Concavity and optimality conditions for continuous time principal-agent problems

Lawrence C. Evans*, Christopher W. Miller†
Mathematics Department
University of California, Berkeley

Insoon Yang‡
Laboratory for Information and Decision Systems
Massachusetts Institute of Technology

Abstract

We present a simple convexity argument that clarifies the effectiveness and scope of Sannikov's approach [S] to continuous time principal-agent problems. We in particular stress the importance of the appropriate concavity of the agent's running payoff function as a function of his actions.

1 Introduction

This paper provides a simple reformulation and reinterpretation of Sannikov's approach [S] to principal-agent problems.

In so-called second-best continuous time principal-agent problems, the agent's effort affects an output process, which is subject also to external noise. The principal however cannot directly observe the agent's effort and can only monitor the stochastic output process. The agent on the other hand has perfect observations. With this asymmetry of information, the principal wants to offer a contract to the agent, comprising both a compensation scheme

*Class of 1961 Collegium Chair. Supported in part by NSF Grant DMS-1301661 and the Miller Institute for Basic Research in Science
†Supported by NSF Grant DGE 1106400
‡Supported by NSF CPS FORCES Grant 1239166

and a recommended effort strategy for the agent, such that (i) the agent has an incentive to enter in the contract and (ii) the agent has no incentive to deviate from the recommended effort strategy given the compensation scheme. The former and latter are, respectively, the agent's *individual rationality* and *incentive compatibility* conditions. The intention is that through proper contracting, the principal can incentivize the agent to control the output process in a way that is beneficial to the principal.

This work provides a reinterpretation of several ideas from Sannikov's breakthrough paper [S] on such principal-agent problems. Our easy proof of incentive compatibility for the agent makes very clear the need for the concavity hypothesis (2.10) and furthermore does not employ changes of measure nor backward SDE (as for instance in Cvitanic–Zhang [C-Z]; see also Cvitanic–Possamai–Touzi [C-P-T]). Our method uses very little probability theory and easily generalizes to higher dimensional problems with general dynamics, and even to the case in which the agent's effort affects the noise in the output process. We consider as well principal-agent problems in which multiple agents simultaneously affect the output process. We show that by using HJB equations, principal can design a contract, a compensation scheme and a recommended effort strategy for each agent, such that the recommended efforts constitute a Nash equilibrium.

We thank I. Ekeland for several email communications, and especially for pointing out errors in a much earlier version of this paper. The paper [M-Y] gives further applications and extensions of our methods.

2 A model problem

In this section we consider a fairly simple one-dimensional problem, to explain our approach in the clearest possible setting. We work on the finite time interval $[0, T]$. We hereafter use $A = A(t)$ to denote the agent's actions, controlling the dynamics, and $P = P(t)$ to denote the principal's ongoing payoffs to the agent, incentivizing him to take certain desired actions. The letter R denotes a payoff reward from the principal to the agent at the terminal time T.

2.1 Statement of the problem.

Dynamics. In this first subsection, we follow Sannikov [S] and assume that the dynamics depend linearly on the agent's effort A and that the noise term depends on neither X nor A:

$$\begin{cases} dX = Adt + dB \qquad (0 \leq t \leq T) \\ X(0) = x. \end{cases} \tag{2.1}$$

Here B is a one-dimensional Brownian motion and we will sometimes write $X = X^A$ to emphasize that the dependence on the control A.

We let $\mathcal{F} = \mathcal{F}^B = \{\mathcal{F}(t) \mid 0 \le t \le T\}$ be the filtration generated by the Brownian motion B, and write $\mathcal{F}^X = \{\mathcal{F}^X(t) \mid 0 \le t \le T\}$ for the filtration generated by the process X.

Payoffs. The payoff J_π for the principal and the payoff J_α for the agent have the forms

$$\begin{cases} J_\pi[A,P,R] := E\Big(\int_0^T r^\pi(A,P)\,dt + q(R)\Big) \\ J_\alpha[A,P,R] := E\Big(\int_0^T r^\alpha(A,P)\,dt + R\Big), \end{cases} \tag{2.2}$$

where $r^\pi = r^\pi(a,p)$ denotes the principal's running payoff and $r^\alpha = r^\alpha(a,p)$ the agent's running payoff. The function $q = q(r)$ is the negative of the cost to the principal of providing at time T a final payout of size r to the agent.

The principal's problem. The goal of the principal is to design A^*, P^*, R^* such that

$$J_\pi[A^*,P^*,R^*] = \max_{A,P,R} J_\pi[A,P,R] \tag{2.3}$$

subject to the constraints that

$$P, R \text{ are } \mathcal{F}^X\text{-adapted,} \tag{2.4}$$

and

$$J_\alpha[A^*,P^*,R^*] = \max_{A} J_\alpha[A,P^*,R^*], \tag{2.5}$$

the maximum taken over $\mathcal{F}$-adapted controls A. This is the agent's incentive compatibility condition.

The key point is that whereas the agent's best control A^* is $\mathcal{F}$-adapted, *the principal's controls P^* and R^* must be $\mathcal{F}^X$-adapted*; that is, the principal must decide what to do based only upon her observations of $X = X^A$ and not upon B, which is unobservable to her.

How can the principal create incentives for the agent to do what the principal wants? In particular, how can the principal encourage the agent to perform a desired action A^*? To repeat, the agent's actions and the Brownian motion are unobservable by the principal; she can only observe X.

2.2 The contract, optimality for the agent. The following payment scheme is inspired by Sannikov [S]; see also Ekeland [E].

Let

$$y(a,p) := -r^\alpha_a(a,p) \tag{2.6}$$

denote the negative of the partial derivative of r^α in the first variable a. Given the smooth, deterministic functions $a = a(w,t)$ and $p = p(w,t)$, we consider the SDE

$$\begin{cases} dW = -(r^\alpha(A^*,P^*) + Y^*A^*)dt + Y^*dX \qquad (0 \le t \le T) \\ W(0) = w. \end{cases} \tag{2.7}$$

where

$$A^* := a(W,t),\ P^* := p(W,t),\ Y^* := y(A^*,P^*). \tag{2.8}$$

Note that the solution of this SDE depends upon the agent's actions A, as recorded in the "dX" term.

The principal agrees to pay the agent at the ongoing rate $P^*(t) := p(W(t),t)$, which at each time t she can compute in terms of the observed dynamics $\{X(s) \mid 0 \le s \le t\}$. In addition, the principal agrees to give the agent the final reward payout

$$R^* := W(T) \tag{2.9}$$

at the terminal time T.

THEOREM 2.1. *Assume the principal uses the control P^* and final payoff R^*, as above.*

(i) *Suppose also*

$$a \mapsto r^\alpha(a,p) \text{ is concave for each } p. \tag{2.10}$$

Then it is optimal for the agent to use the control $A^* = a(W,t)$, and his payoff is then w.

(ii) *Conversely, if the concavity condition* (2.10) *fails, then there exist $A^* := a(W,t)$ and $P^* := p(W,t)$ such that A^* is not optimal for the agent.*

Assertion (i) is a form of Sannikov's optimality condition for agent.

Proof. 1. Let A denote any admissible control for the agent. Since $dX = Adt + dB$, we have

$$dW = (-r^\alpha(A^*,P^*) + Y^*(A - A^*))dt + Y^* dB.$$

Integrate:

$$W(T) - w = \int_0^T -r^\alpha(A^*,P^*) + Y^*(A - A^*)\,dt + \int_0^T Y^*\,dB.$$

We now take expected values and rewrite, recalling that $Y^* = y(A^*,P^*) = -r_a^\alpha(A^*,P^*)$:

$$\begin{aligned} w = E&\left(\int_0^T r^\alpha(A,P^*)\,dt + R^*\right) \\ &+ E\left(\int_0^T r^\alpha(A^*,P^*) + r_a^\alpha(A^*,P^*)(A - A^*) - r^\alpha(A,P^*)\,dt\right). \end{aligned} \tag{2.11}$$

2. When (2.10) holds, the concavity of r^α in the variable a implies that the last term in (2.11) is nonnegative, and is zero for $A = A^* = a(W,t)$. Therefore

$$\begin{aligned}
J_\alpha[A^*, P^*, R^*] &= E\left(\int_0^T r^\alpha(A^*, P^*)\, dt + R^*\right) \\
&= w \\
&\geq E\left(\int_0^T r^\alpha(A, P^*)\, dt + R^*\right) \\
&= J_\alpha[A, P^*, R^*]
\end{aligned}$$

for any other admissible control A the agent may employ. Consequently, the agent's incentive condition (2.5) holds.

3. If the concavity assumption (2.10) fails, then for appropriate A^* and P^* the agent can select an action A for which

$$E\left(\int_0^T r^\alpha(A^*, P^*) + r_a^\alpha(A^*, P^*)(A - A^*) - r^\alpha(A, P^*)\, dt\right) < 0;$$

in which case

$$J_\alpha[A^*, P^*, R^*] = w < J_\alpha[A, P^*, R^*].$$

□

This easy proof makes obvious the necessity of the concavity condition (2.10). Most of the rest of our paper consists of straightforward extensions of the foregoing argument.

Remarks. (i) In this first, simple model it is possible that $W(T)$ is negative; in which case we are assuming that the principal can extract from the agent a payment of size $-W(T) > 0$.

(ii) If we make any other choice of Y^* in the dynamics (2.7) for W, the analog of the last term of (2.11) would read

$$E\left(\int_0^T r^\alpha(A^*, P^*) - Y^*(A - A^*) - r^\alpha(A, P^*)\, dt\right).$$

If Y^* were not as defined in (2.6), (2.8), we could make this expression negative for a suitable choice of A, and thus A^* would not be optimal for the agent. □

2.3 Optimality for the principal. The principal's value function $u = u(w,t)$ is

$$u(w,t) = \sup_{A,P,R} J_\pi[A, P, R], \tag{2.12}$$

where $w \in \mathbb{R}$, $0 \leq t \leq T$,

$$\begin{cases} dW = -(r^\alpha(A^*, P^*) + Y^* A^*)ds + Y^* dX \qquad (t \leq s \leq T) \\ W(t) = w, \end{cases} \tag{2.13}$$

and A, P, R satisfy the analogs of (2.4), (2.5).

THEOREM 2.2. *The principal's value function solves the Hamilton–Jacobi–Bellman (HJB) equation*

$$\begin{cases} u_t + \sup_{a,p} \left\{ \frac{1}{2}(r^\alpha_a(a,p))^2 u_{ww} - r^\alpha(a,p)u_w + r^\pi(a,p) \right\} = 0 \quad \text{on } \mathbb{R} \times [0,T) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad u = q \quad \text{on } \mathbb{R} \times \{t = T\} \end{cases} \tag{2.14}$$

Note that for this particular problem the value function does not depend upon the initial position x of X. The principal selects functions $a = a(w,t)$, $p = p(w,t)$ giving the max in the HJB equation and then sets $A^* := a(W,t)$, $P^* := p(W,t)$, $R^* = W(T)$ as above. (Note very carefully that we are making the strong assumptions (i) that the terminal value problem (2.14) has a sufficiently smooth solution u and (ii) that we can find sufficiently smooth functions a and p that the SDE (2.7), (2.8) has a unique solution. For many examples this need not be so, however, and we hope in future work to extend our methods to such problems.)

Finally, the principal selects, if she can, a point $w^* \in \mathbb{R}$ such that

$$u(w^*, 0) = \max_{\mathbb{R}} u(w, 0) \tag{2.15}$$

to use as the initial condition for the dynamics (2.7):

$$\begin{cases} dW^* = -(r^\alpha(A^*, P^*) + Y^* A^*)dt + Y^* dX \qquad (0 \leq t \leq T) \\ W^*(0) = w^*. \end{cases}$$

Then

$$J_\pi[A^*, P^*, R^*] = u(w^*, 0), \quad J_\alpha[A^*, P^*, R^*] = w^*.$$

We describe the agent's HJB equation in Appendix 1 and discuss more about optimality in Appendix 2.

2.4 What if $a \mapsto r^\alpha$ is not concave? As noted above, if $a \mapsto r^\alpha(a,p)$ is not concave, then for some choices of A^* and P^* it will be possible for the agent to select A for which the last term in (2.11) is negative. In this situation A^* is not optimal for the agent.

But even in this setting *it is still possible for the principal to incentivize the agent to perform certain actions A^**. To see this, let

$$\bar{r}^\alpha(\cdot, p)$$

denote the *concave envelope* of $r^\alpha(\cdot,p)$ for each value of p, the smallest function concave in a that is greater than or equal to $r^\alpha(\cdot,p)$. Write

$$\Gamma := \{(a,p) \mid r^\alpha(a,p) = \bar{r}^\alpha(a,p)\} \tag{2.16}$$

for the set where r^α touches its concave envelope.

Assume for simplicity of exposition that $\bar{r}^\alpha$ is smooth and *redefine*

$$y(a,p) := -\bar{r}^\alpha_a(a,p). \tag{2.17}$$

Given functions $a(w,t)$ and $p(w,t)$, we again solve the SDE (2.7), where

$$Y^* := y(A^*, P^*)$$

for y now given by (2.17).

THEOREM 2.3. *Assume the principal uses the control P^* and final payoff $R^* := W(T)$, as before. Suppose also*

$$(A^*, P^*) \in \Gamma \quad \textit{almost surely.} \tag{2.18}$$

Then it is optimal for the agent to use the control $A^* = a(W,t)$.

Since the original probability measure and P^A are mutually absolutely continuous, the phrase "almost surely" above means with respect to either of these measures.

Proof. As in the proof of Theorem 2.1,

$$\begin{aligned} w = E\left(\int_0^T r^\alpha(A,P^*)\,dt + R^*\right) \\ + E\left(\int_0^T r^\alpha(A^*,P^*) + \bar{r}^\alpha_a(A^*,P^*)(A-A^*) - r^\alpha(A,P^*)\,dt\right). \end{aligned}$$

But $r^\alpha(A^*,P^*) = \bar{r}^\alpha(A^*,P^*)$, owing to (2.18), and therefore

$$\begin{aligned} r^\alpha(A^*,P^*) + \bar{r}^\alpha_a(A^*,P^*)(A-A^*) = \bar{r}^\alpha(A^*,P^*) + \bar{r}^\alpha_a(A^*,P^*)(A-A^*) \\ \geq \bar{r}^\alpha(A,P^*) \geq r^\alpha(A,P^*). \end{aligned}$$

Consequently, the last term above is nonnegative, and is zero for $A = A^*$; thus A^* is optimal for the agent. $\square$

In this case the principal's value function $u = u(w, t)$ solves the modified HJB equation

$$\begin{cases} u_t + \sup_{(a,p)\in\Gamma} \left\{ \frac{1}{2}(r_a^\alpha(a,p))^2 u_{ww} - r^\alpha(a,p)u_w + r^\pi(a,p) \right\} = 0 & \text{on } \mathbb{R} \times [0,T) \\ u = q & \text{on } \mathbb{R} \times \{t = T\}. \end{cases}$$

The additional proviso that $(a, p) \in \Gamma$ diminishes the controls available to the principal, as compared with those computed from the unconstrained HJB equation (2.14). Her value function can therefore be no larger than that computed by solving (2.14), and may well be smaller. □

2.5 Individual rationality constraint for the agent. The model discussed above is extremely simple, but is economically unrealistic in that the values of W may become negative, at which point the agent has no incentive to continue with the contract. We in this section modify the model to allow for possible early termination of the contract in this eventuality.

We keep the same dynamics (2.1) for X, but now extend the controls available to the principal to include an $\mathcal{F}^X$-adapted stopping time τ, at which the contract with the agent ends:

$$\begin{cases} J_\pi[A, P, R, \tau] := E\Big(\int_0^{T\wedge\tau} r^\pi(A, P)\, dt + q(R(T \wedge \tau))\Big) \\ J_\alpha[A, P, R, \tau] := E\Big(\int_0^{T\wedge\tau} r^\alpha(A, P)\, dt + R(T \wedge \tau)\Big). \end{cases} \tag{2.19}$$

In these expressions $\wedge$ means "min" and R now denotes a time-dependent termination payoff to the agent. The dynamics for W are (2.7), as before. Now let

$$\tau^* = \text{first time } W \text{ hits } \{w = 0\}. \tag{2.20}$$

and

$$R^*(T \wedge \tau^*) := W(T \wedge \tau^*). \tag{2.21}$$

So the payoff R^* is 0 if $0 \le \tau^* \le T$, that is, if the W process hits $\{w = 0\}$ before the terminal time T.

THEOREM 2.4. *Assume the principal uses P^*, R^* and τ^*, as above. Suppose also that $a \mapsto r^\alpha(a, p)$ is concave for each p.*

Then it is optimal for the agent to use the control $A^* = a(W, t)$, and his payoff is w.

Proof. The proof is as before, except that now we integrate to the time $T \wedge \tau^*$:

$$\begin{aligned} w = {} & E\left(\int_0^{T\wedge\tau^*} r^\alpha(A, P^*)\, dt + R^*(T \wedge \tau^*)\right) \\ & + E\left(\int_0^{T\wedge\tau^*} r^\alpha(A^*, P^*) + r_a^\alpha(A^*, P^*)(A - A^*) - r^\alpha(A, P^*)\, dt\right). \end{aligned}$$

By concavity, the last term is nonnegative, and is zero for $A = A^*$. Therefore

$$J_\alpha[A^*, P^*, R^*, \tau^*] = w \geq J_\alpha[A, P^*, R^*, \tau^*].$$

$\square$

THEOREM 2.5. *The principal's value function $u = u(w,t)$ solves the HJB equation*

$$\begin{cases} u_t + \sup_{a,p} \left\{ \frac{1}{2}(r_a^\alpha(a,p))^2 u_{ww} - r^\alpha(a,p) u_w + r^\pi(a,p) \right\} = 0 & \text{on } \mathbb{R}_+ \times [0,T) \\ u = 0 & \text{on } \{w = 0\} \times [0,T) \\ u = q & \text{on } \mathbb{R}_+ \times \{t = T\} \end{cases} \tag{2.22}$$

Here $\mathbb{R}_+ := [0, \infty)$.

2.6 Individual rationality constraint for the principal. If his expected payoff w is positive, then the agent has an incentive to accept the contract. But when should the principal agree to this as well? Recall that the terminal value of the HJB equation (2.22) is q, the *negative* of her cost of providing a concluding payout to the agent. Since it is therefore reasonable to assume $q(w) < 0$ for $w > 0$, the principal's value function u is nonpositive at the terminal time T.

Can u be positive earlier in time? This depends on the size of the r^π, the running payoff to the principal, versus the costs of hiring the agent, which depend upon r^α and q. It is generally too hard to solve the terminal value problem (2.22) explicitly, but we can in special cases identify sufficient, but far from optimal, conditions ensuring that the principal's value function is positive somewhere, provided $T > 0$ is large enough:

THEOREM 2.6. *Suppose there exists constants $\hat{a}, \hat{c}$ such that*

$$r^\alpha(\hat{a}, \hat{c}) = 0, \; r_a^\alpha(\hat{a}, \hat{c}) \neq 0, \; r^\pi(\hat{a}, \hat{c}) > 0. \tag{2.23}$$

Assume also

$$q(w) \geq -C(1 + w) \qquad (w \geq 0) \tag{2.24}$$

for some constant $C \geq 0$.

Then for T sufficiently large, we have

$$\max_{\mathbb{R}^+} u(w, 0) > 0. \tag{2.25}$$

To maximize her expected return, the principal should therefore select as the initial condition for the SDE (2.7) a point $w^* \in \mathbb{R}_+$ for which

$$u(w^*, 0) = \sup_{\mathbb{R}_+} u(w, 0).$$

Proof. 1. Write $\sigma := r_a^\alpha(\hat{a},\hat{c})$ and $\rho := r^\pi(\hat{a},\hat{c}) > 0$ and set

$$\Phi(w,t) := \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{w^2}{2t\sigma^2}}.$$

We check that

$$\begin{aligned} v(w,t) := \rho \int_0^t \int_0^\infty & \Phi(w-x,s) - \Phi(w+x,s)\, dxds \\ &+ \int_0^\infty [\Phi(w-x,t) - \Phi(w+x,t)] q(x)\, dx \end{aligned}$$

solves the initial/boundary value problem

$$\begin{cases} v_t = \frac{1}{2}\sigma^2 u_{ww} + \rho & \text{on } \mathbb{R}_+ \times (0,T] \\ v = 0 & \text{on } \{w=0\} \times [0,T) \\ v = q & \text{on } \mathbb{R}_+ \times \{t=0\}. \end{cases} \tag{2.26}$$

Since the principal's value function u solves (2.22), the maximum principle implies

$$u(w,t) \geq v(w, T-t) \qquad (w \geq 0, 0 \leq t \leq T).$$

Then (2.25) will follow, provided

$$\sup_{\mathbb{R}^+} v(w,T) > 0. \tag{2.27}$$

2. To show this, note first that

$$\begin{aligned} v_w(0,T) := \rho \int_0^T \int_0^\infty & \Phi_w(-x,s) - \Phi_w(x,s)\, dxds \\ &+ \int_0^\infty [\Phi_w(-x,T) - \Phi_w(x,T)] q(x)\, dx =: A + B. \end{aligned}$$

Using (2.24), we calculate that

$$A \geq C_1 \rho T^{\frac{1}{2}}, \; B \geq -C_2$$

for appropriate constants $C_1, C_2 > 0$. Consequently $v_w(0,T) > 0$ if T is large enough; and this gives (2.27). $\square$

3 Generalizations

We now generalize, moving to higher dimensions, adding dependence on the spatial variable x and considering more general payoffs and dynamics.

3.1 More complicated dynamics and payoffs. We now allow X to be higher dimensional and, more significantly, assume general drift and the noise terms, that are possibly nonlinear in the agent's effort A and the state X:

$$\begin{cases} dX = f(X,A)dt + \sigma(X,A)dB \qquad (0 \leq t \leq T) \\ X(0) = x. \end{cases} \tag{3.1}$$

Now B is a standard n-dimensional Brownian motion. Here we assume X takes values in $\mathbb{R}^n$, the controls A take values in $\mathbb{R}^m$, the controls P take values in $\mathbb{R}^l$ and

$$f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n,\ \sigma : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{M}^{n\times n},$$

where $\mathbb{M}^{n\times n}$ denotes the space of real, $n \times n$ matrices. We write $\sigma = \sigma(x,a) = ((\sigma^{ij}))$.

The payoffs have the same form as earlier:

$$\begin{cases} J_\pi[A,P,R] = E\left(\int_0^T r^\pi(X,A,P)\,dt + q(X(T),R)\right) \\ J_\alpha[A,P,R] = E\left(\int_0^T r^\alpha(X,A,P)\,dt + R\right), \end{cases} \tag{3.2}$$

where

$$r^\pi, r^\alpha : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^l \to \mathbb{R}.$$

3.2 Concavity hypothesis. Our goal is to design the dynamics W to be used in the compensation scheme for the agent. Given some particular suggested effort A^*, we again need to find a way to ensure that A^* is optimal for the agent. To accomplish this we will need to assume for r^α an appropriate concavity condition that is compatible with the dynamics (3.1).

Definition. Assume $g : \mathbb{R}^m \to \mathbb{R}^n$. We say that a function $\psi : \mathbb{R}^m \to \mathbb{R}$ is **g-concave** if there exists a concave function $\phi : \mathbb{R}^n \to \mathbb{R}$ such that

$$\psi(a) = \phi(g(a)) \qquad (a \in \mathbb{R}^m). \tag{3.3}$$

In other words, ψ is g-concave if it is a concave function of the new variables $\tilde{a} = g(a)$.

We hereafter assume for each x and p that

$$a \mapsto r^\alpha(x,a,p) \text{ is } f(x,a)\text{-concave}. \tag{3.4}$$

We can therefore write

$$r^\alpha(x,a,p) = \phi(x,f(x,a),p) \qquad (x \in \mathbb{R}^n, a \in \mathbb{R}^m)$$

where $s \mapsto \phi(x,s,p)$ is concave, and we assume for simplicity that ϕ is smooth. Then if we define

$$y(x,a,p) := -D_s\phi(x,f(x,a),p), \tag{3.5}$$

the concavity of ϕ implies

$$r^\alpha(x,a,p) - y \cdot (f(x,\hat{a}) - f(x,a)) \geq r^\alpha(x,\hat{a},p) \tag{3.6}$$

for all $x, a, \hat{a}, p$.

3.3 A contract. Given smooth, deterministic functions $a = a(x,w,t)$ and $p = p(x,w,t)$, we solve the SDE

$$\begin{cases} dW = (-r^\alpha(X,A^*,P^*) - f(X,A^*)\cdot Y^*)dt + Y^* \cdot dX \qquad (0 \leq t \leq T) \\ W(0) = w \end{cases} \tag{3.7}$$

for

$$A^* = a(X,W,t),\ P^* = p(X,W,t),\ Y^* := y(X,A^*,P^*), \tag{3.8}$$

the function y defined by (3.5).

The principal can compute $W = W^A$ in terms of the observed values of $X = X^A$; both depend upon the agent's choice of control A. The principal agrees to pay the agent at the rate $P^*(t) = p(X(t),W(t),t)$ for $0 \leq t < T$. In addition she provides the final payout

$$R^* = W(T).$$

We identify next circumstances under which it is optimal for the agent to select $A = A^* = a(X,W,t)$:

THEOREM 3.1. *Suppose for each x and p that*

$$a \mapsto r^\alpha(x,a,p) \text{ is } f(x,a)\text{-concave.}$$

Then it is optimal for the agent to use the control $A^* = a(X,W,t)$.

Proof. Let A be any control for the agent. Then since

$$dX = f(X,A)dt + \sigma(X,A)dB,$$

we have

$$dW = (-r^\alpha(X,A^*,P^*) + Y^* \cdot (f(X,A) - f(X,A^*)))dt + Y^* \cdot \sigma(X,A)dB.$$

Integrate, to find

$$W(T) - w = \int_0^T -r^\alpha(X, A^*, P^*) + Y^* \cdot (f(X,A) - f(X,A^*))\,dt + \int_0^T Y^* \cdot \sigma(X,A)\,dB.$$

Take expected values and rewrite:

$$\begin{aligned} w = E&\left(\int_0^T r^\alpha(X, A, P^*)\,dt + R^*\right) \\ &+ E\left(\int_0^T r^\alpha(X, A^*, P^*) - y(X, A^*, P^*) \cdot (f(X,A) - f(X,A^*)) - r^\alpha(X, A, P^*)\,dt\right). \end{aligned}$$

According to (3.6), the last term is nonnegative, and is zero for $A = A^*$. Therefore

$$\begin{aligned} J_\alpha[A^*, P^*, R^*] &= E\left(\int_0^T r^\alpha(X, A^*, P^*)\,dt + R^*\right) \\ &= w \\ &\geq E\left(\int_0^T r^\alpha(X, A, P^*)\,dt + R^*\right) \\ &= J_\alpha[A, P^*, R^*]. \end{aligned}$$

□

THEOREM 3.2. *The principal's value function $u = u(x, w, t)$ solves the HJB equation*

$$\begin{cases} u_t + \sup_{a,p}\{a^{ij}(x,a)u_{x_ix_j} + y^j(x,a,p)\sigma^{ij}(x,a)u_{x_iw} + \frac{|y(x,a,p)|^2}{2}u_{ww} \\ \qquad +f^i(x,a)u_{x_i} - r^\alpha(x,a,p)u_w + r^\pi(x,a,p)\} = 0 \quad \textit{on } \mathbb{R}^{n+1} \times [0,T) \\ \qquad\qquad u = q \quad \textit{on } \mathbb{R}^{n+1} \times \{t = T\} \end{cases} \tag{3.9}$$

Here $a^{ij} = \frac{1}{2}\sigma^{ik}\sigma^{jk}$ and we recall that $y = y(x, a, p)$ is defined by (3.5). We as before assume that we can select smooth functions $a = a(x, w, t), p = p(x, w, t)$ giving the supremum in the HJB equation.

3.4 How to do better when the noise term depends on A. We consider now explicitly the case that $\sigma = \sigma(x, a)$ truly depends on a; that is, the agent's actions affect the noise term $\sigma(X, A)dB$ in the SDE (3.1). In this case, more information is available to the principal, at least in this highly idealized model, and she can then design a contract that may be better for her than that discussed above.

To see this, note that if $\Phi : \mathbb{R}^n \to \mathbb{R}$, then Itô's formula gives

$$d(\Phi(X)) = \Phi_{x_i} dX^i + \frac{1}{2}\Phi_{x_ix_j}\sigma^{ik}\sigma^{jk}dt.$$

As X, and consequently $\Phi(X)$, are observable to the principal, so therefore is $\sigma^T\sigma(X,A)$. *The principal can therefore enforce the additional contractural requirement of the agent that his effort A satisfy*

$$\sigma^T\sigma(X,A) = \sigma^T\sigma(X,A^*) \qquad (0 \le t \le T), \tag{3.10}$$

where A^* is the suggested control for the agent, as determined below.

To see the advantage to the principal of this information, consider the following new contract. First define

$$\begin{aligned}\Lambda(x,a,p) := \{y \in \mathbb{R}^n \mid r^\alpha(x,a,p) - y\cdot(f(x,\hat{a}) - f(x,a)) \ge r^\alpha(x,\hat{a},p) \\ \text{for all } \hat{a} \in \mathbb{R}^m \text{ such that } \sigma^T\sigma(x,\hat{a}) = \sigma^T\sigma(x,a)\}.\end{aligned} \tag{3.11}$$

Then $\Lambda(x,a,p)$ is a closed, convex (but possibly empty) set.

Consider now the new HJB equation

$$\begin{cases} \tilde{u}_t + \sup_{a,p,y}\{a^{ij}(x,a,p)\tilde{u}_{x_ix_j} + y^j\sigma^{ij}(x,a,p)\tilde{u}_{x_iw} + \frac{|y|^2}{2}\tilde{u}_{ww} + f^i(x,a)\tilde{u}_{x_i} \\ \qquad -r^\alpha(x,a,p)\tilde{u}_w + r^\pi(x,a,p)\} = 0 \quad \text{on } \mathbb{R}^{n+1}\times[0,T) \\ \qquad \tilde{u} = q \quad \text{on } \mathbb{R}^{n+1}\times\{t=T\}, \end{cases} \tag{3.12}$$

where now the variable y over which we take the supremum is subject to the requirement that

$$y \in \Lambda(x,a,p). \tag{3.13}$$

Assuming that we can find smooth enough functions $a = a(x,w,t), p = p(x,w,t)$ and $y = y(x,w,t)$ giving the max in (3.12), the principal can solve the new SDE

$$\begin{cases} dW = (-r^\alpha(X,A^*,P^*) - f(X,A^*)\cdot Y^*)dt + Y^*dX \qquad (0 \le t \le T) \\ W(0) = w, \end{cases} \tag{3.14}$$

now for

$$A^* := a(X,W,t),\ P^* := p(X,W,t),\ Y^* := y(X,W,t). \tag{3.15}$$

She then makes the contract of paying the agent at the rate $P^*(t) = p(X,W,t)$ and also paying him the terminal amount $R^* = W(T)$, provided that she observes that his actions satisfy

$$\sigma^T\sigma(X,A) = \sigma^T\sigma(X,A^*) \qquad (0 \le t \le T) \tag{3.16}$$

where $A^* = a(X,W,t)$. If the agent's actions do not conform to this constraint, he has violated the contract.

THEOREM 3.3. *Assume the principal uses the control P^* and final payoff R^*, as just described.*

(i) *It is then optimal for the agent to use the control $A^* = a(X, W, t)$, and his payoff is w.*

(ii) *Assume for each x and p that $a \mapsto r^\alpha(x, a, p)$ is $f(x, a)$-concave. Suppose that u is the agent's value function for the contract described in Section 3.4 and $\tilde{u}$ is the agent's value function for the contract described above.*

Then

$$\tilde{u} \geq u \qquad \text{on } \mathbb{R}^n \times \mathbb{R} \times [0, T]. \tag{3.17}$$

Thus the principal can use her observations of $\sigma^T \sigma(X, A)$ to come up with a contract that is no more unfavorable, and is perhaps better, to her than that discussed in Section 3.4.

Proof. 1. The proof of the optimality to the agent of the control $A^* = a(X, W, t)$, among all others satisfying (3.16), follows by the usual argument.

2. If $r^\alpha(x, a, p)$ is f-concave in the variable a, the function $y = y(x, p, a)$ defined by (3.5) belongs to the set $\Lambda(x, a, p)$. Hence the supremum in the HJB equation (3.12) for $\tilde{u}$ is over a set that is no smaller than the set over which the sup is computed for the HJB equation (3.9) for u. Therefore the maximum principle implies $\tilde{u} \geq u$. □

4 Many agents

We show next how a single principal can optimally arrange compensation for a collection of N agents, provided they do not collude with each other. For simplicity, we assume that the various agents' actions do not affect the noise in the dynamics.

4.1 Notation, dynamics and payoffs with many agents. The evolution of X is the same as before:

$$\begin{cases} dX = f(X, A)dt + \sigma(X)dB \qquad (0 \leq t \leq T) \\ X(0) = x, \end{cases} \tag{4.1}$$

except that now $A = (A^1, A^2, \ldots, A^N)$, where A^k is the control exercised by the k-th agent for $k = 1, \ldots, N$. So A take values in $\mathbb{R}^{mN}$ and the dynamics (4.1) depend upon all the agents' control choices:

$$f : \mathbb{R}^n \times \mathbb{R}^{mN} \to \mathbb{R}^n, \ \sigma : \mathbb{R}^n \times R^{mN} \to \mathbb{M}^{n \times n}$$

NOTATION. We will write

$$a = (a^1, a^2, \ldots, a^N) \in \mathbb{R}^{mN}, \ p = (p^1, p^2, \ldots, p^N) \in \mathbb{R}^{lN}$$

where $a^k \in \mathbb{R}^m$ and $a^k \in \mathbb{R}^l$ for $k = 1, \dots, N$. Likewise

$$A = (A^1, A^2, \dots, A^N),\ \ P = (P^1, P^2, \dots, P^N),$$

where A^k is the control of the k-th agent and P^k is his running compensation. We also use the convenient notation

$$\begin{aligned} a_{-k} &:= (a^1, \dots, a^{k-1}, a^{k+1}, \dots a^N) \\ (a_{-k}, b^k) &:= (a^1, a^2, \dots, b^k, \dots a^N). \end{aligned}$$

In other words, we obtain (a_{-k}, b^k) by substituting b^k for a^k in $a = (a^1, a^2, \dots, a^N)$.

There are now many payoff functionals, one for the single principal and one for each of the N agents:

$$\begin{cases} J_\pi[A, P, R] = E\left(\int_0^T r^\pi(X, A, P)\, dt + q(X(T), R)\right) \\ J^k_\alpha[A, P, R^k] = E\left(\int_0^T r^\alpha_k(X, A, P)\, dt + R^k\right) \qquad (k = 1, \dots, N) \end{cases}$$

Here R^k is the payoff to the k-th agent at the terminal time T, $R = (R^1, R^2, \dots, R^N)$, and r^π, r^α_k are running costs, where

$$r^\pi, r^\alpha_k : \mathbb{R}^n \times \mathbb{R}^{mN} \times \mathbb{R}^{lN} \to \mathbb{R} \qquad (k = 1, \dots, N).$$

4.2 Concavity hypothesis. We assume for each x, p and a_{-k} that

$$a^k \mapsto r^\alpha_k(x, a, p) \text{ is } f(x, a)\text{-concave} \qquad (k = 1, \dots, N).$$

As in the previous section, it follows that for $k = 1, \dots, N$ there exists a function $y_k(x, a, p)$ such that

$$r^\alpha_k(x, a, p) \le r^\alpha_k(x, a^*, c) - y_k(x, a^*, c) \cdot (f(x, a) - f(x, a^*)) \tag{4.2}$$

for all x, a^*, c and

$$a = (a^*_{-k}, a^k).$$

4.3 The contracts. Given functions $a = a(x, w, t) = (a_1(x, w, t), \dots, a_N(x, w, t))$ and $p = p(x, w, t) = (c_1(x, w, t), \dots, c_N(x, w, t))$, we solve the system of SDE

$$\begin{cases} dW_k = (-r^\alpha_k(X, A^*, P^*) - f(X, A^*)Y^*_k)dt + Y^*_k dX \qquad (0 \le t \le T) \\ W_k(0) = w_k. \end{cases}$$

We use the notation $W = (W_1, \dots, W_N)$.

The principal agrees to pay the k-th agent at the rate $P^*_k = p_k(X, W, t)$ for times $0 \le t < T$, depending upon the values of W, which she can compute in terms of the observed values of X. In addition the principal provides the final payout

$$R^{*k} = W_k(T).$$

THEOREM 4.1 (Feedback Nash equilibrium for the agents)**.** *Under the above assumptions, it is optimal for the k-th agent to use the control*

$$A^{*k} = a_k(X, W, t),$$

assuming that each of the other agents j uses the corresponding control A^{*j} for $j = 1, \ldots, k-1, k+1, \ldots N$.

In other words, $A^* = (A^{*1}, \ldots, A^{*N})$ is a Nash equilibrium for the agents.

Proof. Let A^k be any control for agent k. Then

$$\begin{aligned} dW_k &= (-r_k^\alpha(X, A^*, P^*) - f(X, A^*)Y_k^*)dt + Y_k^* dX \\ &= (-r_k^\alpha(X, A^*, P^*) + Y_k^* \cdot (f(X, A_k) - f(X, A^*)))dt + Y_k^* \cdot \sigma dB, \end{aligned}$$

where $A_k := (A_{-k}^*, A^k)$. Then

$$W_k(T) - w^k = \int_0^T -r_k^\alpha(X, A^*, P^*) + Y_k^* \cdot (f(X, A_k) - f(X, A^*))\, dt + \int_0^T Y_k^* \cdot \sigma dB,$$

and thus

$$\begin{aligned} w^k = E&\left(\int_0^T r_k^\alpha(X, A, P^*)\, dt + R^{*k}\right) \\ &+ E\left(\int_0^T r_k^\alpha(X, A^*, P^*) - r_k^\alpha(X, A_k, P^*) + y_k(X, A^*, P^*) \cdot (f(X, A_k) - f(X, A^*))\, dt\right). \end{aligned}$$

According to (4.2), the last term is nonnegative, and is zero for $A^k = A^{*k}$. Therefore

$$\begin{aligned} J_\alpha^k[A_{-k}^*, A^{*k}, P^*, R^{*k}] &= J_\alpha^k[A^*, P^*, R^{*k}] \\ &= E\left(\int_0^T r_k^\alpha(X, A^*, P^*)\, dt + R^{*k}\right) \\ &= w^k \\ &\geq E\left(\int_0^T r_k^\alpha(X, A_k, P^*)\, dt + R^{*k}\right) \\ &= J_\alpha^k[A_k, P^*, R^{*k}] \\ &= J_\alpha^k[A_{-k}^*, A^k, P^*, R^{*k}]. \end{aligned}$$

Therefore, so long as the other agents stick with the controls

$$A_{-k}^* = (A^{*1}, \ldots, A^{*(k-1)}, A^{*(k+1)}, \ldots, A^{*N}),$$

it is optimal for the k-th agent to employ A^{*k}. $\square$

Write

$$r^\alpha(x,a,p) = (r_1^\alpha(x,a,p),\dots,r_N^\alpha(x,a,p)).$$

THEOREM 4.2. *The principal's value function $u = u(x,w,t)$ solves the HJB equation*

$$\begin{cases} u_t + \sup_{a,p}\{a^{ij}u_{x_ix_j} + y_k^i(x,a,p)\sigma^{ij}(x)u_{x_jw_k} + \frac{1}{2}y_l^i(x,a,p)y_k^i(x,a,p)u_{w_kw_l} \\ \qquad\qquad +f^i(x,a)u_{x_i} - r_k^\alpha(x,a,p)u_{w_k} + r^\pi(x,a,p)\} = 0 \quad \text{on } \mathbb{R}^{n+N}\times[0,T) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad u = q \quad \text{on } \mathbb{R}^{n+N}\times\{t=T\}. \end{cases}$$

5 Appendix 1: The agent's value function and his HJB equation

We have several times above studied the principal's value function and her HJB equation, and it therefore seems natural to discuss as well the agent's value function and his HJB equation. This analysis provides an alternative (but somewhat more complicated) derivation of the agent's optimality condition, Theorem 2.1.

We start by generalizing slightly and assuming the initial time is some given $t\in[0,T]$, and not necessarily $t=0$ as in the previous sections. The dynamics for W are

$$\begin{cases} dW = (-r^\alpha(A^*,P^*) - Y^*A^*)ds + Y^*dX \qquad (t\le s\le T) \\ W(t) = w. \end{cases} \tag{5.1}$$

Below we write $a^* = a(w,t)$ and $p^* = p(w,t)$ to denote functions giving the supremum in the principal's HJB (2.14).

THEOREM 5.1. *If $a\mapsto r^\alpha$ is concave, the agent's value function is*

$$v(w,t)\equiv w. \tag{5.2}$$

Proof. The agent's value function $v = v(w,t)$ solves the agent's HJB equation

$$\begin{cases} v_t + \sup_a\{\frac{1}{2}(r_a^\alpha(a^*,p^*))^2v_{ww} - (a-a^*)r_a^\alpha(a^*,p^*)v_w \\ \qquad\qquad -r^\alpha(a^*,p^*)v_w + r^\alpha(a,p^*)\} = 0 \quad \text{on } \mathbb{R}\times[0,T) \\ \qquad\qquad\qquad\qquad\qquad v = w \quad \text{on } \mathbb{R}\times\{t=T\} \end{cases}$$

We claim that when $a\mapsto r^\alpha$ is concave, a solution is given by (5.2) and that the supremum occurs at $a=a^*$. To confirm this, let $v\equiv w$ above, and check that

$$\begin{aligned} v_t + \sup_a\{&\frac{1}{2}(r_a^\alpha(a^*,p^*))^2v_{ww} - (a-a^*)r_a^\alpha(a^*,p^*)v_w - r^\alpha(a^*,p^*)v_w + r^\alpha(a,p^*)\} \\ &= \sup_a\{r^\alpha(a,p^*) - (r^\alpha(a^*,p^*) + r_a^\alpha(a^*,p^*)(a-a^*))\} \\ &= 0. \end{aligned}$$

And $v(w,t) \equiv w$ satisfies the correct terminal condition on $\mathbb{R} \times \{t = T\}$. By uniqueness, therefore, the agent's value function equals w. $\square$

6 Appendix 2: Better contracts?

We address in this section the subtle question of the actual optimality of the contract devised in Section 2, and, in particular, reproduce and comment upon Sannikov's formulation [S].

To begin, let us suppose that the triple $(\hat{A}, \hat{P}, \hat{R})$ is given, the process $\hat{A}$ being $\mathcal{F}$-adapted. We regard $(\hat{A}, \hat{P}, \hat{R})$ as a proposed contract offered by the principal, where $\hat{A}$ is the suggested effort for the agent. We define then the process

$$\begin{cases} dX = \hat{A}dt + dB \qquad (0 \le t \le T) \\ X(0) = x, \end{cases} \tag{6.1}$$

and assume as well that

$$\hat{P}, \hat{R} \text{ are } \mathcal{F}^X\text{-adapted.} \tag{6.2}$$

Assume instead that the agent's effort is really the process A, although he can claim that in fact $dX = \hat{A}dt + dB$. We then understand the evolution of the observed process X as

$$\begin{cases} dX = \hat{A}dt + dB = Adt + dB^A \qquad (0 \le t \le T) \\ X(0) = x, \end{cases} \tag{6.3}$$

for

$$B^A(t) := \int_0^t \hat{A} - A\, ds + B(t) \qquad (0 \le t \le T).$$

We can invoke Girsanov's Theorem to secure a new probability measure $P^A = P^{A,\hat{A}}$ under which B^A is a Brownian motion. Let us then *redefine* the payoffs (2.2), now to read

$$\begin{cases} J_\pi[A, \hat{P}, \hat{R}] := E^A\Big(\int_0^T r^\pi(A, \hat{P})\, dt + q(\hat{R})\Big) \\ J_\alpha[A, \hat{P}, \hat{R}] := E^A\Big(\int_0^T r^\alpha(A, \hat{P})\, dt + \hat{R}\Big), \end{cases} \tag{6.4}$$

where E^A denotes expectation with respect to P^A. Note carefully that these payoffs therefore also depend upon $\hat{A}$.

We assume further that $\hat{A}$ is incentive compatible for the agent, in the sense that

$$J_\alpha[\hat{A}, \hat{P}, \hat{R}] = \max_A J_\alpha[A, \hat{P}, \hat{R}], \tag{6.5}$$

where A is $\mathcal{F}$-adapted. We call the triple $(\hat{A}, \hat{P}, \hat{R})$ *admissible* if (6.2) and (6.5) hold.

Finally, let (A^*, P^*, R^*) be the contract designed in Section 2. The following theorem describes the sense in which this proposed contract is optimal for the principal:

THEOREM 6.1. (i) *Assume that*

$$a \mapsto r^\alpha(a,p) \text{ is concave for each } p.$$

Then we have the agent's optimality condition

$$J_\alpha[A^*, P^*, R^*] = \max_A J_\alpha[A, P^*, R^*], \tag{6.6}$$

for the payoff J_α defined by (6.4). *Consequently, (A^*, P^*, R^*) is admissible.*

(ii) *If furthermore $(\hat{A}, \hat{P}, \hat{R})$ is any other admissible contract, then*

$$J_\pi[\hat{A}, \hat{P}, \hat{R}] \le J_\pi[A^*, P^*, R^*], \tag{6.7}$$

for J_π defined by (6.4).

Proof. 1. The proof of (6.6) is almost exactly like the proof of Theorem 2.1. Let $P^A = P^{A,A^*}$ now denote the measure for which

$$B^A(t) := \int_0^t A^* - A\,ds + B(t) \qquad (0 \le t \le T).$$

is a Brownian motion. Define the process W by (2.7). We compute

$$W(T) - w = \int_0^T -r^\alpha(A^*, P^*) + Y^*(A - A^*)\,dt + \int_0^T Y^*\,dB^A,$$

where $Y^* = -r^\alpha_a(A^*, P^*)$. Take expected values, but now with respect to the new measure P^A:

$$\begin{aligned} w = E^A &\left(\int_0^T r^\alpha(A, P^*)\,dt + R^* \right) \\ &+ E^A \left(\int_0^T r^\alpha(A^*, P^*) + r^\alpha_a(A^*, P^*)(A - A^*) - r^\alpha(A, P^*)\,dt \right). \end{aligned}$$

By concavity, it follows as before that $w = J_\alpha[A^*, P^*, R^*] \ge J_\alpha[A, P^*, R^*]$.

2 . To prove (6.7), take $(\hat{A}, \hat{P}, \hat{R})$ to be any admissible contract. We follow Sannikov and define

$$\hat{W}(t) := E\left(\int_t^T r^\alpha(\hat{A}, \hat{P})\,ds + \hat{R} \mid \mathcal{F}(t) \right) \quad (0 \le t \le T).$$

In particular, $\hat{W}(T) = \hat{R}$. Then

$$\hat{W}(t) + \int_0^t r^\alpha(\hat{A}, \hat{P})\, ds = E\left(\int_0^T r^\alpha(\hat{A}, \hat{P})\, ds + \hat{R} \mid \mathcal{F}(t)\right),$$

the right hand side of which is an $\mathcal{F}$ martingale. According therefore to the Martingale Representation Theorem (see Karatzas–Shreve [K-S]), there exists an $\mathcal{F}$-adapted process $\hat{Y}$ such that

$$\hat{W}(t) + \int_0^t r^\alpha(\hat{A}, \hat{P})\, ds = \hat{W}(0) + \int_0^t \hat{Y}\, dB \quad (0 \le t \le T). \tag{6.8}$$

3. Put

$$\hat{w} := \hat{W}(0) = E\left(\int_0^T r^\alpha(\hat{A}, \hat{P})\, ds + \hat{R}\right) = J_\alpha[\hat{A}, \hat{P}, \hat{R}];$$

note that $\hat{w}$ is a number, since $\mathcal{F}(0)$ is the trivial σ-algebra.

4. We now claim that

$$\hat{Y} = -r_a^\alpha(\hat{A}, \hat{P}) \quad \text{almost surely.} \tag{6.9}$$

To see this, let A be any other agent effort and observe from (6.8) that

$$d\hat{W} = -r^\alpha(\hat{A}, \hat{P})dt + \hat{Y}\, dB = (-r^\alpha(\hat{A}, \hat{P}) + \hat{Y}(A - \hat{A}))dt + \hat{Y}\, dB^A.$$

We integrate and take expected values with respect to $P^A = P^{A,\hat{A}}$, to find

$$\hat{w} = \hat{W}(0) = \mathbb{E}^A\left(\int_0^T r^\alpha(A, \hat{P})dt + \hat{R}\right) + \mathbb{E}^A\left(\int_0^T r^\alpha(\hat{A}, \hat{P}) - r^\alpha(A, \hat{P}) + \hat{Y}(\hat{A} - A)\, dt\right).$$

Now if (6.9) were false, there would exist A such that the second term on the right is negative. It would then follow that

$$J_\alpha[A, \hat{P}, \hat{R}] > \hat{w} = J_\alpha[\hat{A}, \hat{P}, \hat{R}], \tag{6.10}$$

a contradiction to the admissibility condition (6.5).

4. In view of the foregoing, we have

$$\begin{cases} d\hat{W} = -(r^\alpha(\hat{A}, \hat{P}) + \hat{Y}\hat{A})dt + \hat{Y}\, dX & (0 \le t \le T) \\ \hat{W}(0) = \hat{w}, \end{cases} \tag{6.11}$$

where $\hat{Y} = -r_a^\alpha(\hat{A}, \hat{P})$ and $\hat{W}(T) = \hat{R}$. These are the same as dynamics (2.7) for the process W, except for the initial condition. Consequently,

$$J_\pi[A^*, P^*, R^*] = u(w^*, 0) \ge u(\hat{w}, 0) \ge J_\pi[\hat{A}, \hat{P}, \hat{R}].$$

where the principal's value function u solves the principal's HJB equation (2.14) and w^* verifies (2.15).

□

References

[C-P-T] J. Cvitanic, D. Possamai and N. Touzi, Dynamic programming approach to principal-agent problems, arXiv:1510.07111

[C-Z] J. Cvitanic and J. Zhang, *Contract Theory in Continuous Time Models*, Springer, 2013.

[E] I. Ekeland, How to build stable relationships with people who lie and cheat, Milan Journal of Mathematics 82 (2014), 67–79.

[K-S] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, Graduate Texts in Mathematics 113, Springer, 1988

[M-Y] C. Miller and I. Yang, Optimal dynamic contracts for a large-scale principal-agent hierarchy: a concavity-preserving approach, arXiv:1506.05497 [math.OC] (2015).

[S] Y. Sannikov, A continuous time version of the principal-agent problem, Review of Economic Studies 75 (2008), 957–984.