On dynamic stochastic control with control-dependent information[1][2]

Saul Jacka (Warwick)

Liverpool

2 July, 2015

[1] {see arXiv:1503.02375}

[2] Joint work with Matija Vidmar (Ljubljana)

Outline of talk

Example 1

We have two Brownian Motions, B^0 and B^1. We may only observe one of them at any given time. We acquire a running reward of the difference between the BM we observe and the unobserved one.

Example 1

We have two Brownian Motions, B^0 and B^1. We may only observe one of them at any given time. We acquire a running reward of the difference between the BM we observe and the unobserved one. At any time we may pay a state-dependent cost K to switch our observation to the other BM.

Example 1

We have two Brownian Motions, B^0 and B^1. We may only observe one of them at any given time. We acquire a running reward of the difference between the BM we observe and the unobserved one. At any time we may pay a state-dependent cost K to switch our observation to the other BM. The control process, c, is any cadlag process taking values in $\{0, 1\}$, adapted to the filtration $\mathcal{G}_t^c := \sigma(B_s^{c_s} : s \leq t)$, and representing the index of the BM we choose to observe.

Example 1

We have two Brownian Motions, B^0 and B^1. We may only observe one of them at any given time. We acquire a running reward of the difference between the BM we observe and the unobserved one. At any time we may pay a state-dependent cost K to switch our observation to the other BM. The control process, c, is any cadlag process taking values in $\{0,1\}$, adapted to the filtration $\mathcal{G}^c_t := \sigma(B^{c_s}_s : s \leq t)$, and representing the index of the BM we choose to observe.

We denote by $\sigma^c(t)$ the last time that we changed our observed BM i.e. the last jump time of c before time t, and $\tau^c(t)$ is the lag since the last jump i.e. $\tau^c(t) = t - \sigma^c(t)$. Then we define Z^c as follows:

$$Z^c_t := B^c_t - B^{1-c}_{\sigma^c(t)},$$

so that Z^c is the conditional mean of $B^c - B^{1-c}$ given observations up to time t.

Example 1

We have two Brownian Motions, B^0 and B^1. We may only observe one of them at any given time. We acquire a running reward of the difference between the BM we observe and the unobserved one. At any time we may pay a state-dependent cost K to switch our observation to the other BM. The control process, c, is any cadlag process taking values in $\{0,1\}$, adapted to the filtration $\mathcal{G}^c_t := \sigma(B^{c_s}_s : s \leq t)$, and representing the index of the BM we choose to observe.
We denote by $\sigma^c(t)$ the last time that we changed our observed BM i.e. the last jump time of c before time t, and $\tau^c(t)$ is the lag since the last jump i.e. $\tau^c(t) = t - \sigma^c(t)$. Then we define Z^c as follows:

$$Z^c_t := B^c_t - B^{1-c}_{\sigma^c(t)},$$

so that Z^c is the conditional mean of $B^c - B^{1-c}$ given observations up to time t.
The reward J (which we seek to maximise) is given by

$$J(c) = \int_0^\infty e^{-\alpha t} Z^c_t \, dt - \int_0^\infty e^{-\alpha t} K(Z^c_{t-}, \tau(t-)) |dc_t|$$

Example 2

The controlled process X is merely a process in $\mathbb{R}^n$ (later just $\mathbb{R}$). The control c is the drift of X and is bounded in norm by 1 so

$$dX_t^c = c_t dt + \sigma dW_t,$$

where W is a BM and $||c_t|| \leq 1$.

Example 2

The controlled process X is merely a process in $\mathbb{R}^n$ (later just $\mathbb{R}$). The control c is the drift of X and is bounded in norm by 1 so

$$dX_t^c = c_t dt + \sigma dW_t,$$

where W is a BM and $||c_t|| \leq 1$.
There is an underlying random measure on $\mathbb{R}^n$, μ (which we will later assume to be Poisson with rate λ).
We (progressively) observe the restriction of μ to the closure of the path traced out by the unit ball around X^c.

Example 2

The controlled process X is merely a process in $\mathbb{R}^n$ (later just $\mathbb{R}$). The control c is the drift of X and is bounded in norm by 1 so

$$dX_t^c = c_t dt + \sigma dW_t,$$

where W is a BM and $||c_t|| \leq 1$.
There is an underlying random measure on $\mathbb{R}^n$, μ (which we will later assume to be Poisson with rate λ).
We (progressively) observe the restriction of μ to the closure of the path traced out by the unit ball around X^c.
The objective function J (to be minimised) is

$$J(c) = \int_0^{\tau^c} \mu(B(X_t^c))dt + \kappa 1_{(\tau^c < \infty)},$$

where τ^c is a stopping time (time of retirement) which we also control.

Example 3

N is a Poisson process of unit intensity with arrival times $(S_n)_{n\in\mathbb{N}_0}$, $S_0 := 0$, $S_n < \infty$ for all $n \in \mathbb{N}$. $(R_n)_{n\in\mathbb{N}_0}$ is a sequence of random signs with $\mathrm{P}(R_n = +1) = 1 - \mathrm{P}(R_n = -1) = 2/3$.
The "observed process" is

$$W := W_0 + N + \int_0^{\cdot} \sum_{n\in\mathbb{N}_0} R_n \mathbb{1}_{[S_n, S_{n+1})}(t)dt$$

(so W has 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.

The set of controls, **C**, consists of real-valued, measurable processes, starting at 0, which are adapted to the filtration $\mathcal{F}_t{:=}\sigma(W_{S_n} : \; S_n \leq t)$. 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.
For $X \in \mathbf{C}$ consider the *penalty* functional

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

Let $v := \inf_{X \in \mathbf{C}} \mathrm{E}J(X)$ be the optimal expected penalty; clearly an optimal control is the process $\hat{X}$ which jumps to W_{S_n} at time S_n and assumes a drift of $+1$ in between those instances, so that $v = 1/(3\alpha)$.

Next, for $X \in \mathbf{C}$, let

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

be the "Bellman system".

(1) It is clear 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 $\mathcal{G}$.

(2) Nevertheless the process $(V_{S_n}^X)_{n\in\mathbb{N}_0}$ is a discrete-time submartingale (and martingale when $X = \hat{X}$) with respect to $(\mathcal{G}_{S_n})_{n\in\mathbb{N}_0}$, for all $X \in \mathbf{C}$.

Example 4

Consider coherent risk measures (with a sign change) in multiperiod setting.
So, in the simplest case with no cash flows,

$$\rho(X) := \sup_{\mathbb{Q} \in \mathsf{Q}} \mathsf{E}_{\mathbb{Q}}[X],$$

where Q is some collection of probability measures absolutely continuous with respect to some reference measure $\mathbb{P}$.
We can think of this as a stochastic control problem using the change of measure/Girsanov approach.

Example 4

Consider coherent risk measures (with a sign change) in multiperiod setting.
So, in the simplest case with no cash flows,

$$\rho(X) := \sup_{\mathbb{Q} \in \mathsf{Q}} \mathsf{E}_{\mathbb{Q}}[X],$$

where Q is some collection of probability measures absolutely continuous with respect to some reference measure $\mathbb{P}$.
We can think of this as a stochastic control problem using the change of measure/Girsanov approach.
Now consider the particular case of "expected shortfall" where,

$$\mathsf{Q} = \{\mathbb{Q} : \frac{d\mathbb{Q}}{d\mathbb{P}} \leq \lambda\},$$

for some fixed $\lambda > 1$.

Note: the Basel III accords have fixed on this as the technical measure of risk for reserving and accounting purposes for dervatives and similar contracts.

Note: the Basel III accords have fixed on this as the technical measure of risk for reserving and accounting purposes for dervatives and similar contracts.

Question: is this a stochastic control problem in the dynamic sense? I.e. if we define

$$\rho_t(X) := \mathrm{esssup}_{\mathbb{Q} \in \mathsf{Q}} \mathsf{E}_{\mathbb{Q}}[X|\mathcal{F}_t],$$

is it the case that

$$\rho_t = \rho_t \circ \rho_{t+s}?$$

Note: the Basel III accords have fixed on this as the technical measure of risk for reserving and accounting purposes for dervatives and similar contracts.

Question: is this a stochastic control problem in the dynamic sense? I.e. if we define

$$\rho_t(X) := \operatorname{esssup}_{\mathbb{Q}\in \mathsf{Q}} \mathsf{E}_{\mathbb{Q}}[X|\mathcal{F}_t],$$

is it the case that

$$\rho_t = \rho_t \circ \rho_{t+s}?$$

Answer: no, not in general.

Delbaen gave a nsc for this sort of temporal consistency: m-stability. And in general expected shortfall is not m-stable.

Motivation

In optimal dynamic stochastic control, the output (payoff/penalty):

- is random; the objective being to maximize/minimize its expectation;
- subject to exogenous influence;
- controlled in a way which is adapted to the current and past state of the system.

Information available to controller:

1. May be partial (observable vs. accumulated information).
2. Moreover, may **depend on the control**.

The phenomenon of *control-dependent-information* is common.

Examples: job hunting, quality control, controlled SDEs (loss of information in Tanaka's example), bandit models in economics, etc.

Motivation (cont'd)

Question: Can we offer a consistent general framework for optimal dynamic stochastic control, with an explicit control-dependent informational structure, and that comes equipped with an abstract version of Bellman's optimality (/super/martingale) principle?
Key ingredient is the modeling of information:

- What then is the typical & most important example of a control-dependent filtration?

Motivation (cont'd)

Question: Can we offer a consistent general framework for optimal dynamic stochastic control, with an explicit control-dependent informational structure, and that comes equipped with an abstract version of Bellman's optimality (/super/martingale) principle?
Key ingredient is the modeling of information:

- What then is the typical & most important example of a control-dependent filtration? It is the (completed) natural filtration of an observed (controlled) process.

Motivation (cont'd)

Question: Can we offer a consistent general framework for optimal dynamic stochastic control, with an explicit control-dependent informational structure, and that comes equipped with an abstract version of Bellman's optimality (/super/martingale) principle?
Key ingredient is the modeling of information:

- What then is the typical & most important example of a control-dependent filtration? It is the (completed) natural filtration of an observed (controlled) process.
- *Informational consistency* appears crucial:

  If two controls agree up to a certain time, then what we have observed up to that time should also agree.

Motivation (cont'd)

Question: Can we offer a consistent general framework for optimal dynamic stochastic control, with an explicit control-dependent informational structure, and that comes equipped with an abstract version of Bellman's optimality (/super/martingale) principle?
Key ingredient is the modeling of information:

- What then is the typical & most important example of a control-dependent filtration? It is the (completed) natural filtration of an observed (controlled) process.
- *Informational consistency* appears crucial:

  If two controls agree up to a certain time, then what we have observed up to that time should also agree.

- At the level of random (stopping) times, and in the context of (completed) natural filtrations of processes, this 'obvious' requirement becomes surprisingly non-trivial (at least in continuous time).

Motivation (cont'd)

Question: 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?

Motivation (cont'd)

Question: 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?

Example 5

Let E be an Exp(1) r.v. and I a r.v. uniformly distributed on $\{0, -1, -2\}$. Define the process $X_t := I(t-e)\mathbb{1}_{[0,t]}(e)$, $t \in [0, \infty)$, so X is zero until time E and then has drift I, and the process $Y_t := (-1)(t-e)\mathbb{1}_{[0,t]}(e)\mathbb{1}_{(I<0)}$, $t \in [0, \infty)$, so Y has drift equal to the maximum of -1 and the drift of X.

Motivation (cont'd)

Question: 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?

Example 5

Let E be an Exp(1) r.v. and I a r.v. uniformly distributed on $\{0, -1, -2\}$. Define the process $X_t := I(t-e)\mathbb{1}_{[0,t]}(e)$, $t \in [0, \infty)$, so X is zero until time E and then has drift I, and the process $Y_t := (-1)(t-e)\mathbb{1}_{[0,t]}(e)\mathbb{1}_{(I<0)}$, $t \in [0, \infty)$, so Y has drift equal to the maximum of -1 and the drift of X.

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 $\overline{\mathcal{F}^X}$ as it is of $\overline{\mathcal{F}^Y}$ (but not of $\mathcal{F}^X$ and not of $\mathcal{F}^Y$) ($S = E1_{I\neq 0} + \infty 1_{I=0}$). Moreover, $X^S = 0 = Y^S$.

Motivation (cont'd)

Question: 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?

Example 5

Let E be an Exp(1) r.v. and I a r.v. uniformly distributed on $\{0, -1, -2\}$. Define the process $X_t := I(t-e)\mathbb{1}_{[0,t]}(e)$, $t \in [0,\infty)$, so X is zero until time E and then has drift I, and the process $Y_t := (-1)(t-e)\mathbb{1}_{[0,t]}(e)\mathbb{1}_{(I<0)}$, $t \in [0,\infty)$, so Y has drift equal to the maximum of -1 and the drift of X.

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 $\overline{\mathcal{F}^X}$ as it is of $\overline{\mathcal{F}^Y}$ (but not of $\mathcal{F}^X$ and not of $\mathcal{F}^Y$) ($S = E1_{I\neq 0} + \infty 1_{I=0}$).

Moreover, $X^S = 0 = Y^S$.

Consider now the event $A := \{I = -1\}$. Then it is clear that $A \in \overline{\mathcal{F}^X}_S$ but $A \notin \overline{\mathcal{F}^Y}_S$.

Overview

- On the 'optimal dynamic stochastic control' front:
    1. Of course, the phenomenon of control-dependent information has been studied in the literature in specific situations/problems; but focus there 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*, at least in some sense.
    2. As far as general frameworks go, however, hitherto, only a single, non-control dependent (observable) informational flow appears to have been allowed.
- On the 'informational consistency' front:
    1. What is essentially required is a test "connecting $\sigma(X^S)$ with $\mathcal{F}^X_S$".
    2. In literature this is available for coordinate processes on canonical spaces.
    3. However, coordinate processes are quite restrictive, and certainly not very relevant to stochastic control.

Informal statement of results

- Basically: we answer in the affirmative the two questions posed above.
- Specifically:
  1. We put forward a general stochastic control framework which explicitly allows for a control-dependent informational flow. In particular, we provide a fully general (modulo the relevant (technical) condition) abstract version of Bellman's principle in such a setting.
  2. With respect to the second question, a generalization of (a part of) Galmarino's test to a non-canonical space setting is proved, although full generality could not be achieved.

See [[1], [2], [3], [4], [5], [6], [7], [8]] for some examples!

Convention

The modeling of the informational flow using filtrations, can be done in one of the following two, essentially different, ways: With or without completion. We'll do both with any differences denoted by $\{\}$ braces.

An abstract stochastic control system

We will see a system of stochastic control as consisting of:

$$\left(T, \mathbf{C}, \Omega, (\mathcal{F}^c)_{c\in\mathbf{C}}, (\mathsf{P}^c)_{c\in\mathbf{C}}, J, (\mathcal{G}^c)_{c\in\mathbf{C}}\right),$$

where

(i) A *time set* T with a linear ordering $\leq$. We assume $T = \mathbb{N}_0$ or $\mathbb{R}^+$.

(i) A *time set* T with a linear ordering $\leq$. We assume $T = \mathbb{N}_0$ or $\mathbb{R}^+$.

(ii) A set $\mathbf{C}$ of *admissible controls*.

(i) A *time set* T with a linear ordering $\leq$. We assume $T = \mathbb{N}_0$ or $\mathbb{R}^+$.

(ii) A set $\mathbf{C}$ of *admissible controls*.

(iii) A non-empty sample space Ω endowed with a collection of σ-algebras $(\mathcal{F}^c)_{c\in\mathbf{C}}$. Here $\mathcal{F}^c$ is all the *information accumulated* by the "end of time"/ a "terminal time", when c is the chosen control.

(i) A *time set* T with a linear ordering $\leq$. We assume $T = \mathbb{N}_0$ or $\mathbb{R}^+$.

(ii) A set $\mathbf{C}$ of *admissible controls*.

(iii) A non-empty sample space Ω endowed with a collection of σ-algebras $(\mathcal{F}^c)_{c \in \mathbf{C}}$. Here $\mathcal{F}^c$ is all the *information accumulated* by the "end of time"/ a "terminal time", when c is the chosen control.

(iv) $(\mathsf{P}^c)_{c \in \mathbf{C}}$, a collection of $\{$complete$\}$ probability measures, each P^c having domain which includes the $\{\mathsf{P}^c$-complete$\}$ σ-field $\mathcal{F}^c$ (for $c \in \mathbf{C}$).

(i) A *time set* T with a linear ordering $\leq$. We assume $T = \mathbb{N}_0$ or $\mathbb{R}^+$.

(ii) A set $\mathbf{C}$ of *admissible controls*.

(iii) A non-empty sample space Ω endowed with a collection of σ-algebras $(\mathcal{F}^c)_{c \in \mathbf{C}}$. Here $\mathcal{F}^c$ is all the *information accumulated* by the "end of time"/ a "terminal time", when c is the chosen control.

(iv) $(\mathsf{P}^c)_{c \in \mathbf{C}}$, a collection of {complete} probability measures, each P^c having domain which includes the {P^c-complete} σ-field $\mathcal{F}^c$ (for $c \in \mathbf{C}$).

(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 P^c-a.s. equality}. We further insist $\mathsf{E}^{\mathsf{P}^c} J(c)^- < \infty$ for all $c \in \mathbf{C}$ [we are maximising]. Given the control $c \in \mathbf{C}$, $J(c)$ is the random *payoff*.

(i) A *time set* T with a linear ordering $\leq$. We assume $T = \mathbb{N}_0$ or $\mathbb{R}^+$.

(ii) A set $\mathbf{C}$ of *admissible controls*.

(iii) A non-empty sample space Ω endowed with a collection of σ-algebras $(\mathcal{F}^c)_{c\in\mathbf{C}}$. Here $\mathcal{F}^c$ is all the *information accumulated* by the "end of time"/ a "terminal time", when c is the chosen control.

(iv) $(\mathsf{P}^c)_{c\in\mathbf{C}}$, a collection of $\{$complete$\}$ probability measures, each P^c having domain which includes the $\{\mathsf{P}^c$-complete$\}$ σ-field $\mathcal{F}^c$ (for $c \in \mathbf{C}$).

(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 P^c-a.s. equality$\}$. We further insist $\mathsf{E}^{\mathsf{P}^c} J(c)^- < \infty$ for all $c \in \mathbf{C}$ [we are maximising]. Given the control $c \in \mathbf{C}$, $J(c)$ is the random *payoff*.

(vi) A collection of filtrations $(\mathcal{G}^c)_{c\in\mathbf{C}}$ on Ω. We assume that $\mathcal{G}^c_\infty := \vee_{t\in T}\mathcal{G}^c_t \subset \mathcal{F}^c$, and (for simplicity) that each $\mathcal{G}^c_0$ is P^c-trivial $\{$and P^c-completed$\}$, while $\mathcal{G}^c_0 = \mathcal{G}^d_0$ $\{$i.e. null sets are constant over $\mathbf{C}\}$ and $\mathsf{P}^c|_{\mathcal{G}^c_0} = \mathsf{P}^d|_{\mathcal{G}^d_0}$ for all $c, d \in \mathbf{C}$. $\mathcal{G}^c_t$ is the *information acquired* **by the controller** by time $t \in$, if the control is c.

The dynamical structure

Definition (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 P^c-a.s. equality} stopping time of $\mathcal{G}^c$ for every $c \in \mathbf{C}$.

We assume given: *a* collection $\mathbf{G}$ of controlled times, and also a family $(\mathcal{D}(c,\mathcal{S}))_{(c,\mathcal{S})\in\mathbf{C}\times\mathbf{G}}$ of subsets of $\mathbf{C}$ for which 'certain natural' conditions hold true – stemming from the interpretation that $\mathcal{D}(c,\mathcal{S})$ are the controls agreeing with c up to time $\mathcal{S}$.

We write $c \sim_{\mathcal{S}} d$ **for** $d \in \mathcal{D}(c,\mathcal{S})$.

We **assume** that $(\mathcal{D}(c,\mathcal{S}))_{(c,\mathcal{S})\in\mathbf{C}\times\mathbf{G}}$ is a collection of subsets of $\mathbf{C}$ for which:

(1) $c\in\mathcal{D}(c,\mathcal{S})$ for all $(c,\mathcal{S})\in\mathbf{C}\times\mathbf{G}$.

We **assume** that $(\mathcal{D}(c,\mathcal{S}))_{(c,\mathcal{S})\in \mathbf{C}\times\mathbf{G}}$ is a collection 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 $c,d\in\mathbf{C}$, $d\in\mathcal{D}(c,\mathcal{S})$ implies $\mathcal{S}^c=\mathcal{S}^d$ $\{\mathsf{P}^c$ & P^d-a.s$\}$.

We **assume** that $(\mathcal{D}(c,\mathcal{S}))_{(c,\mathcal{S})\in\mathbf{C}\times\mathbf{G}}$ is a collection 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 $c, d \in \mathbf{C}$, $d \in \mathcal{D}(c,\mathcal{S})$ implies $\mathcal{S}^c = \mathcal{S}^d$ $\{\mathsf{P}^c$ & P^d-a.s$\}$.

(3) If $\mathcal{S}, \mathcal{T} \in \mathbf{G}$, $c \in \mathbf{C}$ and $\mathcal{S}^c = \mathcal{T}^c$ $\{\mathsf{P}^c$-a.s$\}$, then $\mathcal{D}(c,\mathcal{S}) = \mathcal{D}(c,\mathcal{T})$.

We **assume** that $(\mathcal{D}(c,\mathcal{S}))_{(c,\mathcal{S})\in \mathbf{C}\times\mathbf{G}}$ is a collection 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 $c, d \in \mathbf{C}$, $d \in \mathcal{D}(c,\mathcal{S})$ implies $\mathcal{S}^c = \mathcal{S}^d$ $\{\mathsf{P}^c$ & P^d-a.s$\}$.

(3) If $\mathcal{S},\mathcal{T} \in \mathbf{G}$, $c \in \mathbf{C}$ 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 $\mathcal{S},\mathcal{T} \in \mathbf{G}$ and $c \in \mathbf{C}$ is such that $\mathcal{S}^d \le \mathcal{T}^d$ $\{\mathsf{P}^d$-a.s.$\}$ for each $d \in \mathcal{D}(c,\mathcal{T})$, then $\mathcal{D}(c,\mathcal{T}) \subset \mathcal{D}(c,\mathcal{S})$.

We **assume** that $(\mathcal{D}(c,\mathcal{S}))_{(c,\mathcal{S})\in\mathbf{C}\times\mathbf{G}}$ is a collection 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 $c,d\in\mathbf{C}$, $d\in\mathcal{D}(c,\mathcal{S})$ implies $\mathcal{S}^c=\mathcal{S}^d$ $\{\mathsf{P}^c$ & P^d-a.s$\}$.

(3) If $\mathcal{S},\mathcal{T}\in\mathbf{G}$, $c\in\mathbf{C}$ 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 $\mathcal{S},\mathcal{T}\in\mathbf{G}$ and $c\in\mathbf{C}$ is such that $\mathcal{S}^d\leq\mathcal{T}^d$ $\{\mathsf{P}^d$-a.s.$\}$ for each $d\in\mathcal{D}(c,\mathcal{T})$, 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 $\mathbf{C}$.

We **assume** that $(\mathcal{D}(c,\mathcal{S}))_{(c,\mathcal{S})\in\mathbf{C}\times\mathbf{G}}$ is a collection 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 $c, d \in \mathbf{C}$, $d \in \mathcal{D}(c,\mathcal{S})$ implies $\mathcal{S}^c = \mathcal{S}^d$ $\{\mathsf{P}^c$ & P^d-a.s$\}$.

(3) If $\mathcal{S},\mathcal{T} \in \mathbf{G}$, $c \in \mathbf{C}$ 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 $\mathcal{S},\mathcal{T} \in \mathbf{G}$ and $c \in \mathbf{C}$ is such that $\mathcal{S}^d \leq \mathcal{T}^d$ $\{\mathsf{P}^d$-a.s.$\}$ for each $d \in \mathcal{D}(c,\mathcal{T})$, 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 P^c-a.s.$\}$ equal to ∞ (resp. 0).

We **assume** that $(\mathcal{D}(c,\mathcal{S}))_{(c,\mathcal{S})\in\mathbf{C}\times\mathbf{G}}$ is a collection 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 $c,d\in\mathbf{C}$, $d\in\mathcal{D}(c,\mathcal{S})$ implies $\mathcal{S}^c=\mathcal{S}^d$ $\{\mathsf{P}^c$ & P^d-a.s$\}$.

(3) If $\mathcal{S},\mathcal{T}\in\mathbf{G}$, $c\in\mathbf{C}$ 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 $\mathcal{S},\mathcal{T}\in\mathbf{G}$ and $c\in\mathbf{C}$ is such that $\mathcal{S}^d\leq\mathcal{T}^d$ $\{\mathsf{P}^d$-a.s.$\}$ for each $d\in\mathcal{D}(c,\mathcal{T})$, 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 P^c-a.s.$\}$ equal to ∞ (resp. 0).

For example:

> *For all* $\mathcal{S}\in\mathbf{G}$ *and* $c,d\in\mathbf{C}$*:* $d\in\mathcal{D}(c,\mathcal{S})$ *implies* $\mathcal{S}^c=\mathcal{S}^d$ $\{\mathsf{P}^c$ *&* P^d*-a.s*$\}$*.*

Temporal consistency and optimality

Assumption (Temporal consistency)

For all $c, d \in \mathbf{C}$ *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* $\mathsf{P}^c|_{\mathcal{G}^c_{\mathcal{S}^c}} = \mathsf{P}^d|_{\mathcal{G}^d_{\mathcal{S}^d}}$.

Temporal consistency and optimality

Assumption (Temporal consistency)

For all $c, d \in \mathbf{C}$ *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* $\mathsf{P}^c|_{\mathcal{G}^c_{\mathcal{S}^c}} = \mathsf{P}^d|_{\mathcal{G}^d_{\mathcal{S}^d}}$.

Now we can define the optimal payoff.

Definition (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**. $c \in \mathbf{C}$ is said to be **optimal** if $\mathsf{E}^{\mathsf{P}^c} J(c) = v$.

The conditional payoff and the Bellman system

Definition (Conditional payoff & Bellman system)

We define for $c \in \mathbf{C}$ and $\mathcal{S} \in \mathbf{G}$:

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

and then

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

and say $c \in \mathbf{C}$ is **conditionally optimal** at $\mathcal{S} \in \mathbf{G}$, if $V(c, \mathcal{S}) = J(c, \mathcal{S})$ 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**.

A (vital) lattice condition

Assumption (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* (ϵ, 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* P^c*-a.s.*
>
> $$J(d'', \mathcal{S}) \geq (M \wedge J(d, \mathcal{S})) \vee (M \wedge J(d', \mathcal{S})) - \epsilon.$$

This condition represents a direct linking between **C**, **G** and the collection $(\mathcal{G}^c)_{c \in \mathbf{C}}$. In particular, it may fail at deterministic times, as in Example 3.

Theorem (Bellman's principle)

$(V(c,\mathcal{S}))_{(c,\mathcal{S})\in\mathbf{C}\times\mathbf{G}}$ is a $(\mathbf{C},\mathbf{G})$-supermartingale system: that is, whenever $\mathcal{S},\mathcal{T}\in\mathbf{G}$ and $c\in\mathbf{C}$ satisfy $\mathcal{S}^d\leq\mathcal{T}^d$ $\{\mathsf{P}^d\text{-a.s.}\}$ for every $d\in\mathcal{D}(c,\mathcal{T})$, P^{c^*}-a.s.*:*

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

Theorem (Bellman's principle)

$(V(c,\mathcal{S}))_{(c,\mathcal{S})\in\mathbf{C}\times\mathbf{G}}$ is a $(\mathbf{C},\mathbf{G})$-supermartingale system: that is, whenever $\mathcal{S},\mathcal{T}\in\mathbf{G}$ and $c\in\mathbf{C}$ satisfy $\mathcal{S}^d\leq\mathcal{T}^d$ $\{\mathsf{P}^d$-a.s.$\}$ for every $d\in\mathcal{D}(c,\mathcal{T})$, P^{c^*}-a.s.:

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

Moreover, if $c^*\in\mathbf{C}$ is optimal, then $(V(c^*,\mathcal{T}))_{\mathcal{T}\in\mathbf{G}}$ is a $\mathbf{G}$-martingale: i.e. for any $\mathcal{S},\mathcal{T}\in\mathbf{G}$ with $\mathcal{S}^d\leq\mathcal{T}^d$ $\{\mathsf{P}^d$-a.s.$\}$ for each $d\in\mathcal{D}(c^*,\mathcal{T})$, P^{c^*}-a.s.,

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

Theorem (Bellman's principle)

$(V(c,\mathcal{S}))_{(c,\mathcal{S})\in\mathbf{C}\times\mathbf{G}}$ is a $(\mathbf{C},\mathbf{G})$-supermartingale system: that is, whenever $\mathcal{S},\mathcal{T}\in\mathbf{G}$ and $c\in\mathbf{C}$ satisfy $\mathcal{S}^d\leq\mathcal{T}^d$ $\{\mathrm{P}^d\text{-a.s.}\}$ for every $d\in\mathcal{D}(c,\mathcal{T})$, P^{c^*}-a.s.:

$$\mathrm{E}^{\mathrm{P}^c}[V(c,\mathcal{T})|\mathcal{G}^c_{\mathcal{S}^c}]\leq V(c,\mathcal{S}).$$

Moreover, if $c^*\in\mathbf{C}$ is optimal, then $(V(c^*,\mathcal{T}))_{\mathcal{T}\in\mathbf{G}}$ is a $\mathbf{G}$-martingale: i.e. for any $\mathcal{S},\mathcal{T}\in\mathbf{G}$ with $\mathcal{S}^d\leq\mathcal{T}^d$ $\{\mathrm{P}^d\text{-a.s.}\}$ for each $d\in\mathcal{D}(c^*,\mathcal{T})$, P^{c^*}-a.s.,

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

Conversely, and regardless of whether the "upwards lattice assumption" 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 P^{c^*}-expectation and is uniformly integrable, and (iii) $V(c^*,\mathcal{S}_n)\to V(c^*,\infty)$, P^{c^*}-a.s. (or even just in P^{c^*}-probability), as $n\to\infty$, then c^* is optimal.

Theorem (Minimal supermartingale)

Assume the upwards lattice property and that $\infty \in \mathbf{G}$. *Then* V *is the minimal* $(\mathbf{C}, \mathbf{G})$*-supermartingale* W *satisfying the terminal condition* $W(c, \infty) \geq \mathsf{E}^{\mathsf{P}^c}[J(c)|\mathcal{G}^c_\infty]$ *for each* $c \in \mathbf{C}$.

Theorem (Minimal supermartingale)

Assume the upwards lattice property and that $\infty \in \mathbf{G}$*. Then* V *is the minimal* $(\mathbf{C}, \mathbf{G})$*-supermartingale* W *satisfying the terminal condition* $W(c, \infty) \geq \mathsf{E}^{\mathsf{P}^c}[J(c) | \mathcal{G}^c_\infty]$ *for each* $c \in \mathbf{C}$.

Theorem (Observational consistency)

Let X *and* Y *be two processes (on* Ω*, with time domain* T *and values in* $(E, \mathcal{E})$*),* 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* σ*-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 $\mathcal{F}^X_S = \mathcal{F}^Y_S$.

Theorem (Observational consistency II)

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

(1) $(\Omega,\mathcal{G})$ *(resp.* $(\Omega,\mathcal{H})$*) is Blackwell for some* σ*-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 $\overline{\mathcal{F}^Z}^{\mathsf{P}^Z}_T = \overline{\sigma(Z^T)}^{\mathsf{P}^Z} = \overline{\sigma(W^T)}^{\mathsf{P}^W} = \overline{\mathcal{F}^W}^{\mathsf{P}^W}_T$.

Example 1

Recall that c is the control process taking values in $\{0, 1\}$, and representing the index of the BM we choose to observe. The two BMs are B^0 and B^1 and we denote by $\sigma^c(t)$ the last time that we changed our observed BM i.e. the last jump time of c before time t, and $\tau^c(t)$ is the lag since the last jump i.e. $\tau^c(t) = t - \sigma^c(t)$. Then we define Z^c as follows:

$$Z^c_t := B^c_t - B^{1-c}_{\sigma^c(t)},$$

so that Z^c is the conditional mean of $B^c - B^{1-c}$ given $\mathcal{G}^c_t$.
The reward J is given by

$$J(c) = \int_0^\infty e^{-\alpha t} Z^c_t dt - \int_0^\infty e^{-\alpha t} K(Z^c_{t-}, \tau(t-))|dc_t|$$

We will solve this in the following special case:

$$K(z,t)=\infty : z>-l$$

where

$$\begin{aligned} K(z,t) &= \int_{\mathbb{R}} dw\phi(w)[(\psi(z-\sqrt{t}w)-k(\sqrt{t}w-z))1_{(z-\sqrt{t}w>l)} \\ &+\psi(\sqrt{t}w-z)1_{(z-\sqrt{t}w\le l)}]-\psi(-z)+k(z) : z\le -l, \end{aligned}$$

$\psi : z \mapsto \frac{z}{\alpha}+Ae^{-\gamma z}$ and

$$k : z\mapsto \infty : z>-l$$

$$k : z\mapsto [\psi(l)-\psi(-l)]e^{\gamma(z+l)} : z\le -l,$$

with $\gamma=\sqrt{2\alpha}$, $A=\frac{3}{2\gamma\alpha e^{\frac{1}{2}}}$ and $l=\frac{1}{2\gamma}$.
Clearly any c which jumps when Z^c is above level $-l$ attracts infinite cost. For controls which do not do this, define

$$S_t^c = \int_0^t e^{-\alpha s} Z_s^c ds - \int_0^t e^{-\alpha s} K(Z_{s-}^c, \tau(s-))|dc_s| + e^{-\alpha t} V(Z_t^c, \tau^c(t)),$$

for a function V to be defined shortly. Bellman's principle tells us that this should be a $\mathcal{G}_t^c$ supermartingale and a martingale if c is optimal provided that V gives the optimal payoff.

We define V by

$$V(z) = \psi(z) : z \geq -l$$

$$V(z) = \psi(-z) - k(z) : z \leq -l,$$

and observe that V is C^1, and C^2 except at $-l$. We claim that V is the optimal payoff and an "optimal policy" is to switch (i.e. jump c) every time that Z^c hits $-l$. Denote this policy by $\hat{c}$.

Now to show that Bellman's principle holds for this V we need to show:

1. $\frac{1}{2}V''(z) - \alpha V(z) + z \leq 0$;
2. $\frac{1}{2}V''(z) - \alpha V(z) + z = 0$ for $z > -l$;
3. $\int_{\mathbb{R}} dw\phi(w)(V(\sqrt{t}w - z) - K(z,t) - V(z) \leq 0$;
4. $\int_{\mathbb{R}} dw\phi(w)(V(\sqrt{t}w - z) - K(z,t) - V(z) = 0$ for $z \leq -l$;
5. $S_t^{\hat{c}}$ is a ui martingale and converges to $J(\hat{c})$.

Example 2

Recall that (setting $n = 1$ and $\sigma = 0$) the controlled process X is merely a process in $\mathbb{R}$. The control c is the drift of X and is bounded in norm by 1 so

$$dX_t^c = c_t dt.$$

There is an underlying random poisoon measure μ with rate λ.
The filtration $\mathcal{G}^c$ is given by $\mathcal{G}_t^c = \sigma(\mu|_{\overline{\cup_{s \leq t} B(X_s^c)}})$, where $B(x)$ denotes the unit ball centred on x. This means we observe the restriction of μ to the closure of the path traced out by the unit ball around X^c. The objective function J (to be minimised) is

$$J(c) = \int_0^{\tau^c} \mu(B(X_t^c))dt + \kappa 1_{(\tau^c < \infty)},$$

where τ^c is a stopping time (time of retirement) which we also control.
This doesn't *a priori* meet the criterion of having $\mathrm{E}J(d)^+ < \infty$ for all d, so we simply restrict to those c with this property.

Define, for $x, y \in \mathbb{R}$, and ν a measure on $(\mathbb{R}, \mathcal{B})$:

$$f(x, y, \nu) = \int_{x \wedge y}^{x \vee y} \nu(B(z))dz$$

and denote the point mass at x by δ_x. Denote the running iminimum of X^c by I^c and the running maximum by S^c.
For any measure ν on $(\mathbb{R}, \mathcal{B})$ and $i \leq s$, we denote the restriction of ν to the interval $(i-1, s+1)$ by $\nu_{i,s}$; i.e

$$\frac{d\nu_{i,s}}{d\nu} = 1_{(i-1,s+1)}.$$

Now, for any i, s, ν and $x \in [i, s]$, define

$$z_l(x, i, s, \nu) = \sup\{y \leq x : \quad \nu_{i,s}(B(y)) = 0\}$$

and

$$z_r(x, i, s, \nu) = \inf\{y \geq x : \quad \nu_{i,s}(B(y)) = 0\}$$

We actually want to solve the problem with **G** consisting of all fixed times, but first we'll "solve" the problem when **G** consists of a much more restricted set of times: **G** initially consists of controlled times S such that S^c is a $\mathcal{G}^c$-stopping time with the additional property that one of the following four conditions hold:

1. $S^c = 0$
2. or $X^c_{S^c} = I^c_{S^c}$ and $\mu(\{X^c_{S^c}\}) = 1$
3. or $X^c_{S^c} = S^c_{S^c}$ and $\mu(\{X^c_{S^c}\}) = 1$
4. or $\mu(B(X^c_{S^c})) = 0$.

We also restrict controls to taking values in $\{-1,0,1\}$), so X^c may go left or right at unit rate or pause where it is. We refer to this as Problem 2'. It is clear from this setup that at each control time we may choose to pause or retire at time 0, and thereafter we must go left until we hit a zero of $\mu(B(X))$ or a newly revealed point mass of μ, or go right until we hit a zero of $\mu(B(X))$ or a newly revealed point mass of μ.
There is a corresponding "optimality equation":

$$v(x,i,s,\nu)=\min(\kappa,\chi_{\{0\}}(\nu(B(x)),v_l(x,i,s,\nu),v_r(x,i,s,\nu)),$$

where

$$v_l(x,i,s,\nu)=\begin{cases}f(x,z_l,\nu_{i,s})\\ e^{-\lambda(i-z_l)}f(x,z_l,\nu_{i,s})\\ \quad+\int_0^{i-z_l}\lambda e^{-\lambda t}(f(x,i-t,\nu_{i,s})+v(i-t,i-t,s,\nu_{i,s}+\delta_{i-t})dt\end{cases} \tag{1}$$

v_r is defined in a corresponding fashion and χ is the (convex-analytic) indicator function:

$$\chi_A(x) = \begin{cases} 0 & \text{if } x \in A \\ \infty & \text{if } x \notin A \end{cases}$$

Claim: there is a minimal positive solution v to the equation (1.1).
Claim: the optimal payoff to Problem 2 is

$$V(c,t) = \int_0^{t \wedge \tau^c} \mu(B(X_u^c))du + \kappa 1_{(\tau^c \leq t)} + v(X_t^c, I_t^c, S_t^c, \mu) 1_{(\tau^c > t)}.$$

A. Bensoussan.
Stochastic Control of Partially Observable Systems.
Cambridge University Press, 2004.

V. I. Bogachev.
Measure Theory, volume 2 of *Measure Theory*.
Springer, 2007.

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.

W. Fleming and E. Pardoux.
Optimal control for partially observed diffusions.
SIAM Journal on Control and Optimization, 20(2):261–285, 1982.

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.

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.

C. Striebel.
Martingale conditions for the optimal control of continuous time stochastic systems.
Stochastic Processes and their Applications, 18(2):329 – 347, 1984.

W. Wonham.
On the separation theorem of stochastic control.
SIAM Journal on Control, 6(2):312–326, 1968.