

RISK SENSITIVE STOCHASTIC CONTROL AND DIFFERENTIAL GAMES

WENDELL H. FLEMING*

Abstract. We give a concise introduction to risk sensitive control of Markov diffusion processes and related two-controller, zero-sum differential games. The method of dynamic programming for the risk sensitive control problem leads to a nonlinear partial differential equation of Hamilton-Jacobi-Bellman type. In the totally risk sensitive limit, this becomes the Isaacs equation for the differential game. There is another interpretation of the differential game using the Maslov idempotent probability calculus. We call this a max-plus stochastic control problem. These risk sensitive control/differential game methods are applied to problems of importance sampling for Markov diffusions.

1. Introduction. The purpose of this paper is to give a concise introduction to risk sensitive control of Markov diffusion processes and related two controller, zero sum differential games. For simplicity, we consider only risk sensitive control on a fixed finite time horizon. Our formulation is the same as in the book [FS, Chaps. 6 and 11], to which the reader is referred for a more detailed treatment. References [FM2] and [N] are concerned with risk sensitive control on an infinite time horizon.

To motivate the discussion to follow, we begin in Section 2 with an elementary discussion of static optimization. In the risk sensitive stochastic optimization formulation, one seeks to minimize the certainty-equivalent expectation of some criterion \mathcal{J} . This depends on a risk sensitivity parameter θ . In the limit $\theta \rightarrow \infty$, the risk sensitive stochastic optimization problem becomes a deterministic min-max problem. This limiting problem has a kind of “stochastic” interpretation in terms of the Maslov idempotent probability calculus.

In Section 3 we review the dynamic programming approach to risk sensitive control of Markov diffusions on a finite time interval. The state of the process being controlled satisfies the stochastic differential equation (3.1), depending on a parameter ϵ which indicates the intensity of random noise inputs to the state dynamics. The risk sensitivity parameter is $\theta = \epsilon^{-2}$. Dynamic programming leads to the Hamilton-Jacobi-Bellman (HJB) partial differential equation. When written in certainty equivalent form, this is equation (3.5). Under suitable assumptions a unique solution to (3.5) with the boundary data (3.7) exists in either the classical or the viscosity solution sense.

In Section 4 we consider the totally risk sensitive limit $\epsilon \rightarrow 0$. This leads to

*Brown University, Division of Applied Mathematics, Box F / 182 George Street, Providence, RI 02912. E-mail: whf@cfm.brown.edu

a differential game, with state dynamics (4.1). In these game dynamics, the random (Brownian motion) input to the SDE (3.1) is replaced by a disturbance control function chosen by a maximizing controller. This is an approach taken in nonlinear H_∞ -control theory [BB][HJ]. As $\epsilon \rightarrow 0$ the certainty-equivalent value function V^ϵ in the risk sensitive stochastic control problem tends to the upper value function V^+ for the differential game. This upper value function satisfies an Isaacs partial differential equation in the viscosity sense.

In Section 5, another interpretation of the differential game in Section 4 is given, in terms of the Maslov idempotent probability calculus. We call this a “max-plus stochastic control problem.” We discuss the problem of finding suitable classes of minimizing control strategies for the max-plus formulation. These strategies should have a role similar to progressively measurable controls in the usual stochastic model for controlled Markov diffusion. One such class, called Γ_1 in Section 5, consists of strategies which are “strictly progressive.” Since the definition (5.4) of strictly progressive is not very intuitive, it would be interesting to describe other choices for Γ_1 .

Importance sampling is a technique which involves changes of probability measure to reduce the variance in Monte Carlo estimation. In Sections 6 and 7 we outline an approach to importance sampling due to Dupuis and Wang, as it applies to Markov diffusion processes. In this case, the change of probability measure corresponds to a change of drift in the stochastic differential equation which governs the diffusion process dynamics. In principle, the variance can be reduced to 0 (see Theorem 6.1). However, this result is generally not helpful in applications, since it depends on solving the backward PDE (6.6) for the expectation of the random variable being estimated by Monte Carlo. Instead, the importance sampling problem is reformulated as a risk sensitive stochastic control problem of the type in Section 3. In Section 7 we outline the Dupuis-Wang method for obtaining nearly optimal changes of probability measure using approximate supersolutions to the Isaacs PDE for this differential game. Detailed descriptions of this method and applications of it appear in [DW1][DW2][DSW].

The author wishes to thank Hui Wang for helpful suggestions.

2. Static optimization. Let us begin by contrasting, at an abstract level, stochastic and deterministic (min-max) approaches to optimization. To simplify the presentation, we consider in this section “static” problems, in which time plays no role. In the abstract formulation, u denotes a control, v an uncertainty (or disturbance) and $\mathcal{J} = \mathcal{J}(u, v)$ some criterion.

In a stochastic optimization model, v is “random.” Thus $v \in \Omega$, where (Ω, \mathcal{F}, P) is some probability space.

The traditional stochastic optimization objective is to choose u which minimizes

the expectation

$$(2.1) \quad J = E(\mathcal{J}) = \int_{\Omega} \mathcal{J}(u, v)P(dv).$$

In contrast, the traditional min-max approach is to choose u to minimize $\max_v \mathcal{J}(u, v)$. (In this section, the discussion is heuristic. Thus, we omit assumptions needed to ensure integrability, existence of maxima and minima, etc.)

Modified min-approach. The traditional min-max approach is ultraconservative. In taking this approach, the minimizing controller is guarding against all possible uncertainties v which are considered “equally likely.” A less conservative approach is to introduce a function $q(v)$ which measures the “likelihood” of v . The modified min-max problem is to choose u which minimizes the max over v of $\mathcal{J}(u, v) + q(v)$. Let

$$(2.2) \quad V = \min_u \max_v [\mathcal{J}(u, v) + q(v)].$$

If the min occurs at $u = u^*$ and the max at $v = v^*(u)$, then the following inequalities express the *saddle point property*:

$$(2.3) \quad V \geq \mathcal{J}(u^*, v) + q(v) \quad \forall v$$

$$(2.4) \quad V \leq \mathcal{J}(u, v^*(u)) + q(v^*(u)) \quad \forall u.$$

Risk-sensitive minimization. A link between stochastic and deterministic viewpoints is provided by considering risk-sensitive stochastic optimization. Let $F(z)$ be a smooth function, such that

$$(2.5) \quad F'(z) > 0, \quad |F''(z)| \neq 0.$$

The risk-sensitive stochastic optimization problem is to choose u which minimizes the expectation $E[F(\mathcal{J})]$, rather than the traditional criterion $E(\mathcal{J})$.

The *certainty-equivalent expectation* $\mathcal{E}(\mathcal{J})$ is defined by

$$(2.6) \quad \mathcal{E}(\mathcal{J}) = F^{-1}(E[F(\mathcal{J})]).$$

In this paper, we consider only exponential functions F :

$$(2.7) \quad F(z) = e^{\theta z}, \quad \theta > 0.$$

The coefficient of risk sensitivity is defined as

$$(2.8) \quad r_F(z) = \frac{|F''(z)|}{F'(z)}.$$

For the traditional stochastic criterion, $F(z) = z$. This is the “risk neutral” case. Large $r_F(z)$ indicates a great sensitivity to risk. For the exponential, $r_F(z) = \theta$ is constant and hence large θ corresponds to great risk sensitivity. When F is exponential, then (2.6) and (2.7) give

$$(2.9) \quad \mathcal{E}(\mathcal{J}) = \theta^{-1} \log E(e^{\theta \mathcal{J}}).$$

For simplicity, let us assume in this introductory section that u is chosen from a finite set and that v also belongs to some finite set. Then u minimizes $E[F(\mathcal{J})]$ if and only if u minimizes

$$(2.10) \quad \mathcal{E}(\mathcal{J}) = \theta^{-1} \log \sum_v e^{\theta \mathcal{J}(u,v)} p_\theta(v)$$

with $p_\theta(v)$ the probability of v . Suppose that the dependence on θ is such that

$$(2.11) \quad \lim_{\theta \rightarrow \infty} \theta^{-1} \log p_\theta(v) = q(v)$$

where $q(v) \leq 0$ and $q(v_0) = 0$ for some “most likely” v_0 . A version of the Laplace-Varadhan lemma implies that

$$(2.12) \quad \lim_{\theta \rightarrow \infty} \min_u \mathcal{E}(\mathcal{J}) = V$$

with V as in (2.2). See for example [FHH, Lemma A.1]. Thus, the modified min-max problem arises as a limit of the risk sensitive stochastic problem, as the risk parameter $\theta \rightarrow \infty$.

This elementary discussion is intended to motivate the discussion of risk sensitive control of Markov diffusions, and differential game limits in Sections 3 and 4. In the stochastic model in Section 3, the uncertainty enters via a Brownian motion in the stochastic differential equation for the state process. In Section 4, the uncertainty is modelled as a deterministic “disturbance function” v in the state dynamics.

The modified min-max approach is non-stochastic in any traditional sense. However, it does have a kind of “stochastic” interpretation in terms of the Maslov idempotent probability calculus [MS]. Let \oplus and \otimes denote max-plus addition and multiplication:

$$a \oplus b = \max(a, b) \quad a \otimes b = a + b.$$

The *max-plus expectation* $E^+(\mathcal{J})$ is defined as

$$(2.13) \quad \begin{aligned} E^+(\mathcal{J}) &= \max_v [\mathcal{J}(u, v) + q(v)] \\ &= \oplus_v [q(v) \otimes \mathcal{J}(u, v)]. \end{aligned}$$

If $q(v)$ is regarded as the “likelihood” (or “max-plus probability”) of v , then $E^+(\mathcal{J})$ has exactly the same form as the usual expectation $E(\mathcal{J})$ but with usual arithmetic

and probabilities replaced by their max-plus counterparts. With these notations, we can rewrite V in (2.2) as the minimum of max-plus expectations:

$$(2.14) \quad V = \min_u E^+(\mathcal{J}).$$

We will return to this idea in Section 5.

3. Controlled Markov diffusions. In this section we outline very briefly some ideas and results about risk sensitive control of Markov diffusions. For detailed treatments see [FS, Sec. 6.8][BN][J][N][FM1,2] and references cited there.

Notations. \mathbb{R}^n denotes n -dimensional Euclidean space, and $x = (x_1, \dots, x_n)$ denotes a point in \mathbb{R}^n . We consider control on a finite interval $t \leq s \leq T$, with $t \geq 0$. Let $Q_0 = (0, T) \times \mathbb{R}^n$ and \bar{Q}_0 the closure of Q_0 . Let $C^k(\mathbb{R}^n)$ denote the space of k times continuously differentiable functions on \mathbb{R}^n , with values in some Euclidean space, and $C_b^k(\mathbb{R}^n)$ the subspace of those $f \in C^k(\mathbb{R}^n)$ such that f and its partial derivatives of orders up to k are bounded. Similarly, $C^{1,2}(Q)$ is the space of functions $\phi(t, x)$ such that the partial derivatives $\partial\phi/\partial t, \phi_{x_i}, \phi_{x_i x_j}, i, j = 1, \dots, n$ are continuous and $C_b^{1,2}(Q)$ is the subspace of those $\phi \in C^{1,2}(Q)$ such that ϕ and these partial derivatives are bounded. The gradient of $\phi(t, \cdot)$ is denoted by $D_x\phi$ and $D_x^2\phi$ is the matrix of partial derivatives $\phi_{x_i x_j}, i, j = 1, \dots, n$.

The risk sensitive stochastic control problem is formulated as follows: Let $x_s^\epsilon \in \mathbb{R}^n$ denote the state and $u_s \in U$ the control at time s . The state process satisfies the Ito stochastic differential equation (SDE)

$$(3.1) \quad dx_s^\epsilon = f(s, x_s^\epsilon, u_s)ds + \epsilon\sigma(s, x_s^\epsilon, u_s)dw_s, \quad t \leq s \leq T,$$

with initial data $x_t^\epsilon = x$. In (3.1), $\epsilon > 0$ is a parameter and w is a \mathcal{F}_s -adapted Brownian motion of dimension m on some probability space $(\Omega, \mathcal{F}_T, P)$ with $\{\mathcal{F}_s\}$ an increasing family of σ -algebras. Let

$$(3.2) \quad \mathcal{J}^\epsilon = \int_t^T L(s, x_s^\epsilon, u_s)ds + G(x_T^\epsilon)$$

and in the notation of Section 2 let $\theta = \epsilon^{-2}$. The risk sensitive control problem is to choose a $\{\mathcal{F}_s\}$ -progressively measurable control process u to minimize $E_{tx}[\exp(\epsilon^{-2}\mathcal{J}^\epsilon)]$, where \exp denotes the exponential function and the subscript tx refers to the initial data $x_t^\epsilon = x$.

As in [FS, Sec. 6.8] we make the following assumptions:

- (a) $U \subset \mathbb{R}^\ell$ and U is compact;
- (b) $f \in C^1(\overline{Q}_0 \times U)$ and f has bounded first order partial derivatives;
- (3.3) (c) $\sigma, L \in C_b^1(\overline{Q}_0 \times U)$;
- (d) $G \in C_b^1(\mathbb{R}^n)$.

The risk sensitive control problem is to find u , progressively measurable which minimizes $E_{tx}[\exp(\epsilon^{-2} \mathcal{J}^\epsilon)]$. Let $\phi^\epsilon(t, x)$ denote the value function. The HJB equation for ϕ^ϵ is the nonlinear PDE [FS, eqn. 6(8.5)]. It is convenient to rewrite this PDE in the following certainty-equivalent form. Let $V^\epsilon = \epsilon^2 \log \phi^\epsilon$. Thus $V^\epsilon(t, x)$ is the inf over u , of the certainty-equivalent expectation

$$(3.4) \quad \mathcal{E}_{tx}(\mathcal{J}^\epsilon) = \epsilon^2 \log E_{tx}[\exp(\epsilon^{-2} \mathcal{J}^\epsilon)].$$

It turns out that V^ϵ does not depend on the particular reference probability system $(\Omega, \{\mathcal{F}_s\}, P, w)$ in the problem formulation above. See [FS, Sec. 4.2 and 4.7]. In certainty-equivalent form, the associated HJB equation is (see [FS, eqn. 6(8.12)]):

$$(3.5) \quad \frac{\partial V^\epsilon}{\partial t} + H_\epsilon(t, x, D_x V^\epsilon, D_x^2 V^\epsilon) = 0,$$

$$(3.6) \quad H_\epsilon(t, x, p, \pi) = \min_{u \in U} \left[f(t, x, u) \cdot p + \frac{\epsilon^2}{2} \text{tr} a(t, x, u) \pi + \frac{1}{2} a(t, x, u) p \cdot p + L(t, x, u) \right]$$

with $a = \sigma \sigma'$, $\pi = (\pi_{ij})$ any symmetric $n \times n$ matrix and tr the trace. The certainty-equivalent value function V^ϵ is the unique bounded, Lipschitz continuous viscosity solution to (3.4) in Q_0 with the boundary data.

$$(3.7) \quad V^\epsilon(T, x) = G(x).$$

See [FS, Secs. 5.9 and 6.8]. Under stronger assumptions, including uniform parabolicity of the HJB equation (3.4), V^ϵ satisfies (3.5) in the classical sense. See [FS, Remark 6.8.1]. Moreover, an optimal Markov control policy can be obtained by taking arg min in (3.6) with $p = D_x V$ and $\pi = D_x^2 V^\epsilon$. See [FS, Section 4.4].

4. Differential game limit. Instead of the risk-sensitive stochastic control model in Section 3, let us consider the kind of (modified) min-max approach outlined in Section 2. This approach leads to the following two-player, zero-sum differential

game. Let $x_s \in \mathbb{R}^n$ denote the game state at time s , which evolves according to the differential equation

$$(4.1) \quad \frac{dx_s}{ds} = f(s, x_s, u_s) + \sigma(s, x_s, u_s)v_s, \quad t \leq s \leq T$$

with initial data $x_t = x$. In (4.1) the minimizing controller (or player) chooses $u_s \in U$ and the maximizing controller chooses $v_s \in \mathbb{R}^m$. Note that in (4.1), the formal derivative $\epsilon dw_s/ds$ in the stochastic differential equation (3.1) has been replaced by deterministic v_s . In the terminology of H_∞ -control theory, v_s is often called a “disturbance control.” See [BB][HJ]. Let

$$(4.2) \quad \mathcal{J}(t, x; u, v) = \int_t^T L(s, x_s, u_s) ds + G(x_T),$$

$$(4.3) \quad q(v) = -\frac{1}{2} \int_t^T |v_s|^2 ds.$$

This choice of $q(v)$ is motivated by the theory of large deviations for Brownian motions [FW]. The differential game payoff is

$$(4.4) \quad P(t, x; u, v) = \mathcal{J}(t, x; u, v) + q(v).$$

The game description is not yet complete, since the information available to each controller must be specified. Speaking intuitively, both players know the state x_s when u_s and v_s are chosen. Moreover, the maximizing controller has the advantage of knowing u_s when the disturbance control v_s is chosen. This is the so-called “upper” differential game.

There are well known difficulties in making this intuitive formulation precise. To avoid these difficulties, various rigorous definitions of upper differential game have been made (using time discretizations, Elliott-Kalton strategies, etc.). Any “reasonable” definition provides an upper game value $V^+(t, x)$, which satisfies in the viscosity sense the Isaacs PDE

$$(4.5) \quad \frac{\partial V^+}{\partial t} + H(t, x, D_x V^+) = 0, \quad (t, x) \in Q_0,$$

$$(4.6) \quad V^+(T, x) = G(x),$$

$$(4.7) \quad H(t, x, p) = \min_{u \in U} \max_{v \in \mathbb{R}^m} \left[(f(t, x, u) + \sigma(t, x, u)v) \cdot p + L(t, x, u) - \frac{1}{2}|v|^2 \right].$$

We make the same assumptions (3.3) as in Section 3. Then the Elliott-Kalton value function V^+ is the unique bounded Lipschitz continuous viscosity solution to (4.5) - (4.6). See [FS, Sec. 7.11].

By taking the max over v , we rewrite (4.7) as

$$(4.7') \quad H(t, x, p) = \min_{u \in U} \left[(f(t, x, u) \cdot p + \frac{1}{2} a(t, x, u) p \cdot p + L(t, x, u)) \right].$$

with $a = \sigma \sigma'$. Thus $H(t, x, p)$ is the limit as $\epsilon \rightarrow 0$ of $H_\epsilon(t, x, p, \pi)$ in (3.6). Viscosity solution arguments can be used to show that the certainty-equivalent value function tends to the upper differential game value function as $\epsilon \rightarrow 0$:

$$(4.8) \quad V^+(t, x) = \lim_{\epsilon \rightarrow 0} V^\epsilon(t, x),$$

uniformly on compact subsets of \overline{Q}_0 . See [FS, Sec. 11.7].

In addition to the upper value function V^+ , there is a lower value function V^- which satisfies the PDE corresponding to (4.7) with min max replaced by max min. It is important to note that V^+ is obtained in the small-noise ($\epsilon \rightarrow 0$) limit (4.8), and not the lower differential game value V^- . However, if $\sigma = \sigma(t, x)$ does not depend on the control variable u , then the Isaacs minimax condition $\min \max = \max \min$ holds in (4.7). In that case, $V^+ = V^-$.

In the standard treatment of differential games on a finite interval, the controls u_s, v_s are constrained to lie in compact sets. See for example [FS, Chap. 11]. We have assumed that $u_s \in U$, with U compact. However, $v_s \in \mathbb{R}^m$ and \mathbb{R}^m is not compact. It is shown in [FS, Sec. 11.7] that the upper value V^+ is the same if the constraint $|v| \leq R_1$ is imposed, if R_1 is chosen large enough. In fact, one can take $R_1 = K$ where K is a bound for $|\sigma'(t, x, u)p|$ when $|p| \leq M_1$ and M_1 is a Lipschitz constant for $V^+(t, \cdot)$.

5. Max-plus stochastic control. In this section we give another interpretation of the differential game in Section 4, in terms of the Maslov idempotent probability calculus already mentioned in Section 2. With this interpretation, the upper game value function V^+ becomes the value function for what we call a “max-plus stochastic control problem”.

Let us regard $q(v)$ defined by (4.3) as the max-plus likelihood of a disturbance $v \in L^2([t, T]; \mathbb{R}^m)$. As in (2.13), the max-plus expectation for initial data $x_t = x$ and \mathcal{J} as in (4.2) becomes

$$(5.1) \quad E_{tx}^+(\mathcal{J}) = \sup_v P(t, x; u, v).$$

In the max-plus stochastic control problem, we need an analogue of the progressively measurable control processes used to define the certainty-equivalent value function V^ϵ in the risk sensitive stochastic control problem. This analogue will be a suitably defined class of Elliott-Kalton strategies for the minimizing controller in the differential game.

An Elliott-Kalton strategy α for the minimizing controller is a function from $L^2([t, T]; \mathbb{R}^m)$ into $L^2([t, T]; U)$ which is *progressive* in the following sense: for $t < s < T$, $v_r = \tilde{v}_r$ for almost $r \in [t, s]$ implies $\alpha(v)_r = \alpha(\tilde{v})_r$ for almost all $r \in [t, s]$. Let $\Gamma = \Gamma(t, T)$ denote the class of all Elliott-Kalton strategies α . By the Elliott-Kalton definition of the lower value and (5.1) with $u_\cdot = \alpha(v_\cdot)$

$$(5.2) \quad V^-(t, x) = \inf_{\alpha \in \Gamma} E_{tx}^+(\mathcal{J}).$$

In the max-plus stochastic control problem, we wish to replace Γ by a suitable subset $\Gamma_1 \subset \Gamma$ with the property that

$$(5.3) \quad V^+(t, x) = \inf_{\alpha \in \Gamma_1} E_{tx}^+(\mathcal{J}).$$

One choice of Γ_1 is the set Γ_S of *strictly progressive* strategies, defined as follows. An Elliott-Kalton strategy β for the maximizing controller is a progressive function from $L^2([t, T]; U)$ into $L^2([t, T]; \mathbb{R}^m)$. An Elliott-Kalton strategy α for the minimizing controller is *strictly progressive* if: for every Elliott-Kalton strategy β the equations

$$(5.4) \quad u_\cdot = \alpha(v_\cdot), \quad v_\cdot = \beta(u_\cdot)$$

have a solution u_\cdot, v_\cdot see [FS, Sec. 11.9] [F]. One could also take $\Gamma_1 = \bar{\Gamma}_S$, which is the closure of Γ_S in the uniform norm. See [FS, Remark 11.9.1].

It is an interesting open question whether another choice for Γ_1 can be found, which is defined in a more intuitive way than (5.4). Any such Γ_1 must include all progressive strategies α which are piecewise constant in time in the following sense: there is a partition of $[t, T]$ into subintervals I_1, \dots, I_M such that $\alpha(v)_s$ is constant on each subinterval I_j for every v . One could then take $\Gamma_1 = \tilde{\Gamma}_1$, where $\tilde{\Gamma}_1$ is the set of progressive α with the following property: in (5.1) write $\mathcal{J} = \mathcal{J}(\alpha)$, when $u_\cdot = \alpha(v_\cdot)$. Then $\alpha \in \tilde{\Gamma}_1$ if there exists a sequence α_m of piecewise constant strategies, such that

$$(5.5) \quad E_{tx}^+[\mathcal{J}(\alpha)] = \lim_{m \rightarrow \infty} E_{tx}^+[\mathcal{J}(\alpha_m)].$$

Strategies α which are “strictly progressive” according to whatever class Γ_1 is chosen should have a role in the max-plus stochastic control problem which is analogous to progressively measurable controls for systems governed by stochastic differential equations like (3.1). In the usual stochastic formulation, all progressively measurable controls in $L^2([t, T] \times \Omega)$ are mean square limits of processes which are piecewise constant in time. This property is used to show that the infimum of an expected cost criterion taken over progressively measurable control processes is the same as the infimum taken over the subset of piecewise constant control processes. Unfortunately, there does not seem to be a convenient max-plus analogue of this result. Hence, the author has not been able to find a more explicit characterization of the class $\tilde{\Gamma}_1$ of progressive strategies which satisfy (5.5).

6. Importance sampling. Monte Carlo simulation is a basic technique for the approximation of probabilities and expectations of random variables. Let Y be a random variable on some probability space (Ω, \mathcal{F}, P) , and let Y_1, \dots, Y_N be IID copies of Y . The Monte Carlo estimate for $E(Y)$ is the sample mean $\bar{Y}_N = N^{-1}(Y_1 + \dots + Y_N)$. If $Y = 1_A$ is the indicator function of an event A , then this gives the Monte Carlo estimate for the probability $P(A)$.

The theory of large deviations provides asymptotic estimates for probabilities of rare events and associated expectations. If $P(A)$ is very small, then A is rarely observed in a Monte Carlo sample unless the sample size N is very large. One must modify the Monte Carlo technique in order to obtain useful estimates of large deviations rates. The basic idea of importance sampling is to change a probability measure from P to another probability measure \tilde{P} and to consider IID copies η_1, \dots, η_N of $\eta = Y dP/d\tilde{P}$ where $dP/d\tilde{P}$ is the Radon-Nikodym derivative. The sample mean $\bar{\eta}_N$ is taken as an estimator for $E(Y) = \tilde{E}[Y dP/d\tilde{P}]$, where E and \tilde{E} denote expectations under probability measure P and probability measure \tilde{P} respectively.

In this section and Section 7 we outline the Dupuis-Wang approach to importance sampling as it applies to Markov diffusion processes. The method is based on variance reduction ideas. The goal is to choose \tilde{P} in some “admissible” class of probability measures such that the variance under \tilde{P} $\text{var}(\bar{\eta}_N)$ is minimized. Since $\text{var}(\bar{\eta}_N) = N^{-1}\text{var}(\eta)$ and $\bar{\eta}_N$ is unbiased, this is equivalent to choosing \tilde{P} to minimize

$$(6.1) \quad \tilde{E}(\eta^2) = E\left(Y^2 \frac{dP}{d\tilde{P}}\right).$$

Let x_s be a Markov diffusion, $t \leq s \leq T$, which satisfies the stochastic differential equation (SDE)

$$(6.2) \quad dx_s = b(s, x_s)ds + \Sigma(s, x_s)dw_s$$

where w_s is a P -Brownian motion on (Ω, \mathcal{F}_T) adapted to an increasing family of σ -algebras $\{\mathcal{F}_s\}$. We change probability measure from P to \tilde{P} via change of drift from b to another drift \tilde{b} using a Girsanov transformation as in [FS, p. 231]. Let ζ_s be a bounded $\{\mathcal{F}_s\}$ -progressively measurable process and let

$$\tilde{w}_s = w_s + \int_t^s \zeta_r dr.$$

Then \tilde{w}_s is a \tilde{P} -Brownian motion, where the Radon-Nikodym derivative $dP/d\tilde{P}$ satisfies

$$(6.3) \quad \frac{dP}{d\tilde{P}} = \exp \int_t^T \left(\zeta_r d\tilde{w}_r - \frac{1}{2} |\zeta_r|^2 dr \right).$$

Moreover,

$$(6.4) \quad \begin{aligned} dx_s &= \tilde{b}(s, x_s)ds + \Sigma(s, x_s)d\tilde{w}_s \\ \tilde{b}(s, x_s) &= b(s, x_s) - \Sigma(s, x_s)\zeta_s. \end{aligned}$$

Let us now assume that $Y = \Psi(x_T)$ depends on the final state x_T of the Markov diffusion process x_s . Under suitable assumptions on b , Σ and Ψ , the variance can be reduced to 0 by suitable choice of \tilde{b} . See [KP, Sec. 16.2]. The following version of this result is easily obtained, using the Ito differential rule. Let

$$(6.5) \quad \Phi(t, x) = E_{tx}[\Psi(x_T)].$$

THEOREM 6.1. *Assume that $b, \Sigma \in C_b^{1,2}(\overline{Q}_0)$ and that $\Psi \in C_b^2(\mathbb{R}^n)$. Then $\Phi \in C_b^{1,2}(\overline{Q}_0)$ and Φ satisfies the backward PDE*

$$(6.6) \quad \frac{\partial \Phi}{\partial t} + \frac{1}{2}trA(t, x)D_x^2\Phi + b(t, x) \cdot D_x\Phi = 0,$$

where $A = \Sigma\Sigma'$, with $\Phi(T, x) = \Psi(x)$. Moreover, if $\Psi(x) > 0$ and $\Psi^{-1}(x)$ is bounded, then the variance under \tilde{P} of $\eta = YdP/d\tilde{P}$ is 0 if

$$(6.7) \quad \tilde{b}(t, x) = b(t, x) + A(t, x)D_x \log \Phi(t, x).$$

Proof. (Sketch). The proof that $\Phi \in C_b^{1,2}(\overline{Q}_0)$ and satisfies (6.6) uses smooth dependence of solutions x_s to the SDE (6.2) on the initial state $x = x_t$. See [GS, Sec. 11]. To prove the second statement, apply to Ito differential rule to $\log \Phi(s, x_s)$ and use (6.6) to obtain since $\Phi(T, x_T) = \Psi(x_T)$

$$\begin{aligned} \log \Psi(x_T) &= \log \Phi(t, x) - \int_t^T (\zeta_r d\tilde{w}_r - \frac{1}{2}|\zeta_r|^2 dr), \\ \zeta_r &= -\Sigma'(r, x_r)D_x \log \Phi(r, x_r). \end{aligned}$$

By (6.3), $\eta = \Psi(x_T)dP/d\tilde{P}$ equals the constant $\Phi(t, x)$ almost surely. Hence the variance of η under \tilde{P} is 0. □

REMARK 6.1. Results similar to Theorem 6.1 can be obtained under different assumptions. If the matrices $A(t, x)$ are positive definite with bounded inverses $A^{-1}(t, x)$, then results about linear parabolic PDEs can be used. See [FS, Thm. 6.3.1].

The exit probability problem, with $\Phi(t, x) = P_{tx}(x_T \in B)$ can also be considered. Although an indicator function $\Psi(x) = 1_B(x)$ does not satisfy our assumptions, the exit problem can be studied by PDE methods if $A^{-1}(t, x)$ is bounded. See [FS, Secs. 6.6. and 7.10]. For the exit problem, the optimal \tilde{P} can be obtained by conditioning P on the event $x_T \in B$. See, for example, Sec. 6.4 of the first edition of [FS].

From the viewpoint of applications Theorem 6.1 is generally not helpful, since it depends on solving the backward PDE for $\Phi(t, x)$ which is the expectation of the

random variable $\Psi(x_T)$ to be estimated via Monte Carlo. In Section 7 we will outline an alternative technique due to Dupuis and Wang [DW1][DW2][DSW] based on small noise ($\epsilon \rightarrow 0$) asymptotics and approximate supersolutions to certainty-equivalent HJB equations. In many examples this technique gives changes of probability measure which are nearly optimal for small ϵ .

To describe this technique, let us first rewrite the problem of choosing ζ_r to minimize $\tilde{E}(\eta^2)$ in (6.1) as a risk sensitive stochastic control problem. For this purpose, we make another change of probability measure, from P to \hat{P} , via Girsanov such that

$$(6.8) \quad \frac{d\hat{P}}{dP} = \exp \int_t^T \left(\zeta_r dw_r - \frac{1}{2} |\zeta_r|^2 dr \right).$$

Then (6.2) can be rewritten as

$$(6.9) \quad \begin{aligned} dx_s &= (b(s, x_s) + \Sigma(s, x_s)\zeta_s)ds + \Sigma(s, x_s)d\hat{w}_s, \\ w_s &= \hat{w}_s + \int_t^s \zeta_r dr \end{aligned}$$

and \hat{w}_s is a \hat{P} -Brownian motion. Then (6.1) with $Y = \Psi(x_T)$ becomes

$$(6.10) \quad \begin{aligned} \tilde{E}(\eta^2) &= E \left[(\Psi(x_T))^2 \frac{d\hat{P}}{dP} \exp \int_t^T |\zeta_r|^2 dr \right] \\ &= \hat{E}[\exp(\mathcal{J})], \end{aligned}$$

$$(6.11) \quad \mathcal{J} = \int_t^T |\zeta_r|^2 dr + 2 \log \Psi(x_T).$$

In (6.10), \hat{E} is expectation under \hat{P} .

Nearly deterministic Markov diffusions. Let $\epsilon > 0$ be a “small” parameter, and let x_s^ϵ satisfy (6.2) with $\Sigma(t, x) = \epsilon\sigma(t, x)$. Let $u_s = \epsilon\zeta_s$. Then x_s^ϵ satisfies

$$(6.12) \quad dx_x^\epsilon = f(s, x_s^\epsilon, u_s)ds + \epsilon\sigma(s, x_s^\epsilon)d\hat{w}_s$$

$$(6.13) \quad f(t, x, u) = b(t, x) + \sigma(t, x)u.$$

Let us assume that $b, \sigma \in C_b^{1,2}(\overline{Q_0})$. We also suppose that Y is of the form $Y = \exp[(2\epsilon^2)^{-1}G(x_T^\epsilon)]$ with $G \in C_b^2(\mathbb{R}^n)$. These assumptions can be weakened in various ways. See Remark 6.3 below. By (6.10), (6.11)

$$(6.14) \quad \tilde{E}(\eta^2) = \hat{E}[\exp(\epsilon^{-2}\mathcal{J}^\epsilon)]$$

where \mathcal{J}^ϵ is as in (3.2) with $L(u) = |u|^2$. If we take initial data $x_t^\epsilon = x$ and regard u_s as a minimizing control, then the problem of choosing u_s to minimize $\tilde{E}(\eta^2)$ becomes

a risk-sensitive stochastic control problem of the form considered in Section 3. The control space is $U = \mathbb{R}^m$ (no constraints on the control u_s .) The certainty-equivalent Hamiltonian in (3.6) now has the form

$$(6.15) \quad \begin{aligned} H_\epsilon &= b(t, x) \cdot p + \frac{\epsilon^2}{2} \text{tr } a(t, x) \pi + \frac{1}{2} a(t, x) p \cdot p \\ &\quad + \min_u [\sigma(t, x) u \cdot p + |u|^2] \\ &= b(t, x) \cdot p + \frac{\epsilon^2}{2} \text{tr } a(t, x) \pi + \frac{1}{4} a(t, x) p \cdot p. \end{aligned}$$

The min occurs at $u = -\frac{1}{2}\sigma'p$. This provides the following recipe for finding the optimal drift $\tilde{b} = \tilde{b}^\epsilon$ for which the corresponding probability measure \tilde{P} minimizes (6.14).

THEOREM 6.2. *Let $V^\epsilon \in C^{1,2}(\overline{Q}_0)$ be a solution to the certainty-equivalent HJB equation (3.5) with boundary condition (3.7), such that V^ϵ and $D_x V^\epsilon$ are bounded on \overline{Q}_0 . Let*

$$(6.16) \quad \mathbf{u}^\epsilon(t, x) = -\frac{1}{2}\sigma'(t, x)D_x V^\epsilon(t, x)$$

$$(6.17) \quad \tilde{b}^\epsilon = b - \sigma \mathbf{u}^\epsilon = b + \frac{1}{2}aD_x V^\epsilon.$$

Then \tilde{b}^ϵ is the optimal drift and V^ϵ is the certainty-equivalent value function.

Theorem 6.2 can be proved using a standard verification argument in stochastic control theory, applied to the (non-certainty equivalent) value function $\exp(\epsilon^{-2}V^\epsilon)$. See [FS, Thm. 4.3.1 and Remark 4.3.3].

In (6.5) $\Phi = \Phi^\epsilon$ now has the form

$$(6.18) \quad \Phi^\epsilon(t, x) = E_{tx}[\exp(2\epsilon^2)^{-1}G(x_T^\epsilon)].$$

Let

$$(6.19) \quad I^\epsilon(t, x) = \epsilon^2 \log \Phi^\epsilon(t, x) = \mathcal{E}_{tx} \left[\frac{1}{2}G(x_T^\epsilon) \right].$$

THEOREM 6.3. *$V^\epsilon = 2I^\epsilon$ is the certainty-equivalent value function for the risk sensitive stochastic control problem.*

Proof. (Sketch). In (6.6) we take $A = \epsilon^2 a$. Then Theorem 6.1 and a routine calculation shows that $I^\epsilon \in C_b^{1,2}(\overline{Q}_0)$ and $2I^\epsilon$ satisfies the certainty-equivalent HJB equation (3.5) and the boundary condition $2I^\epsilon(T, x) = G(x)$. By Theorem 6.2, $V^\epsilon = 2I^\epsilon$ is the certainty-equivalent value function. \square

Small noise limits. We begin by recalling the following Freidlin-Wentzell large deviations result [FW]:

$$(6.20) \quad \lim_{\epsilon \rightarrow 0} I^\epsilon(t, x) = I^0(t, x),$$

where $I^0(t, x)$ is the value function for the following deterministic control problem. The state ξ_s satisfies

$$(6.21) \quad \frac{d\xi_s}{ds} = b(s, \xi_s) + \sigma(s, \xi_s)v_s, t \leq s \leq T,$$

with control $v_s \in \mathbb{R}^m$ and $\xi_t = x$. Then

$$(6.22) \quad I^0(t, x) = \sup_v \left[-\frac{1}{2} \int_t^T |v_s|^2 ds + \frac{1}{2} G(\xi_T) \right].$$

In large deviations terminology, $-I^0(t, x)$ is the rate function. Moreover, I^0 is the unique bounded, Lipschitz continuous viscosity solution to the PDE

$$(6.23) \quad \frac{\partial I^0}{\partial t} + b(t, x) \cdot D_x I^0 + \frac{1}{2} a(t, x) D_x I^0 \cdot D_x I^0 = 0$$

with boundary data $I^0(T, x) = \frac{1}{2} G(x)$. See [FS, Sec. 6.6]. For a PDE/viscosity solution proof of (6.20) see [FS, Thm. 6.6.2].

Differential game. From the discussion in Section 4, together with (6.14) and $L(u) = |u|^2$, we anticipate that the following differential game should arise from the risk sensitive control problem in the limit $\epsilon \rightarrow 0$. The game state x_s satisfies

$$(6.24) \quad \frac{dx_s}{ds} = b(s, x_s) + \sigma(s, x_s)u_s + \sigma(s, x_s)v_s$$

with initial data $x_t = x$. The game payoff is

$$(6.25) \quad P(t, x; u, v) = \int_t^T \left[|u_s|^2 - \frac{1}{2} |v_s|^2 \right] ds + G(x_T).$$

There are no control constraints. Thus, $u_s \in \mathbb{R}^m$, $v_s \in \mathbb{R}^m$. The Isaacs PDE is

$$(6.26) \quad \frac{\partial V}{\partial t} + H(t, x, D_x V) = 0$$

$$(6.27) \quad \begin{aligned} H(t, x, p) &= b(t, x) \cdot p + \min_u [\sigma(t, x)u \cdot p + |u|^2] \\ &\quad + \max_v \left[\sigma(t, x)v \cdot p - \frac{1}{2} |v|^2 \right] \\ &= b(t, x) \cdot p + \frac{1}{4} a(t, x) p \cdot p. \end{aligned}$$

The min and max are taken over $u \in \mathbb{R}^m$, $v \in \mathbb{R}^m$. Since min and max appear separately in (6.27) the Isaacs minimax condition holds. Thus, we anticipate that upper and lower differential game values should be equal. However, there is a difficulty in defining the differential game value since the control spaces \mathbb{R}^m are not compact. This difficulty can be avoided by introducing artificial constraints $|u_s| \leq R$, $|v_s| \leq R$. Let $V_R(t, x)$ be the Elliott-Kalton differential game value with these constraints.

THEOREM 6.4. *As $\epsilon \rightarrow 0$, the certainty-equivalent value function $V^\epsilon = 2I^\epsilon$ tends to $V^0 = 2I^0$, uniformly on compact subsets of $\overline{Q_0}$. Moreover,*

- (a) V^0 is the unique bounded, Lipschitz continuous viscosity solution of (6.26) with $V^0(T, x) = G(x)$;
- (b) There exists R_0 such that $V^0(t, x) = V_R(t, x)$ for every $R \geq R_0$.

Proof. . (Sketch) By [FS, Thm. 6.6.2], $I^\epsilon \rightarrow I^0$ uniformly on compact sets and I^0 is the unique bounded, Lipschitz continuous viscosity solution of (6.23) with boundary data $I^0(T, x) = \frac{1}{2}G(x)$. Hence, V^0 is a bounded, Lipschitz continuous viscosity solution to (6.26) with $V^0(T, x) = G(x)$. Uniqueness of V^0 and statement (b) follow using arguments in [FS, Sec. 11.7]. In fact, if M is a Lipschitz constant for $V^0(t, \cdot)$ and K is a bound for $|\sigma'(t, x)p|$ when $|p| \leq M$, then it suffices to take $R_0 = K$. □

REMARK 6.2. The constant $R_0 = K$ in the proof is suggested by the following formal argument. At each point (t, x) where V^0 is differentiable, $|D_x V^0(t, x)| \leq M$. If we take $p = D_x V^0(t, x)$ in (6.27), then the min over u and max over v are the same with the constraints $|u| \leq R, |v| \leq R$ for $R \geq K$. This is not a rigorous argument since $D_x V^0(t, x)$ only exists for almost all x .

REMARK 6.3. The rather strong assumptions that $b, \sigma \in C_b^{1,2}(\overline{Q}_0)$ and $G \in C_b^2(\mathbb{R}^n)$ can be weakened in various ways. In Theorem 6.3 it suffices to assume that G is bounded and uniformly continuous, since any such G is the uniform limit as $m \rightarrow \infty$ of a sequence of functions $G_m \in C_b^2(\mathbb{R}^n)$. However, the optimal drift \mathbf{u}^ϵ in (6.16) should then be replaced by nearly optimal \mathbf{u}_m^ϵ obtained when G is replaced by G_m . Theorem 6.4 remains true if $b, \sigma \in C_b^1(\overline{Q}_0)$ and $G \in C_b^1(\mathbb{R}^m)$. However, for V^ϵ to be a smooth solution to (3.5) the matrices $a(t, x)$ should be assumed to have bounded inverses. Otherwise, V^ϵ is a solution to (3.5) in the viscosity sense.

7. Supersolution method. In this section we describe a method for obtaining changes of probability measure for the importance sampling problem in Section 6. The method is based on considering functions which are approximately supersolutions of the HJB equation for the certainty-equivalent value function V^ϵ . This method is due to Dupuis and Wang. In their formulation, certainty-equivalent expectations are replaced by their negatives, and supersolutions are replaced by subsolutions. This method was applied in [DW1][DW2][DSW] to obtain approximately optimal changes of probability measure for various types of importance sampling problems, including large deviations for sums of IID random variables and rare buffer overflows in queuing systems.

For $\delta > 0$ let us call $W(t, x)$ a smooth δ -supersolution to (6.26) and the boundary data if $W \in C^{1,2}(\overline{Q}_0)$ and

$$(7.1) \quad \frac{\partial W}{\partial t} + H(t, x, D_x W) \leq \delta$$

$$(7.2) \quad G(x) \leq W(T, x) + \delta.$$

The min in (6.27) with $p = D_x W$ is attained at

$$(7.3) \quad \mathbf{u}(t, x) = -\frac{1}{2}\sigma'(t, x)D_x W(t, x).$$

The corresponding change of probability measure from P to \hat{P} corresponds to a change of drift from b to \hat{b} (see (6.9)), where

$$(7.4) \quad \hat{b} = b - \frac{1}{2}aD_x W.$$

We recall from Section 6 the notations

$$\eta = Y \frac{dP}{d\hat{P}}, \quad Y = \exp \left[(2\epsilon^2)^{-1} G(x_T^\epsilon) \right].$$

THEOREM 7.1. *Assume that $W \in C_b^{1,2}(\bar{Q}_0)$ is a smooth δ -supersolution to (6.26) and the boundary data. Then there exist $\epsilon_0 > 0$, $C > 0$ such that, for $0 < \epsilon < \epsilon_0$*

$$(7.5) \quad \tilde{\mathcal{E}}_{tx}(\eta^2) \leq W(t, x) + C\delta.$$

Proof. . (Sketch) Let $\phi = \exp(\epsilon^{-2}W)$. Then for small enough ϵ

$$(7.6) \quad \frac{\partial \phi}{\partial t} + \hat{b} \cdot D_x \phi + \frac{\epsilon^2}{2} \text{tr } a D_x^2 \phi + \epsilon^{-2}(|u|^2 - 2\delta)\phi \leq 0.$$

Let x_s^ϵ be the solution of (6.12) with $u_s = \mathbf{u}(s, x_s^\epsilon)$ and initial data $x_t^\epsilon = x$. By (7.6) and the Feynman-Kac formula [FS, p. 233]

$$\hat{E}_{tx} \left\{ \phi(T, x_T^\epsilon) \exp \left[\epsilon^{-2} \int_t^T (|u_s|^2 - 2\delta) ds \right] \right\} \leq \phi(t, x).$$

The proof is completed by using (6.14), (7.2) and taking $\epsilon^2 \log$. \square

The usefulness of Theorem 7.1 depends on finding good choices for W without solving PDEs. To find $I^0(t_1, x_1)$ for a particular set (t_1, x_1) of initial data, the solution to the control problem (6.21)–(6.22) is needed. This does not require solving the PDE (6.23). Suppose that a δ -supersolution W is found with $W(t_1, x_1) \leq 2I^0(t_1, x_1) + \beta$. The variance of η under \hat{P} is

$$\begin{aligned} \text{var}(\eta) &= \tilde{E}_{t_1 x_1}(\eta^2) - [\tilde{E}_{t_1 x_1}(\eta)]^2 \\ &= \tilde{E}_{t_1 x_1}(\eta^2) - [\Phi^\epsilon(t_1, x_1)]^2 \end{aligned}$$

with Φ^ϵ as in (6.18). Suppose that

$$W(t_1, x_1) \leq 2I^0(t_1, x_1) + \beta.$$

Since $V^\epsilon = 2I^\epsilon = 2\epsilon^2 \log \Phi^\epsilon$ tends to $2I^0$ as $\epsilon \rightarrow 0$, we obtain for small ϵ the inequality

$$(7.7) \quad \epsilon^2 \log \left\{ \frac{(\text{var}(\eta))^{\frac{1}{2}}}{\Phi^\epsilon(t_1, x_1)} \right\} \leq \frac{1}{2} C_1 \delta + \beta.$$

For small δ and β , this gives a sharp upper bound for the standard deviation/mean ratio. [DW2] has a more extended discussion of the supersolution method, including several examples in which $W(t, x)$ is a smoothed version of the maximum of functions W_1, \dots, W_K , where each W_k is an affine function. The method was applied in [DSW] to problems of rare buffer overflows in tandem queues.

REFERENCES

- [BB] T. BASAR AND P. BERNHARD, *H[∞]-Optimal Control and Related Minimax Design Problems*, 2nd Edition, Birkhauser, Boston, 1995.
- [BN] A. BENSOUSSAN AND H. NAGAI, *Min-max characterization of a small noise limit on risk sensitive control*, SIAM J. Control Optim., 35(1997), pp. 1093-1115.
- [DSW] P. DUPUIS, A.D. SEZER AND H. WANG, *Dynamic importance sampling for queuing networks*, preprint.
- [DW1] P. DUPUIS AND H. WANG, *Importance sampling. large deviations and differential games*, Stochastics and Stochastics Reports, 76(2004), pp. 481–508.
- [DW2] P. DUPUIS AND H. WANG, *Subsolutions of an Isaacs equation and efficient schemes for importance sampling*, Math. of Operations Research, to appear.
- [F] W. H. FLEMING, *Max-plus stochastic control*, In: Stochastic Theory and Control (ed. B. Pasik-Duncan) Lecture Notes in Control Info. Sci., 280(2002), Springer Verlag, New York, pp. 111–119.
- [FHH] W. H. FLEMING AND D. HERNANDEZ-HERNANDEZ, *Risk sensitive control of finite state machines on an infinite horizon*, SIAM J. Control Optim., 35(1997), pp. 1790–1810.
- [FM1] W. H. FLEMING AND W. M. MCEANEY, *Risk sensitive optimal control and differential games*, Lecture Notes in Control Info. Sci., 184(1992), Springer Verlag, New York, pp. 185–197.
- [FM2] W. H. FLEMING AND W. M. MCEANEY, *Risk sensitive control on an infinite time horizon*, SIAM J. Control Optim., 33(1995), pp. 1881–1915.
- [FS] W. H. FLEMING AND H. M. SONER, *Controlled Markov Processes and Viscosity Solutions*, Second Edition, Springer Verlag, New York, 2006.
- [FW] M. I. FREIDLIN AND A. D. WENTZELL, *Random Perturbations of Dynamical Systems*, Springer Verlag, New York, 1984.
- [GS] I. I. GIKHMAN AND A. V. SKOROKHOD, *Stochastic Differential Equations*, Springer Verlag, New York, 1972.
- [HJ] J. W. HELTON AND M. R. JAMES, *Extending H[∞] Control to Nonlinear Systems*, SIAM, Philadelphia, 1999.
- [J] M. R. JAMES, *Asymptotic analysis of nonlinear stochastic risk sensitive control and differential games*, Math. of Control, Signals and Syst., 5(1992), pp. 401–417.
- [KP] P. E. KLOEDEN AND E. PLATEN, *Numerical Solution of Stochastic Differential Equations*, Springer Verlag, New York, 1992.
- [MS] V.P. MASLOV AND S.M. SAMBORSKI (eds), *Idempotent Analysis*, Advances in Soviet Mathematics, 13, American Math. Society, Providence, RI, 1992.
- [N] H. NAGAI, *Bellman equations of risk sensitive control*, SIAM J. Control Optim., 34(1996), pp. 74–101.