

## A note on the strong formulation of stochastic control problems with model uncertainty\*

Mihai Sîrbu <sup>†</sup>

### Abstract

We consider a Markovian stochastic control problem with model uncertainty. The controller (intelligent player) observes only the state, and, therefore, uses feedback (closed-loop) strategies. The adverse player (nature) who does not have a direct interest in the payoff, chooses open-loop controls that parametrize Knightian uncertainty. This creates a two-step optimization problem (like half of a game) over feedback strategies and open-loop controls. The main result is to show that, under some assumptions, this provides the same value as the (half of) the zero-sum symmetric game where the adverse player also plays feedback strategies and actively tries to minimize the payoff. The value function is independent of the filtration accessible to the adverse player. Aside from the modeling issue, the present note is a technical companion to a previous work.

**Keywords:** model uncertainty; stochastic games; elementary feedback strategies; stochastic Perron's method; viscosity solutions.

**AMS MSC 2010:** 91A05; 91A15; 49L20; 49L25.

Submitted to ECP on April 7, 2014, final version accepted on November 3, 2014.

## 1 Introduction

We consider a stochastic control problem with model uncertainty. At first, the problem looks identical to the symmetric zero-sum game in [S14b]. However, here, only one player is a true optimizer (intelligent player) who tries to maximize the payoff. The other control variable is chosen by an adverse player (nature) who does not have a vested interest in minimizing the payoff, and models Knightian uncertainty. We argue that the two apparently identical problems (the symmetric zero-sum game in [S14b] and the model uncertainty) should be rigorously defined differently.

More precisely, we interpret the control problem with model uncertainty as a two-step optimization problem. The controller (intelligent players) observes the state process only, so he/she chooses feedback (closed-loop) strategies. The adverse player chooses open-loop controls, and such controls are actually adapted to a possibly larger filtration than the one generated by the Brownian motion. In other words, the adverse player, while not acting strategically against the controller, has access to the Brownian motion and other information and may choose a parametrization of the model which **just happens** to be totally adverse to the controller.

---

\*Supported in part by the National Science Foundation under Grant DMS 1211988. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

<sup>†</sup>University of Texas at Austin, Department of Mathematics, USA. E-mail: sirbu@math.utexas.edu

A similar model of robust control over feedback/closed-loop/positional strategies for the controller and open loop-controls for the adverse player has been considered in [KS88] in deterministic setting. However, our discretization of time for the feedback strategies is different, and, arguably, better fitted to the present case where the system is stochastic, and allows for strong solutions of the state system. In addition, our note deals with the (important, in our view) issue of the information available to the adverse player. A part of our contribution is to prove that the value function does not depend on the filtration accessible to the adverse player. This is not obvious a-priori.

There is a vast literature on robust optimization/model uncertainty, and we do not even attempt to scratch the surface in presenting the history of the problem. However, we have not encountered this very particular way to represent stochastic optimization problems with model uncertainty, i.e. a **strong** formulation over **elementary feedback strategies** for the controller vs. **open-loop controls for the nature**, nor the technical result about the equality of the value functions we obtain.

The message of the present note is two-fold: first, an optimization problem with model uncertainty is **not** the same as a zero-sum game, so it should be modeled differently. We propose to use feedback strategies for the controller and open-loop controls for the adverse player, obtaining a two-step/sup-inf optimization problem over **strong solutions** of the state system. Second, with this formulation, the value function is, indeed, equal to the (lower) value of the zero-sum game, where the adverse player is symmetric to the controller and also plays pure feedback strategies. Beyond the modeling issue, the mathematical statement does not seem obvious, and the proof is based on verification by Stochastic Perron's Method, along the lines of [S14b]. It is unclear how one could prove directly, using only the probabilistic representation of the value functions, such statement.

## 2 Stochastic Control with Model Uncertainty

### 2.1 The Stochastic System

We consider a stochastic differential system of the form:

$$\begin{cases} dX_t = b(t, X_t, u_t, v_t)dt + \sigma(t, X_t, u_t, v_t)dW_t, \\ X_s = x \in \mathbb{R}^d, \end{cases} \quad (2.1)$$

starting at an initial time  $0 \leq s \leq T$  at some position  $x \in \mathbb{R}^d$ . Here, the control  $u$  chosen by the controller (intelligent player) belongs to some compact metric space  $(U, d_U)$  and the parameter  $v$  (chosen by the adverse player/nature) belongs to some other compact metric space  $(V, d_V)$  and represents the model uncertainty. In other words, the Brownian motion  $W$  represents the “known unknowns”, and the process  $v$  stands for the “unknown unknowns”, a.k.a. “Knightian uncertainty”. The state  $X$  lives in  $\mathbb{R}^d$  and the process  $(W_t)_{s \leq t \leq T}$  is a  $d'$ -dimensional Brownian motion on a *fixed* probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  with respect to some filtration  $\mathbb{F} = (\mathcal{F}_t)_{s \leq t \leq T}$ . The filtration  $\mathbb{F}$  satisfies the usual conditions and is usually larger than the augmented natural filtration generated by the Brownian motion, by which we mean,  $\mathcal{F}_t^W = \sigma(W_u, s \leq u \leq t) \vee \mathcal{N}(\mathbb{P}, \mathcal{F})$  for  $s \leq t \leq T$ . The space  $(\Omega, \mathcal{F}, \mathbb{P})$ , the Brownian motion  $W$  and the filtration  $\mathbb{F}$  may depend on  $s$ . To keep the notation simple, we do not emphasize the dependence on  $s$ , unless needed. The coefficients  $b : [0, T] \times \mathbb{R}^d \times U \times V \rightarrow \mathbb{R}^d$  and  $\sigma : [0, T] \times \mathbb{R}^d \times U \times V \rightarrow \mathcal{M}^{d \times d'}$  satisfy the **Standing assumption**:

1. **(C)**  $b, \sigma$  are jointly continuous on  $[0, T] \times \mathbb{R}^d \times U \times V$
2. **(L)**  $b, \sigma$  satisfy a uniform local Lipschitz condition in  $x$ , i.e.

$$|b(t, x, u, v) - b(t, y, u, v)| + |\sigma(t, x, u, v) - \sigma(t, y, u, v)| \leq L(K)|x - y|$$

$\forall |x|, |y| \leq K, t \in [0, T], u \in U, v \in V$  for some  $L(K) < \infty$ , and

3. **(GL)**  $b, \sigma$  satisfy a global linear growth condition in  $x$

$$|b(t, x, u, v)| + |\sigma(t, x, u, v)| \leq C(1 + |x|)$$

$\forall |x|, |y| \in \mathbb{R}^d, t \in [0, T], u \in U, v \in V$  for some  $C < \infty$ .

Now, given a bounded and continuous function  $g : \mathbb{R}^d \rightarrow \mathbb{R}$ , the controller is trying to maximize  $\mathbb{E}[g(X_T^{s,x;u,v})]$ . Since  $v$  is "uncertain", optimizing "robustly", means optimizing the functional  $\inf_v \mathbb{E}[g(X_T^{s,x;u,v})]$ , leading to the two-step optimization problem

$$\sup_u \left( \inf_v \mathbb{E}[g(X_T^{s,x;u,v})] \right).$$

It is not yet clear what  $u, v$  mean in the formulation above, and giving a precise meaning to this is one of the goals of the present note.

## 2.2 Modeling a Zero-Sum Game

For an identical stochastic system, imagine that  $v$  represents the choice of another intelligent player and  $g(X_T^{s,x;u,v})$  is the amount paid by the  $v$  player to the  $u$  player. For this closely related, but **different** problem it was argued in [S14b] that, as long as both players only observe the state process, they should both play, symmetrically, as strategies, some feedback functionals  $u, v$  of restricted form.

We denote by  $C([s, T]) \triangleq C([s, T], \mathbb{R}^d)$  and endow this path space with the natural (and raw) filtration  $\mathbb{B}^s = (\mathcal{B}_t^s)_{s \leq t \leq T}$  defined by  $\mathcal{B}_t^s \triangleq \sigma(y(u), s \leq u \leq t), s \leq t \leq T$ . The elements of the path space  $C([s, T])$  will be denoted by  $y(\cdot)$  or  $y$ . The stopping times on the space  $C([s, T])$  with respect with the filtration  $\mathbb{B}^s$ , i.e. mappings  $\tau : C([s, T]) \rightarrow [s, T]$  satisfying  $\{\tau \leq t\} \in \mathcal{B}_t^s \forall s \leq t \leq T$  are called stopping rules, following [KS01]. We denote by  $\mathbb{B}^s$  the class of such stopping rules starting at  $s$ .

**Definition 2.1** (Elementary Feedback Strategies). Fix  $0 \leq s \leq T$ . An elementary strategy  $\alpha$  starting at  $s$ , for the first intelligent player/controller is defined by

- a finite non-decreasing sequence of stopping rules, i.e.  $\tau_k \in \mathbb{B}^s$  for  $k = 1, \dots, n$  and

$$s = \tau_0 \leq \dots \tau_k \leq \dots \leq \tau_n = T,$$

- for each  $k = 1 \dots n$ , a constant value of the strategy  $\xi_k$  in between the times  $\tau_{k-1}$  and  $\tau_k$ , which is decided based only on the knowledge of the past state up to  $\tau_{k-1}$ , i.e.  $\xi_k : C([s, T]) \rightarrow U$  such that  $\xi_k \in \mathcal{B}_{\tau_{k-1}}^s$ .

The strategy is to hold  $\xi_k$  in between  $(\tau_{k-1}, \tau_k]$ , i.e.  $\alpha : (s, T] \times C([s, T]) \rightarrow U$  is defined by

$$\alpha(t, y(\cdot)) \triangleq \sum_{k=1}^n \xi_k(y(\cdot)) \mathbf{1}_{\{\tau_{k-1}(y(\cdot)) < t \leq \tau_k(y(\cdot))\}}.$$

An elementary strategy  $\beta$  for the second player is defined in an identical way, but takes values in  $V$ . We denote by  $\mathcal{A}(s)$  and  $\mathcal{B}(s)$  the collections of all possible elementary strategies for the  $u$ -player and the  $v$ -player, respectively, given the initial deterministic time  $s$ .

The main result in [S14b] is the description of the lower and upper values of such a **zero-sum symmetric game over elementary feedback strategies**. We recall below the result, for convenience:

**Theorem 2.2.** Under the standing assumption, we have

1. for each  $\alpha \in \mathcal{A}(s)$ ,  $\beta \in \mathcal{B}(s)$ , there exists a unique strong solution  $(X_t^{s,x;\alpha,\beta})_{s \leq t \leq T}$  (such that  $X_t^{s,x;\alpha,\beta} \in \mathcal{F}_s^W$ ) of the closed-loop state system

$$\begin{cases} dX_t = b(t, X_t, \alpha(t, X.), \beta(t, X.))dt + \sigma(t, X_t, \alpha(t, X.), \beta(t, X.))dW_t, & s \leq t \leq T, \\ X_s = x \in \mathbb{R}^d. \end{cases} \quad (2.2)$$

2. the functions

$$V^-(s, x) \triangleq \sup_{\alpha \in \mathcal{A}(s)} \inf_{\beta \in \mathcal{B}(s)} \mathbb{E}[g(X_T^{s,x;\alpha,\beta})] \leq V^+(s, x) \triangleq \inf_{\beta \in \mathcal{B}(s)} \sup_{\alpha \in \mathcal{A}(s)} \mathbb{E}[g(X_T^{s,x;\alpha,\beta})]$$

are the unique bounded continuous viscosity solutions of the Isaacs equations (for  $i = -$  and  $i = +$ ) to the game

$$\begin{cases} -v_t - H^i(t, x, v_x, v_{xx}) = 0 & \text{on } [0, T) \times \mathbb{R}^d, \\ v(T, \cdot) = g(\cdot), & \text{on } \mathbb{R}^d. \end{cases} \quad (2.3)$$

where,

$$\begin{aligned} H^-(t, x, p, M) &\triangleq \sup_{u \in U} \inf_{v \in V} L(t, x, p, M; u, v) \\ &\leq H^+(t, x, p, M) \triangleq \inf_{v \in V} \sup_{u \in U} L(t, x, p, M; u, v), \end{aligned}$$

using the notation

$$L(t, x, p, M; u, v) \triangleq b(t, x, u, v) \cdot p + \frac{1}{2} \text{Tr} (\sigma(t, x, u, v) \sigma(t, x, u, v)^T M).$$

### 2.3 Back to Control with Model Uncertainty

In our setting,  $v$  does not represent an intelligent player: we can think about it as nature, which does not have a payoff to minimize (or a vested interest from playing against player  $u$ ). The controller (player  $u$ ) does have a payoff to maximize. It is still natural to assume that, the controller only observes the state of the system, so he/she uses **the same elementary feedback strategies**  $\alpha \in \mathcal{A}(s)$ . On the other hand, the adverse player, the nature, can choose any parameter  $v$ , and, can actually do so using the whole information available in the filtration  $\mathbb{F}$ . In other words, we treat as the possible (uncertain) choices of the model to be all **open-loop** control processes  $v_t$ . We define

$$\mathcal{V}(s) \triangleq \{(v_t)_{s \leq t \leq T} | \text{predictable with respect to } \mathbb{F}\},$$

and set up the optimization problem under model uncertainty as

$$V(s, x) \triangleq \sup_{\alpha \in \mathcal{A}(s)} \inf_{v \in \mathcal{V}(s)} \mathbb{E}[g(X_T^{s,x;\alpha,v})].$$

The above formulation represents the modeling contribution of the present note. We emphasize one last time that, in our model,

- nature uses open-loop controls  $v \in \mathcal{V}(s)$ , while the controller uses feedback strategies  $\alpha \in \mathcal{A}(s)$ ,
- the nature's controls are adapted to the filtration  $\mathbb{F}$  which may be strictly larger than the one generated by the Brownian motion.

Before even studying the well posed-ness of the state equation over one feedback strategy  $\alpha$  and one open loop control  $v$ , it is expected (proven rigorously below), that  $V \leq V^-$ . The main result of the present note is:

**Theorem 2.3.** Under the standing assumption, we have

1. for each  $\alpha \in \mathcal{A}(s)$  and  $v \in \mathcal{V}(s)$ , the state equation has a unique strong solution  $(X_t^{s,x;\alpha,v})_{s \leq t \leq T}$  with  $X_t^{s,x;\alpha,v} \in \mathcal{F}_t \supset \mathcal{F}_t^W$ ,
2.  $V = V^-$  is the unique continuous viscosity solution of the lower Isaacs equation, i.e. equation (2.3) for  $i = -$ ,
3. the value function  $V$  satisfies the Dynamic Programming Principle

$$V(s, x) = \sup_{\alpha \in \mathcal{A}(s)} \inf_{v \in \mathcal{V}(s)} \mathbb{E}[V(\rho(X^{s,x;\alpha,v}), X_{\rho(X^{s,x;\alpha,v})}^{s,x;\alpha,v})] \quad \forall \rho \in \mathbb{B}^s.$$

It is important, in our view, to obtain strong solutions of the state equation, and this is the main reason to restrict feedback strategies to the class of elementary strategies. Mathematically, our result states that the use of open-loop controls by the stronger player (here, the nature), even adapted to a much larger filtration than the one generated by the “known randomness”  $W$ , does not change the value function, from the one where the stronger player only observes the state process, as long as the weaker player **only observes the state**. More precisely, the technical contribution of the note is to show that

$$\sup_{\alpha \in \mathcal{A}(s)} \inf_{v \in \mathcal{V}(s)} \mathbb{E}[g(X_T^{s,x;\alpha,v})] = \sup_{\alpha \in \mathcal{A}(s)} \inf_{\beta \in \mathcal{B}(s)} \mathbb{E}[g(X_T^{s,x;\alpha,\beta})].$$

In our understanding, this is not entirely obvious.

- Remark 2.4.*
1. one possible way to model the robust control problem is to assume that  $\alpha$  is an Elliott-Kalton strategy (like in [EK72] or [FS89]) and  $v$  is an open-loop control. While such an approach is present in the literature, we find it quite hard to justify the assumption that the controller can observe the changes in model uncertainty **in real time**, i.e. really observe  $v_t$  right at time  $t$ . Locally (over an infinitesimal time period), this amounts for the nature to **first** choose the uncertainty parameter  $v$ , **then, after observing**  $v$  for the controller to choose  $u$ . This contradicts the very idea of Knightian uncertainty we have in mind. If one actually went ahead and modeled our control problem in such a way, than  $V$  would be equal to  $V^+$ , since the Elliott-Kalton player is the stronger player as described above (see [FS89] for the mathematics, under stronger assumptions on the system).
  2. another way would be to model the “nature” as the Elliott-Kalton strategy player  $\beta$  and let the controller/intelligent player use open loop controls  $u$ . This does not seem too appealing either, since nature does not have any payoff/vested interest. Why would nature be able to **observe the controller’s actions and act strategically against** him/her? In addition, if the controller chooses open-loop controls, he/she needs to have the whole information in  $\mathbb{F}$  available. The controller does not usually observe directly even the noise  $W$ , leave alone the other possible information in  $\mathbb{F}$ . However, with such a model, mathematically, the resulting value function is expected to be the same,  $V = V^-$  (see, again, [FS89], up to technical details).

### 3 Proofs

The proposition below contains the proof of the first item in Theorem 2.3.

**Proposition 3.1.** Fix  $s, x$  and  $\alpha \in \mathcal{A}(s)$  and  $v \in \mathcal{V}(s)$ . Then, there exists a unique strong (and square integrable) solution  $(X_t^{s,x;\alpha,v})_{s \leq t \leq T}$ ,  $X_t^{s,x;\alpha,v} \in \mathcal{F}_t$  of the state equation

$$\begin{cases} dX_t = b(t, X_t, \alpha(t, X.), v_t)dt + \sigma(t, X_t, \alpha(t, X.), v_t) dW_t, & s \leq t \leq T, \\ X_s = x \in \mathbb{R}^d. \end{cases} \quad (3.1)$$

Proof: The proof of the above proposition (both existence and uniqueness) is based on solving the equation, successively on  $[\tau_k(X^{s,x;\alpha,v}), \tau_{k+1}(X^{s,x;\alpha,v})]$  for  $k = 1, \dots, n$ . The details are rather obvious and, even in [S14b], the proof of a similar lemma (where both players choose elementary feedback strategies, unlike here) was only sketched.  $\diamond$

Before we proceed, let  $\alpha \in \mathcal{A}(s), \beta \in \mathcal{B}(s)$ . We can consider

$$(v_t)_{s \leq t \leq T} \triangleq (\beta(t, X^{s,x;\alpha,\beta}))_{s \leq t \leq T} \in \mathcal{V}(s),$$

such that  $X^{s,x;\alpha,\beta} = X^{s,x;\alpha,v}$ . This means that, for a fixed  $\alpha$ , there are more open-loop controls nature can use, than feedback strategies an adverse zero-sum player could use. This shows that

$$V(s, x) = \sup_{\alpha \in \mathcal{A}(s)} \inf_{v \in \mathcal{V}(s)} \mathbb{E}[g(X_T^{s,x;\alpha,v})] \leq \sup_{\alpha \in \mathcal{A}(s)} \inf_{\beta \in \mathcal{B}(s)} \mathbb{E}[g(X_T^{s,x;\alpha,\beta})] = V^-(s, x).$$

The goal is to prove the inequality above is actually a true equality. The proof of the main Theorem 2.3 relies on a similar adaptation of the Perron's Method that was introduced in [S14b] for symmetric zero-sum games played over elementary feedback strategies. As mentioned, the present note is a technical companion to [S14b]. The main (but not only) technical difference is that the stochastic sub-solutions of the robust control problem need to be defined differently, to account for the fact that the adverse player is using open-loop controls.

Following [S14b], we first define elementary feedback strategies starting at sequel times to the initial (deterministic) time  $s$ . The starting time is a stopping rule.

**Definition 3.2** (Elementary Strategies starting later). Fix  $s$  and let  $\tau \in \mathbb{B}^s$  be a stopping rule. An elementary strategy, denoted by  $\alpha \in \mathcal{A}(s, \tau)$ , for the first player, starting at  $\tau$ , is defined by

- (again) a finite non-decreasing sequence of stopping rules, i.e.  $\tau_k \in \mathbb{B}^s, k = 1, \dots, n$  for some finite  $n$ , and with  $\tau = \tau_0 \leq \dots \leq \tau_k \leq \dots \leq \tau_n = T$ .
- for each  $k = 1 \dots n$ , a constant action  $\xi_k$  in between the times  $\tau_{k-1}$  and  $\tau_k$ , which is decided based only on the knowledge of the past state up  $\tau_{k-1}$ , i.e.  $\xi_k : C([s, T]) \rightarrow U$  such that  $\xi_k \in \mathcal{B}_{\tau_{k-1}}^s$ .

The strategy is, again, to hold  $\xi_k$  in between  $(\tau_{k-1}, \tau_k]$ , i.e..

$$\alpha : \{(t, y) | \tau(y) < t \leq T, y \in C([s, T])\} \rightarrow U$$

with

$$\alpha(t, y(\cdot)) \triangleq \sum_{k=1}^n \xi_k(y(\cdot)) \mathbb{1}_{\{\tau_{k-1}(y(\cdot)) < t \leq \tau_k(y(\cdot))\}}.$$

The notation is consistent with  $\mathcal{A}(s) = \mathcal{A}(s, s)$ .

We recall, still from [S14b], that strategies in  $\mathcal{A}(s, \tau)$  cannot be used by themselves for the game starting at  $s$ , but have to be concatenated with other strategies.

**Proposition 3.3** (Concatenated elementary feedback strategies). Fix  $s$  and let  $\tau \in \mathbb{B}^s$  be a stopping rule and  $\tilde{\alpha} \in \mathcal{A}(s, \tau)$ . Then, for each  $\alpha \in \mathcal{A}(s, s)$ , the mapping  $\alpha \otimes_{\tau} \tilde{\alpha} : (s, T] \times C([s, T]) \rightarrow U$  defined by

$$(\alpha \otimes_{\tau} \tilde{\alpha})(t, y(\cdot)) \triangleq \alpha(t, y(\cdot)) \mathbb{1}_{\{s < t \leq \tau(y(\cdot))\}} + \tilde{\alpha}(t, y(\cdot)) \mathbb{1}_{\{\tau(y(\cdot)) < t \leq T\}}$$

is a simple strategy starting at  $s$ , i.e.  $\alpha \otimes_{\tau} \tilde{\alpha} \in \mathcal{A}(s, s)$ .

Compared to [S14b] the definition below has to be carefully modified.

**Definition 3.4** (Stochastic Sub-Solution). A function  $w : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$  is called a stochastic sub-solution of the lower Isaacs equation, i.e. equation (2.3) for  $i = -$ , if

1. it is bounded, continuous and  $w(T, \cdot) \leq g(\cdot)$ ,
2. for each  $s$  and for each stopping rule  $\tau \in \mathbb{B}^s$  there exists an elementary strategy  $\tilde{\alpha} \in \mathcal{A}(s, \tau)$  such that, for any  $\alpha \in \mathcal{A}(s)$ , any  $v \in \mathcal{V}(s)$ , any  $x$  and each stopping rule  $\rho \in \mathbb{B}^s$ ,  $\tau \leq \rho \leq T$ , with the simplifying notation  $X \triangleq X^{s,x,\alpha \otimes \tau \tilde{\alpha}, v}$  and  $\tau' \triangleq \tau(X)$ ,  $\rho' \triangleq \rho(X)$ , we have

$$w(\tau', X_{\tau'}) \leq \mathbb{E}[w(\rho', X_{\rho'}) | \mathcal{F}_{\tau'}] \quad \mathbb{P} - a.s.$$

Let  $w$  a stochastic sub-solution. Fix  $s$ . There exists  $\tilde{\alpha} \in \mathcal{A}(s)$  such that, for each  $x$ , each  $\rho \in \mathbb{B}^s$  and each  $v \in \mathcal{V}(s)$  we have

$$w(s, x) \leq \mathbb{E} \left[ w(\rho(X^{s,x,\tilde{\alpha},v}), X_{\rho(X^{s,x,\tilde{\alpha},v})}^s) | \mathcal{F}_s \right], \quad \mathbb{P} - a.s. \quad (3.2)$$

Taking the expectation it is obvious that, if  $w$  is a stochastic sub-solution, then we have the half DPP/sub-optimality principle

$$w(s, x) \leq \sup_{\alpha \in \mathcal{A}(s)} \inf_{v \in \mathcal{V}(s)} \mathbb{E} \left[ w(\rho(X^{s,x,\alpha,v}), X_{\rho(X^{s,x,\alpha,v})}^s) \right], \quad \forall \rho \in \mathbb{B}^s. \quad (3.3)$$

Since  $w(T, \cdot) \leq g(\cdot)$ , we obtain  $w(s, x) \leq V(s, x) \leq V^-(s, x)$ .

We have already characterized  $V^-$  as the unique solution of the lower Isaacs equation in [S14b]. Therefore, we actually need **only half** of the Perron construction here. We denote by  $\mathcal{L}$  the set of stochastic sub-solutions in Definition 3.4 (non-empty from the boundedness assumptions). Define

$$w^- \triangleq \sup_{w \in \mathcal{L}} w \leq V \leq V^-.$$

**Proposition 3.5** (Stochastic Perron for Robust Control). Under the standing assumptions,  $w^-$  is a LSC viscosity super-solution of the lower Isaacs equation, up to  $t = 0$ .

The following lemmas are very similar to their counterparts in [S14b].

**Lemma 3.6.** If  $w_1, w_2 \in \mathcal{L}$  then  $w_1 \vee w_2 \in \mathcal{L}$ .

Fix  $\tau \in \mathbb{B}^s$  a stopping rule. Let  $\tilde{\alpha}_1, \tilde{\alpha}_2 \in \mathcal{A}(s, \tau)$  be the two feedback strategies of the controller, starting at  $\tau$  corresponding the the sub-solutions  $w_1$  and  $w_2$  for the Definition 3.4. The new strategy starting at  $\tau$ , defined for any  $y \in C([s, T])$  by

$$\begin{aligned} \tilde{\alpha}(t, y) &= \tilde{\alpha}_1(t, y) \mathbb{1}_{\{w_1(\tau(y), y(\tau(y))) \geq w_2(\tau(y), y(\tau(y)))\}} \\ &\quad + \tilde{\alpha}_2(t, y) \mathbb{1}_{\{w_1(\tau(y), y(\tau(y))) < w_2(\tau(y), y(\tau(y)))\}} \end{aligned}$$

does the job for the definition of  $w \triangleq w_1 \vee w_2$  as a stochastic sub-solution .  $\diamond$

**Lemma 3.7.** There exists a non-decreasing sequence  $\mathcal{L} \ni w_n \nearrow w^-$ .

Proof: according to Proposition 4.1 in [BS12], there exist  $\tilde{w}_n \in \mathcal{L}$  such that  $w^- = \sup_n \tilde{w}_n$ . Now, we can just define  $w_n = \tilde{w}_1 \vee \dots \vee \tilde{w}_n \in \mathcal{L} \nearrow w^-$ .  $\diamond$

**Proof of Proposition 3.5** The proof is similar to [S14b]. Since Itô formula applies the same, **regardless of filtration**, it produces sub-martingales in a similar way, even though  $v$  is an open-loop control, and the filtration may be larger than the one generated by  $W$ . This is the key point that allows us to obtain the result. We only sketch some key points of the proof, in order to avoid repeating all the similar arguments in [S14b].

The interior super-solution property for  $w^-$ : Let  $(t_0, x_0)$  in the parabolic interior  $[0, T] \times \mathbb{R}^d$  such that a smooth function  $\varphi$  strictly touches  $w^-$  from below at  $(t_0, x_0)$ .

Assume, by contradiction, that  $\varphi_t + H^-(t, x, \varphi_x, \varphi_{xx}) > 0$  at  $(t_0, x_0)$ . In particular, there exists  $\hat{u} \in U$  and  $\varepsilon > 0$  such that

$$\begin{aligned} & \varphi_t(t_0, x_0) + \inf_{v \in V} \left[ b(t_0, x_0, \hat{u}, v) \cdot \varphi_x(t_0, x_0) \right. \\ & \left. + \frac{1}{2} \text{Tr}(\sigma(t_0, x_0, \hat{u}, v) \sigma(t_0, x_0, \hat{u}, v)^T \varphi_{xx}(t_0, x_0)) \right] > \varepsilon. \end{aligned}$$

To simplify notation, all small balls here are actually included in (i.e. intersected with) the parabolic interior. Since  $b, \sigma$  are continuous, and  $V$  is compact, the uniform continuity of the above expression in  $(t, x, v)$  for  $(t, x)$  around  $(t_0, x_0)$  implies that there exists a smaller  $\varepsilon > 0$  such that

$$\varphi_t(t, x) + \inf_{v \in V} \left[ b(t, x, \hat{u}, v) \cdot \varphi_x(t, x) + \frac{1}{2} \text{Tr}(\sigma(t, x, \hat{u}, v) \sigma(t, x, \hat{u}, v)^T \varphi_{xx}(t, x)) \right] > \varepsilon,$$

on  $B(t_0, x_0, \varepsilon)$ . Now, on the compact (rectangular) torus  $\mathbb{T} = \overline{B(t_0, x_0, \varepsilon)} - B(t_0, x_0, \varepsilon/2)$  we have that  $\varphi < w^-$  and the max of  $\varphi - w^-$  is attained, therefore it is strictly negative. In other words  $\varphi < w^- - \eta$  on  $\mathbb{T}$  for some  $\eta > 0$ . Since  $w_n \nearrow w^-$ , a Dini type argument similar to [BS14] and [BS13] shows that, for  $n$  large enough we have  $\varphi < w_n - \eta/2$  on  $\mathbb{T}$ . For simplicity, fix such an  $n$  and call  $v = w_n$ . Now, define, for small  $\delta \ll \eta/2$

$$v^\delta \triangleq \begin{cases} (\varphi + \delta) \vee v & \text{on } B(t_0, x_0, \varepsilon), \\ v & \text{outside } B(t_0, x_0, \varepsilon). \end{cases}$$

Since  $v^\delta(t_0, x_0) > w^-(t_0, x_0)$ , we have a contradiction if  $v^\delta \in \mathcal{L}$ . To begin with, we emphasize that  $v^\delta = v$  on  $\mathbb{T}$ . Fix  $s$  and let  $\tau \in \mathbb{B}^s$  be a stopping rule for the initial time  $s$ . We need to construct an elementary strategy  $\tilde{\alpha} \in \mathcal{A}(s, \tau)$  in the Definition 3.4 of stochastic sub-solution for  $v^\delta$ . We do that as follows:

1. with the notation  $\partial \triangleq \partial B(t_0, x_0, \varepsilon/2)$  we define the stopping rule  $\tau_1 : C([s, T]) \rightarrow [s, T]$ ,  $\tau \leq \tau_1 \in \mathbb{B}_s$  by

$$\tau_1(y) \triangleq \begin{cases} \tau(y), & \text{if } v(\tau(y), y(\tau(y))) = v^\delta(\tau(y), y(\tau(y))) \\ \inf\{t \mid \tau(y) \leq t \leq T, (t, y(t)) \in \partial\}, & \text{if } v(\tau(y), y(\tau(y))) < v^\delta(\tau(y), y(\tau(y))). \end{cases}$$

Recall that  $v = v^\delta$  on  $\partial$ .

2. starting at  $\tau$  and up to the stopping rule  $\tau_1$ , follow the constant action  $\hat{u}$
3. starting at  $\tau_1$ , (when  $v^\delta = v$ , by construction) follow the strategy  $\alpha_1 \in \mathcal{A}(s, \tau_1)$  corresponding to the Definition 3.4 of the stochastic sub-solution  $v$  with respect to the starting stopping rule  $\tau_1$ .

More precisely, we define the strategy  $\tilde{\alpha} : \{(t, y) \mid \tau(y) < t \leq T, y \in C([s, T])\} \rightarrow U$ , by

$$\tilde{\alpha}(t, y) \triangleq \hat{u} 1_{\{\tau(y) < t \leq \tau_1(y)\}} + \alpha_1(t, y) 1_{\{\tau_1(y) < t \leq T\}}.$$

We follow similar arguments to [S14b] to show that  $\tilde{\alpha} \in \mathcal{A}(s, \tau)$  is a strategy that satisfies the Definition 3.4 for  $v^\delta$  as a stochastic sub-solution corresponding to the stopping rule  $\tau$ . Actually, the construction of  $\tilde{\alpha}$  above is slightly simpler than the construction in [S14b] (and fixes some typos there). To summarize, we obtained that  $v^\delta \in \mathcal{L}$ , so we reached a contradiction. The terminal condition property for  $w^-$  is proved very similarly.  $\diamond$

**Proof of Theorem 2.3:** Recall that the first part was proved by Proposition 3.1.

Next, the proof of the second item is finished, once we use the comparison result from Lemma 4.1 in [S14b]. More precisely, we know that  $w^- \leq V \leq V^-$  and  $w^-$  is a viscosity super-solution and  $V^-$  is a viscosity solution of the lower Isaacs equation



(from [S14b][Theorem 4.1]). Therefore, according to [S14b][Lemma 4.1], we also have  $V^- \leq w^-$ , so  $V = V^-$  is the unique viscosity solution.

Finally, the DPP in Item 3 of Theorem 2.3 is actually an easy observation based on the fact that the value function  $V^- = V$  satisfies a similar (but not identical) DPP (see [S14b]) and the half DPP (3.2). $\diamond$

#### 4 Additional Modeling Comments

In our (strong) model of robust control, the value function of the intelligent player turns out to be  $V = V^-$ . Obviously, one can ask the question: should this player try to randomize feedback strategies somehow, to get the potentially better value  $V^{mix}$  of the value over mixed strategies (for both players) obtained in [S14a] (but in a martingale symmetric formulation)?

Modeling mixed feedback strategies for the controller, and open loop-strategies controls for the adverse player is a highly non-trivial issue, and not obviously possible in strong formulation (see [S14a] for some comments along these lines, for the case of a zero-sum symmetric game). In our formulation of optimization with model uncertainty, the maximizing player has to settle with the value  $V = V^-$ . However, the controller couldn't do better anyway in one of the two situations:

1. when the Isaacs condition over pure strategies is satisfied, i.e.

$$\sup_{u \in U} \inf_{v \in V} L(t, x, p, M; u, v) = \inf_{v \in V} \sup_{u \in U} L(t, x, p, M; u, v)$$

so  $V^- = V^{mix} = V^+$

2. in any additional situation when  $V^- = V^{mix} < V^+$ , i.e. all situations in which (even at the formal level) potential randomization for the  $u$  player does not change the Hamiltonian. More precisely, if

$$\sup_{u \in U} \inf_{\nu \in \mathcal{P}(V)} \int L(t, x, p, M, u, v) \nu(dv) = \inf_{\nu \in \mathcal{P}(V)} \sup_{u \in U} \int L(t, x, p, M, u, v) \nu(dv),$$

since

$$\sup_{u \in U} \inf_{v \in V} L(t, x, p, M; u, v) = \sup_{u \in U} \inf_{\nu \in \mathcal{P}(V)} \int L(t, x, p, M, u, v) \nu(dv)$$

and

$$\inf_{\nu \in \mathcal{P}(V)} \sup_{u \in U} \int L(t, x, p, M, u, v) \nu(dv) = \inf_{\nu \in \mathcal{P}(V)} \sup_{\mu \in \mathcal{P}(U)} \int L(t, x, p, M, u, v) \mu(dv) \nu(dv)$$

we have

$$H^- = H^{mix} \leq H^+$$

although the Isaacs condition over pure strategies may not be satisfied ( $H^- < H^+$ ). In such a situation, the robust controller cannot expect to get a better value than  $V = V^-$ . A sufficient condition for this is for the map

$$u \rightarrow L(t, x, p, M; u, v)$$

to be concave. Up to different modeling of strategies, this is exactly the case in the interesting recent contribution [TTU13].

## References

- [BS12] E. Bayraktar and M. Sirbu, *Stochastic Perron's method and verification without smoothness using viscosity comparison: the linear case*, Proceedings of the American Mathematical Society **140** (2012), 3645–3654. MR-2929032
- [BS13] ———, *Stochastic Perron's method for Hamilton-Jacobi-Bellman equations*, SIAM Journal on Control and Optimization **51** (2013), no. 6, 4274–4294. MR-3124891
- [BS14] ———, *Stochastic Perron's method and verification without smoothness using viscosity comparison: obstacle problems and Dynkin games*, Proceedings of the American Mathematical Society **142** (2014), no. 4, 1399–1412. MR-3162260
- [EK72] R. J. Elliott and N. J. Kalton, *Values in differential games*, Bull. Amer. Math. Soc **72** (1972), no. 3, 427–431. MR-0295775
- [FS89] W. H. Fleming and P. E. Souganidis, *On the existence of value functions of two-player, zero-sum stochastic differential games*, Indiana University Mathematics Journal **38** (1989), no. 2, 293–314. MR-0997385
- [KS88] N. N. Krasovskii and A. I. Subbotin, *Game-theoretical control problems*, Springer Series in Soviet Mathematics, Springer-Verlag, New York, 1988, Translated from the Russian by Samuel Kotz. MR-0918771
- [KS01] I. Karatzas and W. Sudderth, *The controller-and-stopper game for a linear diffusion*, Annals of Probability **29** (2001), no. 3, 1111–1127. MR-1872738
- [S14a] M. Sirbu, *On martingale problems with continuous-time mixing and values of zero-sum games without Isaacs conditions*, SIAM Journal on Control and Optimization **52** (2014), no. 5, 2877–2890. MR-3261608
- [S14b] ———, *Stochastic Perron's method and elementary strategies for zero-sum differential games*, SIAM Journal on Control and Optimization **52** (2014), no. 3, 1693–1711. MR-3206980
- [TTU13] Revaz Tevzadze, Teimuraz Toronjadze, and Tamaz Uzunashvili, *Robust utility maximization for a diffusion market model with misspecified coefficients*, Finance Stoch. **17** (2013), no. 3, 535–563. MR-3066987