\mathrm{Im}X^S, \mathcal{E}^{\otimes T}|_{\mathrm{Im}X^S})$ is separable).*
   *(c) $X^S$ is constant on the atoms of $\sigma(X^S)$, i.e. $(\mathrm{Im}X^S, \mathcal{E}^{\otimes T}|_{\mathrm{Im}X^S})$ is Hausdorff.*

*Remark* 6.4.

1. Condition (2) is clearly not very innocuous, but will typically be met when $X$ is the coordinate process on a canonical space.
2. Condition ((3)b) is verified, if there is a $\mathcal{D} \subset E^T$, with $\mathrm{Im}X^S \subset \mathcal{D}$, such that the trace $\sigma$-field $\mathcal{E}^{\otimes T}|_{\mathcal{D}}$ is separable. For example (when $T = [0,\infty)$) this is the case if $E$ is a second countable (e.g. separable metrizable) topological space endowed with its (then separable) Borel $\sigma$-field, and the sample paths of $X^S$ are, say, all left- or all right-continuous (take $\mathcal{D}$ to be all the left- or all the right-continuous paths from $E^{[0,\infty)}$).
3. Finally, condition ((3)c) follows, if $(E, \mathcal{E})$ is Hausdorff and so, in particular, when the singletons of $E$ belong to $\mathcal{E}$.

*Proof.* Assume first (1). Let $A \in \mathcal{F}^X_S$ and $n \in \mathbb{N}_0 \cup \{\infty\}$. Then $A \cap \{S = n\} \in \mathcal{F}^X_n$, so $A \cap \{S = n\} = (X|_{[0,n]})^{-1}(Z)$ (resp. $A \cap \{S = n\} = X^{-1}(Z)$) for some $Z \in \mathcal{E}^{\otimes[0,n]}$ (resp. $Z \in \mathcal{E}^{\otimes \mathbb{N}_0}$), when $n < \infty$ (resp. $n = \infty$). But then $A \cap \{S = n\} = (X^S|_{[0,n]})^{-1}(Z) \cap \{S = n\}$ (resp. $A \cap \{S = n\} = (X^S)^{-1}(Z) \cap \{S = n\}$). Thanks to Lemma 6.2, $\{S = n\} \in \sigma(X^S)$, and we are done.

Assume next (2). Let $A \in \mathcal{F}^X_S$. Then $\mathbb{1}_A = F \circ X$ for some $\mathcal{E}^{\otimes T}/\mathcal{B}(\{0,1\})$-measurable mapping $F$. Since $\mathrm{Im}X^S \subset \mathrm{Im}X$, for any $\omega \in \Omega$, there is an $\omega' \in \Omega$ with $X(\omega') = X^S(\omega)$, and then thanks to Lemma 6.1 $X^S(\omega') = X^S(\omega)$, moreover, $F \circ X^S(\omega) = F \circ X(\omega') = \mathbb{1}_A(\omega') = \mathbb{1}_A(\omega)$. It follows that $\mathbb{1}_A = F \circ X^S$.

Assume now (3). We apply Blackwell's Theorem. Specifically, on account of ((3)a) $(\Omega, \mathcal{G})$ is a Blackwell space and $\mathcal{F}^X_S$ is a sub-$\sigma$-field of $\mathcal{G}$; on account of ((3)b), $\sigma(X^S)$ is a separable sub-$\sigma$-field of $\mathcal{G}$. Finally, thanks to ((3)c) and Lemma 6.1, every atom (equivalently, every element) of $\mathcal{F}^X_S$ is a union of atoms of $\sigma(X^S)$. It follows that $\mathcal{F}^X_S \subset \sigma(X^S)$. $\square$

The continuous-time analogue of Lemma 6.2 is now as follows:

**Proposition 6.5** (Stopping times). *Let $X$ be a process (on $\Omega$, with time domain $[0,\infty)$ and values in $E$), $S : \Omega \to [0,\infty]$ a time. Suppose:*

*(1) $\sigma(X|_{[0,t]})$ and $\sigma(X^{S\wedge t})$ are separable, $(\mathrm{Im}X|_{[0,t]}, \mathcal{E}^{\otimes[0,t]})$ and $(\mathrm{Im}X^{S\wedge t}, \mathcal{E}^{\otimes T}|_{\mathrm{Im}X^{S\wedge t}})$ Hausdorff for each $t \in [0,\infty)$.*
*(2) $X^S$ and $X$ are both measurable with respect to a Blackwell $\sigma$-field $\mathcal{G}$ on $\Omega$.*

*Then the following are equivalent:*

*(a) $S$ is an $\mathcal{F}^X$-stopping time.*
*(b) $S$ is an $\mathcal{F}^{X^S}$-stopping time.*

*Proof.* Suppose first $S$ is an $\mathcal{F}^{X^S}$-stopping time. Let $t \in [0,\infty)$. Then $\{S \le t\} \in \mathcal{F}^{X^S}_t$. But $\mathcal{F}^{X^S}_t = \sigma(X^S|_{[0,t]}) \subset \sigma(X|_{[0,t]}) = \mathcal{F}^X_t$. This follows from the fact that every atom of $\sigma(X^S|_{[0,t]})$ is a union of atoms of $\sigma(X|_{[0,t]})$ (whence one can apply Blackwell's Theorem). To see this, note that if $\omega$ and $\omega'$ belong to the same atom of $\sigma(X|_{[0,t]})$, then $X|_{[0,t]}(\omega) = X|_{[0,t]}(\omega')$ (since $(\mathrm{Im}X|_{[0,t]}, \mathcal{E}^{\otimes[0,t]})$ is Hausdorff). But then $X^S_s(\omega) = X^S_s(\omega')$ for all $s \in [0, (S(\omega)\wedge t) \wedge (S(\omega')\wedge t)]$, and so by Lemma 6.1 (applied to the process $X^S$ and the stopping time $S \wedge t$ of $\mathcal{F}^{X^S}$), $(X^S)^{S\wedge t}(\omega) = (X^S)^{S\wedge t}(\omega')$, i.e. $X^S|_{[0,t]}(\omega) = X^S|_{[0,t]}(\omega')$. We conclude that $\omega$ and $\omega'$ belong to the same atom of $\sigma(X^S|_{[0,t]})$.

Conversely, assume $S$ is an $\mathcal{F}^X$-stopping time. Let $t \in [0,\infty)$. Then $\{S \le t\} \in \mathcal{F}^X_{S\wedge t}$, $S \wedge t$ is an $\mathcal{F}^X$-stopping time and thanks to Proposition 6.3, $\mathcal{F}^X_{S\wedge t} \subset \sigma(X^{S\wedge t}) = \sigma(X^S|_{[0,t]})$. $\square$

What finally follows is the main result of this section. As mentioned in the Introduction, it generalizes canonical-space results already available in literature.

**Theorem 6.6** (Generalized Galmarino's test). *Let $X$ be a process, $S$ an $\mathcal{F}^X$-stopping time. If $T = \mathbb{N}_0$, then $\sigma(X^S) = \mathcal{F}^X_S$. Moreover, if $X^S$ is $\mathcal{F}^X_S/\mathcal{E}^{\otimes T}$-measurable (in particular, if it is adapted to the stopped filtration $(\mathcal{F}^X_{t\wedge S})_{t\in T}$) and either one of the conditions:*

*(1) $\mathrm{Im}X^S \subset \mathrm{Im}X$.*
*(2) (a) $(\Omega, \mathcal{G})$ is Blackwell for some $\sigma$-field $\mathcal{G} \supset \mathcal{F}^X_\infty$.*
*(b) $\sigma(X^S)$ is separable.*
*(c) $(\mathrm{Im}X^S, \mathcal{E}^{\otimes T}|_{\mathrm{Im}X^S})$ is Hausdorff.*

*is met, then the following statements are equivalent:*

*(i) $A \in \mathcal{F}^X_S$.*
*(ii) $\mathbb{1}_A$ is constant on every set on which $X^S$ is constant and $A \in \mathcal{F}^X_\infty$.*
*(iii) $A \in \sigma(X^S)$.*

*Proof.* The first claim, which assumes $T = \mathbb{N}_0$, follows from Proposition 6.3 and the fact that automatically $X^S$ is $\mathcal{F}^X_S/\mathcal{E}^{\otimes T}$-measurable in this case.

In general, implication (i)⇒(ii) follows from Lemma 6.1. Implication (ii)⇒(iii) proceeds as follows.

Suppose first (1). Let $\mathbb{1}_A$ be constant on every set on which $X^S$ is constant, $A \in \mathcal{F}^X_\infty$. Then $\mathbb{1}_A = F \circ X$ for some $\mathcal{E}^{\otimes T}/\mathcal{B}(\{0,1\})$-measurable mapping $F$. Next, from $\mathrm{Im}X^S \subset \mathrm{Im}X$, for any $\omega \in \Omega$, there is an $\omega' \in \Omega$ with $X(\omega') = X^S(\omega)$, and then thanks to Lemma 6.1, $X^S(\omega') = X^S(\omega)$, so that by assumption, $\mathbb{1}_A(\omega) = \mathbb{1}_A(\omega')$, also. Moreover, $F \circ X^S(\omega) = F \circ X(\omega') = \mathbb{1}_A(\omega') = \mathbb{1}_A(\omega)$. It follows that $\mathbb{1}_A = F \circ X^S$.

Assume now (2). Again apply Blackwell's Theorem. Specifically, on account of (2)(a) $(\Omega, \mathcal{G})$ is a Blackwell space and $\mathcal{F}^X_S$ is a sub-$\sigma$-field of $\mathcal{G}$; on account of (2)(b), $\sigma(X^S)$ is a separable sub-$\sigma$-field of $\mathcal{G}$. Finally, if $\mathbb{1}_A$ is constant on every set on which $X^S$ is constant and $A \in \mathcal{F}^X_\infty$, then $\mathbb{1}_A$ is a $\mathcal{G}$-measurable function (by (2)(a)), constant on every atom of $\sigma(X^S)$ (by (2)(c)). It follows that $\mathbb{1}_A$ is $\sigma(X^S)$-measurable.

The implication (iii)⇒(i) is just one of the assumptions. □

As for our original motivation into this investigation, we obtain:

**Corollary 6.7** (Observational consistency)**.** *Let $X$ and $Y$ be two processes (on $\Omega$, with time domain $T$ and values in $E$), $S$ an $\mathcal{F}^X$ and an $\mathcal{F}^Y$-stopping time. Suppose furthermore $X^S = Y^S$. If any one of the conditions*

*(1)* $T = \mathbb{N}_0$.

*(2)* $\mathrm{Im}X = \mathrm{Im}Y$.

*(3) (a)* $(\Omega, \mathcal{G})$ *(resp.* $(\Omega, \mathcal{H})$*) is Blackwell for some $\sigma$-field $\mathcal{G} \supset \mathcal{F}^X_\infty$ (resp. $\mathcal{H} \supset \mathcal{F}^Y_\infty$).*

*(b)* $\sigma(X^S)$ *(resp.* $\sigma(Y^S)$*) is separable and contained in $\mathcal{F}^X_S$ (resp. $\mathcal{F}^Y_S$).*

*(c)* $(\mathrm{Im}X^S, \mathcal{E}^{\otimes T}|_{\mathrm{Im}X^S})$ *(resp.* $(\mathrm{Im}Y^S, \mathcal{E}^{\otimes T}|_{\mathrm{Im}Y^S})$*) is Hausdorff.*

*is met, then $\mathcal{F}^X_S = \mathcal{F}^Y_S$.*

*Remark* 6.8. If $T = \mathbb{N}_0$, then in place of $S$ being a stopping time of both $\mathcal{F}^X$ and $\mathcal{F}^Y$, it is sufficient (*ceteris paribus*) to insist on $S$ being a stopping time of just one of them. It is so by Lemma 6.2. The same is true when (3) obtains, as long as the conditions of Proposition 6.5 are met for the time $S$ and the processes $X$ and $Y$ alike.

*Proof.* If (1) or (3) hold, then the claim follows immediately from Theorem 6.6.

If (2) holds, let $A \in \mathcal{F}^X_S$, $t \in T \cup \{\infty\}$. Then $\mathbb{1}_{A\cap\{S\leq t\}} = F \circ X|_{[0,t]}$ (resp. $\mathbb{1}_{A\cap\{S\leq t\}} = F \circ X$) for some $\mathcal{E}^{\otimes[0,t]}/\mathcal{B}(\{0,1\})$-measurable (resp. $\mathcal{E}^{\otimes T}/\mathcal{B}(\{0,1\})$-measurable) $F$, when $t < \infty$ (resp. $t = \infty$). Moreover, if $\omega \in \Omega$, there is an $\omega' \in \Omega$ with $X(\omega') = Y(\omega)$, hence $X(\omega)$ agrees with $Y(\omega)$ and $X(\omega')$ on $T \cap [0, S(\omega)]$, and thus thanks to Lemma 6.1, $S(\omega) = S(\omega')$ and $\mathbb{1}_A(\omega) = \mathbb{1}_A(\omega')$. We obtain $F \circ Y|_{[0,t]}(\omega) = F \circ X|_{[0,t]}(\omega') = \mathbb{1}_{A\cap\{S\leq t\}}(\omega') = \mathbb{1}_{A\cap\{S\leq t\}}(\omega)$, i.e. $\mathbb{1}_{A\cap\{S\leq t\}} = F \circ Y|_{[0,t]}$ (resp. $F \circ Y(\omega) = F \circ X(\omega') = \mathbb{1}_{A\cap\{S\leq t\}}(\omega') = \mathbb{1}_{A\cap\{S\leq t\}}(\omega)$, i.e. $\mathbb{1}_{A\cap\{S\leq t\}} = F \circ Y$). □

We also have:

**Proposition 6.9** (Monotonicity of information)**.** *Let $Z$ be a process (on $\Omega$, with time domain $T$ and values in $E$), $U \leq V$ two stopping times of $\mathcal{F}^Z$. If either $T = \mathbb{N}_0$ or else the conditions:*

1. *$(\Omega, \mathcal{G})$ is Blackwell for some $\sigma$-field $\mathcal{G} \supset \sigma(Z^V) \vee \sigma(Z^U)$.*
2. *$(\mathrm{Im}Z^V, \mathcal{E}^{\otimes T}|_{\mathrm{Im}Z^V})$ is Hausdorff.*
3. *$\sigma(Z^V)$ is separable.*

*are met, then $\sigma(Z^U) \subset \sigma(Z^V)$.*

*Proof.* In the discrete case the result follows at once from Theorem 6.6. In the opposite instance, we claim that the assumptions imply that every atom of $\sigma(Z^U)$ is a union of the atoms of $\sigma(Z^V)$: Let $\omega$ and $\omega'$ belong to the same atom of $\sigma(Z^V)$; then since $(\mathrm{Im}Z^V, \mathcal{E}^{\otimes T}|_{\mathrm{Im}Z^V})$ is Hausdorff $Z^V(\omega) = Z^V(\omega')$, hence by Lemma 6.1 $V(\omega) = V(\omega')$ and $U(\omega) = U(\omega')$, and so *a fortiori* $Z^U(\omega) = Z^U(\omega')$, which implies that $\omega$ and $\omega'$ belong to the same atom of $\sigma(Z^U)$. Apply Blackwell's Theorem. □

## 7. The case with completions

We have studied in the previous section natural filtrations *proper* — it is sometimes convenient to augment the latter by sets of probability zero[7] — we turn our attention to their completions. Notation-wise, for a filtration $\mathcal{G}$ on $\Omega$ and a *complete* probability measure $\mathsf{P}$, whose domain includes $\mathcal{G}_\infty$, thereon, we denote by $\overline{\mathcal{G}}^{\mathsf{P}}$ the **completed filtration** given by (for $t \in T$) $\overline{\mathcal{G}}^{\mathsf{P}}_t = \overline{\mathcal{G}_t}^{\mathsf{P}} = \mathcal{G}_t \vee \mathcal{N}$; $\mathcal{N}$ being the collection of precisely all $\mathsf{P}$-null sets; likewise if the domain of $\mathsf{P}$ includes a $\sigma$-field $\mathcal{A}$ on $\Omega$, then $\overline{\mathcal{A}}^{\mathsf{P}} := \mathcal{A} \vee \mathcal{N}$. For any other unexplained notation, that we shall use, we refer the reader to the beginning of Section 6. And while /for ease of language/ we will continue to work in the sequel with processes/stopping times, their equivalence classes (with respect to indistinguishability/a.s. equality), would of course (as appropriate) suffice.

First, all is well in the discrete case.

**Lemma 7.1.** *Let $T = \mathbb{N}_0$, $\mathcal{G}$ a filtration on $\Omega$. Let furthermore $\mathsf{P}$ be a complete probability measure on $\Omega$, whose domain includes $\mathcal{G}_\infty$; $S$ a $\overline{\mathcal{G}}^{\mathsf{P}}$-stopping time. Then $S$ is $\mathsf{P}$-a.s. equal to a stopping time $S'$ of $\mathcal{G}$; and for any $\mathcal{G}$-stopping time $U$, $\mathsf{P}$-a.s. equal to $S$, $\overline{\mathcal{G}_U}^{\mathsf{P}} = \overline{\mathcal{G}}^{\mathsf{P}}_S$. Moreover, if $U$ is another random time, $\mathsf{P}$-a.s equal to $S$, then it is a $\overline{\mathcal{G}}^{\mathsf{P}}$-stopping time, and $\overline{\mathcal{G}}^{\mathsf{P}}_S = \overline{\mathcal{G}}^{\mathsf{P}}_U$.*

*Proof.* For each $n \in \mathbb{N}_0$, we may find an $A_n \in \mathcal{G}_n$, such that $\{S = n\} = A_n$, $\mathsf{P}$-a.s. Then $S' := (\cup_{n\in\mathbb{N}_0} A_n \times \{n\}) \cup ((\Omega \backslash \cup_{m\in\mathbb{N}_0} A_m) \times \{\infty\})$ is a $\mathcal{G}$-stopping time, $\mathsf{P}$-a.s. equal to $S$. Let now

[7] Making them also (in the temporally continuous case, if they are not automatically already) right-continuous, is less interesting from the point of view of stochastic control, since the stopping times one is really interested in are (usually) foretellable/predictable, anyway. In general, this is also less of an innocuous operation. For, one might well concede to being unable to act on a null set; one cannot but feel apprehensive about having to 'peak infinitesimally into the future' before being able to act in the present. Indeed, we will see in the sequel that even the act of completion alone is less harmless than might seem at first glance.

$U$ be any time with these two properties of $S'$. To show $\overline{\mathcal{G}_U}^{\mathsf{P}} \subset \overline{\mathcal{G}}^{\mathsf{P}}_S$, it suffices to note that (i) $\mathcal{N}$, the collection of all $\mathsf{P}$-null sets, is contained in $\overline{\mathcal{G}}^{\mathsf{P}}_S$ and (ii) $\mathcal{G}_U \subset \overline{\mathcal{G}}^{\mathsf{P}}_S$, both of which are easy to see. Conversely, if $A \in \overline{\mathcal{G}}^{\mathsf{P}}_S$, then for each $n \in \mathbb{N}_0 \cup \{\infty\}$, $A \cap \{S = n\} = B_n$, $\mathsf{P}$-a.s., for some $B_n \in \mathcal{G}_n$, and hence the event $\cup_{n\in\mathbb{N}_0\cup\{\infty\}} B_n \cap \{U = n\}$ belongs to $\mathcal{G}_U$, and is $\mathsf{P}$-a.s. equal to $A$.

Finally, let $U$ be another random time, $\mathsf{P}$-a.s equal to $S$. For each $n \in \mathbb{N}_0 \cup \{\infty\}$, there is then a $C_n \in \mathcal{G}_n$ with $\{U = n\} = \{S = n\} = C_n$, $\mathsf{P}$-a.s., whence $U$ is a $\overline{\mathcal{G}}^{\mathsf{P}}$-stopping time. It follows, by what we have shown already, that we can find $Z$, a $\mathcal{G}$-stopping time, $\mathsf{P}$-a.s. equal to $S$, hence $U$, and thus with $\overline{\mathcal{G}}^{\mathsf{P}}_S = \overline{\mathcal{G}_Z}^{\mathsf{P}} = \overline{\mathcal{G}}^{\mathsf{P}}_U$. □

**Corollary 7.2.** *Let $T = \mathbb{N}_0$; $X$ and $Y$ processes (on $\Omega$, with time domain $\mathbb{N}_0$ and values in $E$); $\mathsf{P}^X$ and $\mathsf{P}^Y$ be complete probability measures on $\Omega$ whose domains contain $\mathcal{F}^X_\infty$ and $\mathcal{F}^Y_\infty$, respectively, and sharing their null sets. Suppose furthermore $S$ is an $\overline{\mathcal{F}^X}^{\mathsf{P}^X}$ and an $\overline{\mathcal{F}^Y}^{\mathsf{P}^Y}$ stopping time, with $X^S = Y^S$, $\mathsf{P}^X$ and $\mathsf{P}^Y$-a.s. Then $\overline{\mathcal{F}^X}^{\mathsf{P}^X}_S = \overline{\sigma(X^S)}^{\mathsf{P}^X} = \overline{\sigma(Y^S)}^{\mathsf{P}^Y} = \overline{\mathcal{F}^Y}^{\mathsf{P}^Y}_S$.*[8]

*Proof.* From Lemma 7.1 we can find stopping times $U$ and $V$ of $\mathcal{F}^X$ and $\mathcal{F}^Y$, respectively, both $\mathsf{P}^X$ and $\mathsf{P}^Y$-a.s. equal to $S$. The event $\{X^U = Y^V\}$ is $\mathsf{P}^X$ and $\mathsf{P}^Y$-almost certain. It then follows further from Theorem 6.6 and Lemma 7.1 again, that $\overline{\mathcal{F}^X}^{\mathsf{P}^X}_S = \overline{\mathcal{F}^X_U}^{\mathsf{P}^X} = \overline{\sigma(X^U)}^{\mathsf{P}^X} = \overline{\sigma(X^S)}^{\mathsf{P}^X} = \overline{\sigma(Y^S)}^{\mathsf{P}^Y} = \overline{\sigma(Y^V)}^{\mathsf{P}^Y} = \overline{\mathcal{F}^Y_V}^{\mathsf{P}} = \overline{\mathcal{F}^Y}^{\mathsf{P}^Y}_S$, as desired. □

**Corollary 7.3.** *Let $T = \mathbb{N}_0$, $X$ a process (on $\Omega$, with time domain $\mathbb{N}_0$ and values in $E$), $\mathsf{P}$ a complete probability measure on $\Omega$ whose domain contains $\mathcal{F}^X_\infty \vee \mathcal{F}^{X^S}_\infty$, $S : \Omega \to T \cup \{\infty\}$ a random time. Then the following are equivalent:*

*(1) $S$ is an $\overline{\mathcal{F}^X}^{\mathsf{P}}$-stopping time.*
*(2) $S$ is an $\overline{\mathcal{F}^{X^S}}^{\mathsf{P}}$-stopping time.*

*Proof.* That the first implies the second is clear from Lemma 7.1, Lemma 6.2 and the fact that two processes, which are versions of each other, generate the same filtration, up to null sets. For the converse, one resorts to re-doing the relevant part of the proof of Lemma 6.2, adding $\mathsf{P}$-a.s. qualifiers as appropriate; the details are left to the reader. □

The temporally continuous case is much more involved. Indeed, we have the following significant negative results.

*Example* 7.4. Let $\Omega = (0, \infty) \times \{0, 1\}$; $\mathcal{F}$ be the product of the Lebesgue $\sigma$-field on $(0, \infty)$ and of the power set on $\{0, 1\}$; thereon $\mathsf{P} = \mathrm{Exp}(1) \times \mathrm{Unif}(\{0, 1\})$ be the product law (which is complete; any law on the first coordinate with a continuous distribution function would also do); $e$ (respectively $a$) be the projection onto the first (respectively second) coordinate. Define the process $N_t = a(t - e)\mathbb{1}_{[0,t]}(e)$, $t \in [0, \infty)$ (starting at zero, the process $N$ departs from zero at time $e$ with unit

[8]Of course, all these completions really only depend on the null sets, which the two measures $\mathsf{P}^X$ and $\mathsf{P}^Y$ share by assumption.

positive drift, or remains at zero, for all times, with equal probability, independently of $e$). Its completed natural filtration, $\overline{\mathcal{F}^N}^{\mathsf{P}}$, is already right-continuous.

For, if $t \in [0,\infty)$, $\overline{\mathcal{F}^N}^{\mathsf{P}}_{t+} = \overline{\mathcal{F}^N_{t+}}^{\mathsf{P}}$; so let $A \in \mathcal{F}^N_{t+}$, we show $A \in \overline{\mathcal{F}^N_t}^{\mathsf{P}}$. (i) $A \cap \{e = t\}$ is $\mathsf{P}$-negligible. (ii) For sure $A = N^{-1}(G)$, for some $G \subset \mathbb{R}^{[0,\infty)}$, measurable. Then define for each natural $n \geq 1/t$ (when $t > 0$), $L_n : \mathbb{R}^{[0,t]} \to \mathbb{R}^{[0,+\infty)}$, by demanding

$$L_n(\omega)(u) = \begin{cases} \omega(u), & \text{for } u \leq t \\ \omega(t) + (u-t)\frac{\omega(t)-\omega(t-1/n)}{1/n}, & \text{for } u > t \end{cases}$$

($u \in [0,\infty)$, $\omega \in \mathbb{R}^{[0,t]}$), a measurable mapping. It follows that for $t > 0$, for each natural $n \geq 1/t$, $N^{-1}(G) \cap \{e \leq t - 1/n\} = N|_{[0,t]}^{-1}(L_n^{-1}(G)) \cap \{e \leq t-1/n\} \in \mathcal{F}^N_t$. (iii) For each natural $n$, $A \cap \{e > t + 1/n\} = N|_{[0,t+1/n]}^{-1}(G_n) \cap \{e > t+1/n\}$ for some measurable $G_n \subset \mathbb{R}^{[0,t+1/n]}$, so $A \cap \{e > t+1/n\}$ is $\emptyset$ or $\{e > t + 1/n\}$ according as 0 is an element of $G_n$ or not (note this is a "monotone" condition, in the sense that as soon as we once get a non-empty set for some natural $n$, we subsequently get $\{e > t + 1/m\}$ for all natural $m \geq n$). It follows that $A \cap \{e > t\} = \cup_{n\in\mathbb{N}}(A \cap \{e > t + 1/n\}) \in \{\emptyset, \{e > t\}\} \subset \mathcal{F}^N_t$.

Let further $U$ be the first entrance time of the process $N$ to $(0,\infty)$. By the Début Theorem, this is a stopping time of $\overline{\mathcal{F}^N}^{\mathsf{P}}$, but is $\mathsf{P}$-a.s. equal to no stopping time of $\mathcal{F}^N$ at all.

For, suppose that it were $\mathsf{P}$-a.s. equal to a stopping time $V$ of $\mathcal{F}^N$. Then there would be a set $\Omega'$, belonging to $\mathcal{F}$, of full $\mathsf{P}$-measure, and such that $V = U$ on $\Omega'$. Tracing everything ($\mathcal{F}$, $\mathsf{P}$, $N$, $a$, $e$, $V$) onto $\Omega'$, we would obtain ($\mathcal{F}'$, $\mathsf{P}'$, $N'$, $a'$, $e'$, $V'$), with (i) $V'$ equal to the first entrance time of $N'$ to $(0,\infty)$ and (ii) $V'$ a stopping time of $\mathcal{F}^{N'}$, the natural filtration of $N'$. Still $N'_t = a'(t-e')\mathbb{1}_{[0,t]}(e')$, $t \geq 0$. Take now $\{\omega,\omega'\} \subset \Omega'$ with $a(\omega) = 1$, $a(\omega') = 0$, denote $t := e(\omega)$. Then $H|_{[0,t]}(\omega) = H|_{[0,t]}(\omega')$, so $\omega$ and $\omega'$ should belong to the same atom of $\mathcal{F}^{N'}_t$; yet $\{V' \leq t\} \in \mathcal{F}^{N'}_t$, with $\mathbb{1}_{\{V'\leq t\}}(\omega) = 1$ and $\mathbb{1}_{\{V'\leq t\}}(\omega') = 0$, a contradiction.

Moreover, $\overline{\mathcal{F}^N}^{\mathsf{P}}_U \neq \overline{\sigma(N^U)}^{\mathsf{P}}$, since the event $A := \{U < \infty\} = \{a = 1\}$ that $N$ ever assumes a positive drift belongs to $\overline{\mathcal{F}^N}^{\mathsf{P}}_U$ (which fact is clear), but not to $\overline{\sigma(N^U)}^{\mathsf{P}} = \overline{\sigma(0)}^{\mathsf{P}}$, the trivial $\sigma$-field (it is also obvious; $\mathsf{P}(a = 1) \notin \{0,1\}$). ◇

*Example* 7.5. It is worse, still. Let $\Omega = (0,\infty) \times \{-2,-1,0\}$ be endowed with the law $\mathsf{P} = \mathrm{Exp}(1) \times \mathrm{Unif}(\{-2,-1,0\})$, defined on the tensor product of the Lebesgue $\sigma$-field on $(0,\infty)$ and the power set of $\{-2,-1,0\}$. Denote by $e$, respectively $I$, the projection onto the first, respectively second, coordinate. Define the process $X_t := I(t-e)\mathbb{1}_{[0,t]}(e)$ , $t \in [0,\infty)$, and the process $Y_t := (-1)(t-e)\mathbb{1}_{[0,t]}(e)\mathbb{1}_{\{-1,-2\}} \circ I$, $t \in [0,\infty)$. The completed natural filtrations of $X$ and $Y$ are already right-continuous. The first entrance time $S$ of $X$ into $(-\infty,0)$ is equal to the first entrance time of $Y$ into $(-\infty,0)$, and this is a stopping time of $\overline{\mathcal{F}^X}^{\mathsf{P}}$ as it is of $\overline{\mathcal{F}^Y}^{\mathsf{P}}$ (but not of $\mathcal{F}^X$ and not of $\mathcal{F}^Y$). Moreover, $X^S = 0 = Y^S$.

Consider now the event $A := \{I = -1\}$. Then $A \in \overline{\mathcal{F}^X}^{\mathsf{P}}_S$ (it is clear). However, $A \notin \overline{\mathcal{F}^Y}^{\mathsf{P}}_S$. For, assuming the converse, we should have, $\mathsf{P}$-a.s., $\mathbb{1}_{A\cap\{S\leq 1\}} = F \circ Y|_{[0,1]}$ for some, measurable, $F$. In particular, since $A \cap \{S \leq 1\}$ has positive probability, there should be an $\omega \in A \cap \{S \leq 1\}$ with $F(Y|_{[0,1]}(\omega)) = 1$. But also the event $\{I = -2\} \cap \{S \leq 1\}$ has positive probability and is disjoint from $A \cap \{S \leq 1\}$, so there should be an $\omega' \in \{I = -2\} \cap \{S \leq 1\}$ having $F(Y|_{[0,1]}(\omega')) = 0$. A contradiction, since nevertheless $Y|_{[0,1]}(\omega') = Y|_{[0,1]}(\omega)$. ◇

The problem here, as it were, is that in completing the natural filtration the (seemingly innocuous) operation of adding all the events negligible under $\mathsf{P}$ is done uncountably many times (once for every deterministic time). In particular, this cannot be recovered by a single completion of the sigma-field generated by the stopped process. Completions are not always harmless.

Furthermore, it does not appear immediately clear to us, what a sensible *direct* 'probabilistic' analogue of Lemma 6.1 should be /nor, indeed, how to go about proving one, and then using it to produce the relevant counter-parts to the results of Section 6/.

However, the situation is not so bleak, since positive results can be got at least for foretellable/predictable stopping times: As in the case of discrete time – by an *indirect* method; reducing the 'probabilistic' to the 'measure-theoretic' case. We use here the terminology of [4, p. 127, Definitions IV.69 & IV.70]; given a filtration $\mathcal{G}$ and a probability measure $\mathsf{Q}$ on $\Omega$, whose domain includes $\mathcal{G}_\infty$:

> A random time $S : \Omega \to [0, \infty]$ is **predictable** relative to $\mathcal{G}$ if the stochastic interval $[\![T, \infty[\![$ is predictable. It is **$\mathsf{Q}$-foretellable** relative to $\mathcal{G}$ if there exists a $\mathsf{Q}$-a.s. nondecreasing sequence $(S_n)_{n\geq 1}$ of $\mathcal{G}$-stopping times with $S_n \leq S$, $\mathsf{Q}$-a.s for all $n \geq 1$ and such that, again $\mathsf{Q}$-a.s.,
>
> $$\lim_{n\to\infty} S_n = S, S_n < S \text{ for all } n \text{ on } \{S > 0\};$$
>
> **foretellable**, if the a.s. qualifications can be omitted.

Note that in a $\mathsf{P}$-complete filtration ($\mathsf{P}$ itself assumed complete), the notions of predictable, foretellable and $\mathsf{P}$-foretellable stopping times coincide [4, p. 127, IV.70; p. 128, Theorem IV.71 & p. 132, Theorem IV.77].

The following is now a complement to [4, p. 120, Theorem IV.59 & p. 133, Theorem IV.78] [9, p. 5, Lemma 1.19], and an analogue of the discrete statement of Lemma 7.1:

**Proposition 7.6.** *Let $T = [0, \infty)$, $\mathcal{G}$ be a filtration on $\Omega$. Let furthermore $\mathsf{P}$ be a complete probability measure on $\Omega$, whose domain includes $\mathcal{G}_\infty$; $S$ a predictable stopping time relative to $\overline{\mathcal{G}}^{\mathsf{P}}$. Then $S$ is $\mathsf{P}$-a.s. equal to a predictable stopping time $P$ of $\mathcal{G}$. Moreover, if $U$ is any $\mathcal{G}$-stopping time, $\mathsf{P}$-a.s. equal to $S$, then $\overline{\mathcal{G}}^{\mathsf{P}}_S = \overline{\mathcal{G}_U}^{\mathsf{P}}$. Finally, if $S'$ is another random time, $\mathsf{P}$-a.s equal to $S$, then it is a $\overline{\mathcal{G}}^{\mathsf{P}}$-stopping time, and $\overline{\mathcal{G}}^{\mathsf{P}}_S = \overline{\mathcal{G}}^{\mathsf{P}}_{S'}$.*

*Proof.* The first claim is contained in [4, p. 133, Theorem IV.78].

Now let $U$ be any $\mathcal{G}$-stopping time, $\mathsf{P}$-a.s. equal to $S$. The inclusion $\overline{\mathcal{G}}^{\mathsf{P}}_S \supset \overline{\mathcal{G}_U}^{\mathsf{P}}$ is obvious. Then take $A \in \overline{\mathcal{G}}^{\mathsf{P}}_S$. Since $A \in \overline{\mathcal{G}}^{\mathsf{P}}_\infty = \overline{\mathcal{G}_\infty}^{\mathsf{P}}$, there is an $A' \in \mathcal{G}_\infty$, such that $A' = A$, $\mathsf{P}$-a.s. Furthermore, since $S$ is foretellable,

$$S_A := \begin{cases} S & \text{on } A \\ \infty & \text{on } \Omega\backslash A \end{cases}$$

is foretellable also (if $(S_n)_{n\geq 1}$ $\mathsf{P}$-foretells $S$, then $((S_n)_A \wedge n)_{n\geq 1}$ $\mathsf{P}$-foretells $S_A$). Hence, by what we have already shown, there exists $V$, a $\mathcal{G}$-stopping time, with $V = S_A$, $\mathsf{P}$-a.s. So, $\mathsf{P}$-a.s., $A = (A' \cap \{U = \infty\}) \cup \{V = U < \infty\} \in \mathcal{G}_U$.

Finally let $S'$ be a random time, $\mathsf{P}$-a.s. equal to $S$. Clearly, it is a $\overline{\mathcal{G}}^{\mathsf{P}}$-stopping time. Moreover, we have found $P$, a $\mathcal{G}$-stopping time, $\mathsf{P}$-a.s. equal to $S$ and (by way of corollary) $S'$. It follows from what we have just shown that $\overline{\mathcal{G}}^{\mathsf{P}}_S = \overline{\mathcal{G}_P}^{\mathsf{P}} = \overline{\mathcal{G}}^{\mathsf{P}}_{S'}$. □

From this we can obtain easily a couple of useful counter-parts to the findings of Section 6 in the continuous case. They (Corollaries 7.7, 7.9 and 7.10 that follow) should be used in conjunction with (in this order) (i) the fact that a standard Borel space[9]-valued random element measurable with respect to the completed domain of the probability measure $\mathsf{Q}$, is $\overline{\mathsf{Q}}$-a.s. equal to a random element measurable with respect to the uncompleted domain of $\mathsf{Q}$ ($\overline{\mathsf{Q}}$ being the completion of $\mathsf{Q}$) [10, p. 13, Lemma 1.25] and (ii) the existence part of Proposition 7.6. Loosely speaking one imagines working on the completion of a nice (Blackwell) space. Then the quantities measurable with respect to the completed sigma-fields are a.s. equal to quantities measurable with respect to the uncompleted sigma-fields, and to them the 'measure-theoretic' results apply. Taking completions again, we arrive at the relevant 'probabilistic' statements. The formal results follow.

**Corollary 7.7.** *Let $T = [0,\infty)$, $Z$ a process and $\mathsf{P}$ a complete probability measure on $\Omega$, whose domain includes $\mathcal{F}^Z_\infty$, $P$ an $\overline{\mathcal{F}^Z}^{\mathsf{P}}$ predictable stopping time. If further for some process $X$ $\mathsf{P}$-indistinguishable from $Z$ and a stopping time $S$ of $\mathcal{F}^X$, $\mathsf{P}$-a.s. equal to $P$:*

*(1) $(\Omega, \mathcal{G})$ is Blackwell for some σ-field $\mathcal{G} \supset \mathcal{F}^X_S \vee \sigma(X^S)$.*

*(2) $\sigma(X^S)$ is separable.*

*(3) $(\mathrm{Im}X^S, \mathcal{E}^{\otimes[0,\infty)}|_{\mathrm{Im}X^S})$ is Hausdorff.*

*then $\overline{\mathcal{F}^Z}^{\mathsf{P}}_P \subset \overline{\sigma(Z^P)}^{\mathsf{P}}$.*

*Remark* 7.8. The reverse inclusion $\overline{\sigma(Z^P)} \subset \overline{\mathcal{F}^Z}^{\mathsf{P}}$ is usually trivial (compare the remarks on this in the second bullet point entry of Section 6, p. 22).

*Proof.* According to Proposition 6.3, $\mathcal{F}^X_S \subset \sigma(X^S)$. Also $\overline{\mathcal{F}^X}^{\mathsf{P}} = \overline{\mathcal{F}^Z}^{\mathsf{P}}$ and $\overline{\sigma(X^S)}^{\mathsf{P}} = \overline{\sigma(Z^P)}^{\mathsf{P}}$. Taking completions in $\mathcal{F}^X_S \subset \sigma(X^S)$, we obtain by Proposition 7.6, as applied to the stopping time $P$ of $\overline{\mathcal{F}^X}^{\mathsf{P}}$, $\mathsf{P}$-a.s. equal to the stopping time $S$ of $\mathcal{F}^X$:

$$\overline{\mathcal{F}^Z}^{\mathsf{P}}_P = \overline{\mathcal{F}^X}^{\mathsf{P}}_P = \overline{\mathcal{F}^X_S}^{\mathsf{P}} \subset \overline{\sigma(X^S)}^{\mathsf{P}} = \overline{\sigma(Z^P)}^{\mathsf{P}},$$

[9]One that is Borel isomorphic to a Borel subset of $[0,1]$ [10, p. 7], equivalently a Borel subset of a Polish space [2, p. 12, Definition 6.2.10]. Recall the spaces of Euclidean space-valued càdlàg (resp. continuous) paths endowed with the Skorohod topology [9, Section VI.1] (resp. the topology of locally uniform convergence [11, p. 60]) are Polish, hence standard Borel, spaces.

as desired. $\square$

**Corollary 7.9.** *Let $Z$ and $W$ be two processes (on $\Omega$, with time domain $[0,\infty)$ and values in $E$); $\mathsf{P}^Z$ and $\mathsf{P}^W$ probability measures on $\Omega$, sharing their null sets, and whose domain includes $\mathcal{F}^Z_\infty$ and $\mathcal{F}^W_\infty$, respectively; $P$ a predictable $\overline{\mathcal{F}^Z}^{\mathsf{P}^Z}$ and $\overline{\mathcal{F}^W}^{\mathsf{P}^W}$-stopping time. Suppose furthermore $Z^P = W^P$, $\mathsf{P}^Z$ and $\mathsf{P}^W$-a.s. If for two processes $X$ and $Y$, $\mathsf{P}^Z$ and $\mathsf{P}^W$-indistinguishable from $Z$ and $W$, respectively, and some stooping times $S$ and $U$ of $\mathcal{F}^X$ and $\mathcal{F}^Y$, respectively, $\mathsf{P}^Z$ and $\mathsf{P}^W$-a.s. equal to $P$:*

*(1) $(\Omega,\mathcal{G})$ (resp. $(\Omega,\mathcal{H})$) is Blackwell for some $\sigma$-field $\mathcal{G}\supset\mathcal{F}^X_\infty$ (resp. $\mathcal{H}\supset\mathcal{F}^Y_\infty$).*

*(2) $\sigma(X^S)$ (resp. $\sigma(Y^U)$) is separable and contained in $\mathcal{F}^X_S$ (resp. $\mathcal{F}^Y_U$).*

*(3) $(\mathrm{Im}X^S,\mathcal{E}^{\otimes[0,\infty)}|_{\mathrm{Im}X^S})$ (resp. $(\mathrm{Im}Y^U,\mathcal{E}^{\otimes[0,\infty)}|_{\mathrm{Im}Y^U})$) is Hausdorff.*

*then $\overline{\mathcal{F}^Z}^{\mathsf{P}^Z}_P = \overline{\sigma(Z^P)}^{\mathsf{P}^Z} = \overline{\sigma(W^P)}^{\mathsf{P}^W} = \overline{\mathcal{F}^W}^{\mathsf{P}^W}_P$.*

*Proof.* The claim follows from Corollary 7.7, and the fact that again (similarly as in the proof of Corollary 7.7) $\sigma(X^S)\subset\mathcal{F}^X_S$ implies $\overline{\sigma(Z^P)}^{\mathsf{P}^Z}\subset\overline{\mathcal{F}^Z}^{\mathsf{P}^Z}_P$; likewise for $W$. $\square$

**Corollary 7.10.** *Let $X$ be a process (on $\Omega$, with time domain $[0,\infty)$ and values in $E$); $\mathsf{P}$ a complete probability measure on $\Omega$, whose domain includes $\mathcal{F}^X_\infty$; $S$ and $P$ two predictable stopping times of $\overline{\mathcal{F}^X}^{\mathsf{P}}$ with $S\le P$. Let $U$ and $V$ be two stopping times of the natural filtration of a process $Z$, $\mathsf{P}$-indistinguishable from $X$, $\mathsf{P}$-a.s. equal to $S$ and $P$, respectively, with $U\le V$, and such that: $(\Omega,\mathcal{G})$ is Blackwell for some $\sigma$-field $\mathcal{G}\supset\sigma(Z^V)\vee\sigma(Z^U)$, $(\mathrm{Im}Z^V,\mathcal{E}^{\otimes T}|_{\mathrm{Im}Z^V})$ is Hausdorff and $\sigma(Z^V)$ is separable. Then $\overline{\sigma(X^S)}^{\mathsf{P}}\subset\overline{\sigma(X^P)}^{\mathsf{P}}$.*

*Remark* 7.11. the existence part of Proposition 7.6 here, is as follows: There are $U$ and $V'$, stopping times of $\mathcal{F}^X$, $\mathsf{P}$-a.s. equal to $S$ and $P$, respectively. Then, true, $U\le V'$ only $\mathsf{P}$-a.s. But $V := V'\mathbb{1}(U\le V') + U\mathbb{1}(U>V')$ is also a stopping time of $\mathcal{F}^X$, $\mathsf{P}$-a.s. equal to $P$, and *it* satisfies $U\le V$ with certainty.

*Proof.* We find that $\overline{\sigma(X^S)}^{\mathsf{P}} = \overline{\sigma(Z^U)}^{\mathsf{P}}$ and $\overline{\sigma(X^P)}^{\mathsf{P}} = \overline{\sigma(Z^V)}^{\mathsf{P}}$. Then apply Proposition 6.9. $\square$

We are not able to provide a (in conjunction with Proposition 7.6) *useful* counter-part to Proposition 6.5. (True, Proposition 7.6 says that given a predictable stopping time $P$ of $\overline{\mathcal{F}^{X^P}}^{\mathsf{P}}$, there is a predictable stopping time $U$ of $\mathcal{F}^{X^P}$, $\mathsf{P}$-a.s. equal to $P$. But this is not to say $U$ is a stopping time of $\mathcal{F}^{X^U}$, so one cannot directly apply Proposition 6.5.) This is open to future research.

## References


1. A. Bensoussan. *Stochastic Control of Partially Observable Systems.* Cambridge University Press, 2004.
2. V. I. Bogachev. *Measure Theory*, volume 2 of *Measure Theory.* Springer, 2007.
3. Z. Brzeźniak and S. Peszat. Space-time continuous solutions to SPDE's driven by a homogeneous Wiener process. *Studia Mathematica*, 137(3):261–299, 1999.
4. C. Dellacherie and P. A. Meyer. *Probabilities and Potential A.* North-Holland Mathematics Studies. Herman Paris, 1978.

5. N. El Karoui. Les Aspects Probabilistes Du Contrôle Stochastique. In P. L. Hennequin, editor, *Ecole d'Eté de Probabilités de Saint-Flour IX-1979*, volume 876 of *Lecture Notes in Mathematics*, chapter 2, pages 73–238. Springer Berlin / Heidelberg, 1981.
6. N. El Karoui, D. Nguyen, and M. Jeanblanc-Picqu. Existence of an Optimal Markovian Filter for the Control under Partial Observations. *SIAM Journal on Control and Optimization*, 26(5):1025–1061, 1988.
7. W. Fleming and E. Pardoux. Optimal control for partially observed diffusions. *SIAM Journal on Control and Optimization*, 20(2):261–285, 1982.
8. R. G. Fryer and P. Harms. Two-Armed Restless Bandits with Imperfect Information: Stochastic Control and Indexability. NBER Working Papers 19043, National Bureau of Economic Research, Inc, May 2013.
9. J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, Berlin Heidelberg, 2003.
10. O. Kallenberg. *Foundations of Modern Probability*. Probability and Its Applications. Springer, New York Berlin Heidelberg, 1997.
11. I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics. Springer New York, 1991.
12. P. E. Protter. *Stochastic Integration and Differential Equations: Version 2.1*. Applications of mathematics. U.S. Government Printing Office, 2004.
13. L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes, and Martingales: Volume 1, Foundations*. Cambridge Mathematical Library. Cambridge University Press, 2000.
14. H. M. Soner and N. Touzi. Dynamic programming for stochastic target problems and geometric flows. *Journal of the European Mathematical Society*, 4(3):201–236, 2002.
15. C. Striebel. Martingale conditions for the optimal control of continuous time stochastic systems. *Stochastic Processes and their Applications*, 18(2):329 – 347, 1984.
16. W. Wonham. On the separation theorem of stochastic control. *SIAM Journal on Control*, 6(2):312–326, 1968.

## Appendix A. Miscellaneous technical results

Throughout this appendix $(\Omega, \mathcal{F}, \mathsf{P})$ is a probability space; $\mathsf{E}$ denotes expectation with respect to $\mathsf{P}$.

**Lemma A.1** (On conditioning). *Let $X : \Omega \to [-\infty, +\infty]$ be a random variable, and $\mathcal{G}_i \subset \mathcal{F}$, $i = 1, 2$, two sub-$\sigma$-fields of $\mathcal{F}$ agreeing when traced on $A \in \mathcal{G}_1 \cap \mathcal{G}_2$. Then, $\mathsf{P}$-a.s. on $A$, $\mathsf{E}[X|\mathcal{G}_1] = \mathsf{E}[X|\mathcal{G}_2]$, whenever $X$ has a $\mathsf{P}$-integrable positive or negative part.*

*Proof.* $\mathbb{1}_A Z$ is $\mathcal{G}_2$-measurable, for any $Z$ $\mathcal{G}_1$-measurable, by an approximation argument. Then, $\mathsf{P}$-a.s., $\mathbb{1}_A \mathsf{E}[X|\mathcal{G}_1] = \mathsf{E}[\mathbb{1}_A X|\mathcal{G}_2]$, by the very definition of conditional expectation. □

**Lemma A.2** (Generalised conditional Fatou and Beppo Levi). *Let $\mathcal{G} \subset \mathcal{F}$ be a sub-$\sigma$-field and $(f_n)_{n\geq 1}$ a sequence of $[-\infty, +\infty]$-valued random elements, whose negative parts are dominated $\mathsf{P}$-a.s by a single $\mathsf{P}$-integrable random variable. Then, $\mathsf{P}$-a.s.,*

$$\mathsf{E}[\liminf_{n\to\infty} f_n|\mathcal{G}] \leq \liminf_{n\to\infty} \mathsf{E}[f_n|\mathcal{G}].$$

*If, moreover, $(f_n)_{n\geq 1}$ is $\mathsf{P}$-a.s. nondecreasing, then, $\mathsf{P}$-a.s.,*

$$\mathsf{E}[\lim_{n\to\infty} f_n|\mathcal{G}] = \lim_{n\to\infty} \mathsf{E}[f_n|\mathcal{G}].$$

*Proof.* Just apply conditional Fatou (resp. Beppo Levi) to the $\mathsf{P}$-a.s. nonnegative (resp. nonnegative nondecreasing) sequence $f_n + g$ where $g$ is the single $\mathsf{P}$-integrable random variable which $\mathsf{P}$-a.s. dominates the negative parts of $f$. Then use linearity and subtract the $\mathsf{P}$-a.s. finite quantity $\mathsf{E}[g|\mathcal{G}]$. □

The following is a slight generalization of [15, Theorem A2].

**Lemma A.3** (Essential supremum and the upwards lattice property)**.** *Let $\mathcal{G} \subset \mathcal{F}$ be a sub-$\sigma$-field and $X = (X_\lambda)_{\lambda\in\Lambda}$ a collection of $[-\infty, +\infty]$-valued random variables with integrable negative parts. Assume furthermore that for each $\{\epsilon, M\} \subset (0,\infty)$, $X$ has the "$(\epsilon, M)$-upwards lattice property", i.e. for all $\{\lambda, \lambda'\} \subset \Lambda$, one can find a $\lambda'' \in \Lambda$ with $X_{\lambda''} \geq (M \wedge X_\lambda) \vee (M \wedge X_{\lambda'}) - \epsilon$ $\mathsf{P}$-a.s. Then, $\mathsf{P}$-a.s.,*

$$\mathsf{E}[\mathsf{P}\text{-esssup}_{\lambda\in\Lambda} X_\lambda|\mathcal{G}] = \mathsf{P}\text{-esssup}_{\lambda\in\Lambda}\mathsf{E}[X_\lambda|\mathcal{G}], \tag{A.1}$$

*where on the right-hand side the essential supremum may of course equally well be taken with respect to the measure $\mathsf{P}|_{\mathcal{G}}$.*

*Proof.* It is assumed without loss of generality that $\Lambda \neq \emptyset$, whence remark that $\mathsf{P}\text{-esssup}_{\lambda\in\Lambda} X_\lambda$ has an integrable negative part. Then the "$\geq$-inequality" in (A.1) is immediate.

Conversely, we show first that it is sufficient to establish the "$\leq$-inequality" in (A.1) for each truncated $(X_\lambda \wedge N)_{\lambda\in\Lambda}$ family, as $N$ runs over $\mathbb{N}$. Indeed, suppose we have $\mathsf{P}$-a.s.

$$\mathsf{E}[\mathsf{P}\text{-esssup}_{\lambda\in\Lambda} X_\lambda \wedge N|\mathcal{G}] \leq \mathsf{P}\text{-esssup}_{\lambda\in\Lambda}\mathsf{E}[X_\lambda \wedge N|\mathcal{G}]$$

for all $N \in \mathbb{N}$. Then *a fortiori* $\mathsf{P}$-a.s. for all $N \in \mathbb{N}$,

$$\mathsf{E}[\mathsf{P}\text{-esssup}_{\lambda\in\Lambda} X_\lambda \wedge N|\mathcal{G}] \leq \mathsf{P}\text{-esssup}_{\lambda\in\Lambda}\mathsf{E}[X_\lambda|\mathcal{G}]$$

and generalised conditional monotone convergence (Lemma A.2) allows to pass to the limit:

$$\mathsf{E}[\lim_{N\to\infty} \mathsf{P}\text{-esssup}_{\lambda\in\Lambda} X_\lambda \wedge N|\mathcal{G}] \leq \mathsf{P}\text{-esssup}_{\lambda\in\Lambda}\mathsf{E}[X_\lambda|\mathcal{G}]$$

$\mathsf{P}$-a.s. But clearly, $\mathsf{P}$-a.s., $\lim_{N\to\infty} \mathsf{P}\text{-esssup}_{\lambda\in\Lambda} X_\lambda \wedge N \geq \mathsf{P}\text{-esssup}_{\lambda\in\Lambda} X_\lambda$, since for all $\lambda \in \Lambda$, we have, $\mathsf{P}$-a.s., $X_\lambda \leq \lim_{N\to\infty} X_\lambda \wedge N \leq \lim_{N\to\infty} \mathsf{P}\text{-esssup}_{\mu\in\Lambda} X_\mu \wedge N$.

Thus it will indeed be sufficient to establish the "$\leq$-inequality" in (A.1) for the truncated families, and so it is assumed without loss of generality (take $M = N$) that $X$ enjoys, for each $\epsilon \in (0,\infty)$, the "$\epsilon$-upwards lattice property": for all $\{\lambda, \lambda'\} \subset \Lambda$, one can find a $\lambda'' \in \Lambda$ with $X_{\lambda''} \geq X_\lambda \vee X_{\lambda'} - \epsilon$ $\mathsf{P}$-a.s.

Then take $(\lambda_n)_{n\geq 1} \subset \Lambda$ such that, $\mathsf{P}$-a.s., $\mathsf{P}\text{-esssup}_{\lambda\in\Lambda} X_\lambda = \sup_{n\geq 1} X_{\lambda_n}$ and fix $\delta > 0$. Recursively define $(\lambda'_n)_{n\geq 1} \subset \Lambda$ so that, $X_{\lambda'_1} = X_{\lambda_1}$ while for $n \in \mathbb{N}$, $\mathsf{P}$-a.s., $X_{\lambda'_{n+1}} \geq X_{\lambda'_n} \vee X_{\lambda_{n+1}} - \delta/2^n$. Prove by induction that $\mathsf{P}$-a.s. for all $n \in \mathbb{N}$, $X_{\lambda'_n} \geq \max_{1\leq k\leq n}(X_{\lambda_k} - \sum_{l=1}^{n-1} \delta/2^l)$, so that $\liminf_{n\to\infty} X_{\lambda'_n} \geq \sup_{n\in\mathbb{N}} X_{\lambda_n} - \delta$, $\mathsf{P}$-a.s. Note next that the negative parts of $(X_{\lambda'_n})_{n\in\mathbb{N}}$ are dominated $\mathsf{P}$-a.s. by a single $\mathsf{P}$-integrable random variable. By the generalised conditional Fatou's lemma (Lemma A.2) we therefore obtain, $\mathsf{P}$-a.s., $\mathsf{P}\text{-esssup}_{\lambda\in\Lambda}\mathsf{E}[X_\lambda|\mathcal{G}] \geq \liminf_{n\to\infty} \mathsf{E}[X_{\lambda'_n}|\mathcal{G}] \geq$

$\mathsf{E}[\liminf_{n\to\infty} X_{\lambda'_n}|\mathcal{G}] \geq \mathsf{E}[\mathsf{P}\text{-esssup}_{\lambda\in\Lambda} X_\lambda|\mathcal{G}] - \delta$. Finally, let $\delta$ descend to 0 (over some sequence descending to 0). □

Department of Mathematics, University of Ljubljana, Slovenia
*E-mail address*: matija.vidmar@fmf.uni-lj.si

Department of Statistics, University of Warwick, UK
*E-mail address*: s.d.jacka@warwick.ac.uk