INTERNATIONAL JOURNAL OF ROBUST AND NONLINEAR CONTROL, VOL. 6, 1–19 (1996)

NONLINEAR DISCRETE-TIME RISK-SENSITIVE OPTIMAL CONTROL

MARCO C. CAMPI

Dipartimento di Electtronica per l'Automazione, Univerità degli Studi di Brescia, via Branze 38, 25123 Brescia, Italy

AND

MATTHEW R. JAMES

Department of Systems Engineering and Cooperative Research Centre for Robust and Adaptive Systems, Australian National University, Canberra, ACT 0200, Australia

SUMMARY

This paper is devoted to the study of the connections among risk-sensitive stochastic optimal control, dynamic game optimal control, risk-neutral stochastic optimal control and deterministic optimal control in a nonlinear, discrete-time context with complete state information. The analysis worked out sheds light on the profound links among these control strategies, which remain hidden in the linear context. In particular, it is shown that, under suitable parameterizations, risk-sensitive control can be regarded as a control methodology which combines features of both *stochastic* risk-neutral control and *deterministic* dynamic game control.

KEY WORDS nonlinear stochastic systems; risk-sensitive stochastic optimal control; dynamic games; large deviations; asymptotic series; robust control

1. INTRODUCTION

Over the last decade, robust control has attracted increasing interest. A huge stream of literature has been devoted to H_∞ control theory, which has shown to be a systematic and effective way to design robust controllers such that bounds on the H_∞-norm of certain transfer functions are achieved. The H_∞ approach leads to robust controllers of the worst-case type, in the sense that focuses attention on that disturbance input which produces the largest effect on the system output and tries to minimize such an effect. The min-max nature of H_∞ has been widely emphasized by connecting H_∞ control with deterministic game theory, see, for example, Doyle *et al.*,[7] Limebeer *et al.*[16] and Basar and Bernhard.[4]

The H_∞ approach is alternative to (and somewhat in contrast with) the well-developed H_2 design procedure. In H_2, uncertainties on the model are described as statistical variations with respect to a nominal situation. The control goal is then to minimize the mean value of certain cost functions over all the possible statistical occurrences. This approach is somewhat more

This paper was recommended for publication by editor M. J. Grimble

CCC 1049–8923/96/010001–19

Received 8 April 1993
Revised 22 April 1994

optimistic than H_∞ in that one does not assume that disturbances are necessarily malicious and act so as to hinder the controller objectives. The difference between H_∞ and H_2 is made even sharper when one considers the mathematical formulation of the two problems: the first one involves no probability, whereas the second one is inherently stochastic.

In 1973, Jacobson[11] replaced the usual quadratic cost function used in the H_2 control problem with a modified criterion obtained through an exponential transformation. More precisely, he considered the linear system

$$x_{j+1} = Ax_j + Bu_j + \xi_j$$

where u_j is the control variable and $\{\xi_j\}$ is a white Gaussian disturbance and introduced the cost function

$$V = \mathbf{E}[\exp J]$$

where

$$J = \sum_{j=0}^{N-1} \left(\tfrac{1}{2}u_j' R u_j + \tfrac{1}{2}x_j' Q x_j\right) + \tfrac{1}{2}x_N' \bar{Q} x_N$$

Since the exponential function is convex, the penalty in the occurrence of values of J larger than $E[J]$ outweighs the alleviation in penalty caused by the occurrence of some values less than $E[J]$. This corresponds to a pessimistic viewpoint in the control design: one acts as though all the uncertainties ξ_j affecting the state x_{j+1} were likely to turn out to one's disadvantage (risk-sensitive control). As the second derivative of the exponential function is increasing with its argument, the larger the value of J, the greater the predisposition to pessimism. As a consequence of this fact, one will try to keep J small. This leads to a conservative control policy with robust stability characteristics.

Jacobson,[11] studies the risk-sensitive control problem in the case of perfect state observation and shows that the feedback control law is linear in the state, the gain of the controller being computed from a generalized Riccati equation. Moreover, via explicit calculations, he shows that the structure of the corresponding controller is the same as that for a dynamic game control problem. This interesting result established a link for the first time between deterministic dynamic game problems and control problems based on the minimization of stochastic cost functions. The case of imperfect state observation was first addressed in Reference 12. However, it was not until Whittle[19] that a satisfactory treatment of this case was worked out. In this same paper, Whittle introduces a risk-sensitivity parameter and demonstrates that selecting this parameter to be too large may lead to situations in which, regardless of the control policy, the cost function is infinite. This causes paralysis in the control design in the conviction of inability to control. Instead, when the risk-sensitivity parameter goes to zero, the usual linear quadratic gaussian control (risk-neutral control) is recovered. Glover and Doyle[9] see also Bernstein and Haddad,[5] study the connection between risk-sensitive control and H_∞ control. By focusing on linear time-invariant regulators, they show that in the infinite horizon case the risk-sensitive criterion enforces a bound on the H_∞-norm of the closed-loop transfer function. Moreover, the optimal controller minimizes the entropy integral over the set of all controllers meeting the H_∞-norm bound. The interested reader is referred to Whittle[20] for a comprehensive presentation of risk-sensitive optimal control theory.

It is interesting to note that all the above connections between dynamic game and H_∞ control theories and risk-sensitive optimal control theory are a consequence of the linear quadratic context in which they have been worked out. In fact, if one turns to nonlinear systems and/or to cost functions which are not quadratic (or exponential of quadratic, in the risk-sensitive case), different solutions are obtained for the different control problems. Only in the linear-quadratic case, these solutions collapse and turn out to be coincident. This could make one to think that all these links are just an accidental consequence of special circumstances.

In the context of nonlinear systems with complete state information, the link was established with the aid of an asymptotic analysis in James[13] and Fleming and McEneaney.[8] These two papers treat the mathematically technical continuous-time case, making use of PDE viscosity solution methods. In this paper we are concerned with risk-sensitive optimal control for nonlinear *discrete-time* systems with complete state information.

By suitably parametrizing the corresponding cost function, we will be able to clarify the relationships between this control problem and various other optimal control strategies. To be specific, the value function for the risk-sensitive control problem we consider is

$$S^{\mu,\varepsilon}(x;k) = \inf_u \mathbf{E}\left[\exp \frac{\mu}{\varepsilon}\left\{\sum_{j=k}^{N-1} L(x_j^\varepsilon, u_j) + \Phi(x_N^\varepsilon)\right\}\right]$$

where u_j is the control variable and x_j^ε is the controlled process governed by the dynamical system

$$x_{j+1}^\varepsilon = b(x_j^\varepsilon, u_j) + \sqrt{\varepsilon}\,\xi_j, \quad \{\xi_j\} \text{ white Gaussian sequence}$$

After applying the logarithmic transformation

$$W^{\mu,\varepsilon}(x;k) = \frac{\varepsilon}{\mu} \log S^{\mu,\varepsilon}(x;k)$$

we study the limits of $W^{\mu,\varepsilon}(x;k)$ when the parameters μ and ε tend to zero (Section 3). For $\varepsilon\downarrow 0$ the value function of the dynamic game control problem is obtained, whereas the value function of the risk-neutral control problem is achieved for $\mu\downarrow 0$. When both $\mu\downarrow 0$ and $\varepsilon\downarrow 0$, $W^{\mu,\varepsilon}(x;k)$ tends to the value function of a deterministic optimal control problem. All these limits are illustrated in Figure 1.

These results shed light on the profound links between deterministic dynamic games and stochastic optimal control problems in the nonlinear context, where the solutions of the different problems are not coincident due to accidental circumstances. As can be seen from Figure 1, the risk-sensitive regulator problem can be regarded as a generalization of the

$$\begin{tikzcd}
W^{\mu,\varepsilon} \arrow[r, "\varepsilon \downarrow 0"] \arrow[d, "\mu \downarrow 0"'] & W^{\mu} \arrow[d, "\mu \downarrow 0"] \\
W^{\varepsilon} \arrow[r, "\varepsilon \downarrow 0"'] & W
\end{tikzcd}$$

Figure 1

deterministic control problem which combines features of both dynamic games and stochastic control. To better clarity this point, in Section 4 we shall work out the following first order expansion for $W^{a\rho,b\rho}(x,k)$:

$$W^{a\rho,b\rho}(x;k) = W(x;k) + \rho(W_g(x;k),\ W_n(x;k))\begin{pmatrix}a\\b\end{pmatrix} + o(\rho), \text{ as } \rho \downarrow 0 \tag{1}$$

In equation (1), $W(x:k)$ is the value function of the deterministic optimal control problem, $W_g(x;k)$ will be interpreted as an incremental cost due to a deterministic opposing player whereas $W_n(x;k)$ is a term due to the diffusion effect of stochastic noise.

The linkage between deterministic dynamic games and risk-sensitive stochastic control for nonlinear discrete-time systems with incomplete state information has recently been established by James *et al.*[14,15] We also refer the reader to Ball and Helton,[2] van der Shaft,[17] Isidori and Astolfi,[10] and Baras and James[3] for the formulation of H_∞ control problems in a nonlinear context.

2 THE OPTIMAL CONTROL PROBLEMS

In this section, we formulate the four optimal control problems which will be studied in the forthcoming sections.

2.1. Risk-sensitive stochastic optimal control problem

Consider the nonlinear discrete-time stochastic system described by

$$x_{j+1}^{\varepsilon} = b(x_j^{\varepsilon}, u_j) + \sqrt{\varepsilon}\,\xi_j \tag{2}$$

where $x_j^{\varepsilon} \in \mathbf{R}^n$ is the state vector, $u_j \in U \subset \mathbf{R}^m$ is the control variable and $\{\xi_j\}$ is a sequence of independently distributed Gaussian random variables with probability density $p(z) = (2\pi)^{-n/2} \exp\{-\frac{1}{2}|z|^2\}$. Moreover, consider two functions $L(\cdot,\cdot):\mathbf{R}^n \times U \to \mathbf{R}$ and $\Phi(\cdot):\mathbf{R}^n \to \mathbf{R}$ and assume that $b(\cdot,\cdot)$, $L(\cdot,\cdot)$ and $\Phi(\cdot)$ satisfy the following regularity conditions

(i) $b(\cdot,\cdot)$ is bounded ad continuous and $b(\cdot,u)$ is uniformly Lipschitz continuous, uniformly in $u \in U$.
(ii) $L(\cdot,\cdot)$ is bounded and continuous and $L(\cdot,u)$ is uniformly Lipschitz continuous, uniformly in $u \in U$.
(iii) $\Phi(\cdot)$ is bounded and Lipschitz continuous.

Denoting by $U(k)$ the class of control sequences $u = \{u_k, \ldots, u_{N-1}\}$ such that $u_j \in U$, $j = k, \ldots, N-1$, is measurable with respect to the σ-algebra generated by $\{x_k^{\varepsilon}, \ldots, x_j^{\varepsilon}\}$, the value function for the risk-sensitive optimal control problem in the interval $[k, N]$ is defined by

$$S^{\mu,\varepsilon}(x;k) = \inf_{u \in U(k)} \mathbf{E}_{x,k}\left[\exp \frac{\mu}{\varepsilon}\left\{\sum_{j=k}^{N-1} L(x_j^{\varepsilon}, u_j) + \Phi(x_N^{\varepsilon})\right\}\right]$$

where x_j^{ε} is generated by system (2) initialized at time k with state x.

2.2. Dynamic game optimal control problem

Consider the nonlinear discrete-time deterministic system described by

$$x_{j+1} = b(x_j, u_j) + z_j \tag{3}$$

where $x_j \in \mathbf{R}^n$ is the state vector, $u_j \in U \subset \mathbf{R}^m$ is the control variable, and $z_j \in \mathbf{R}^n$ is the input variable to be chosen by an opposing player. In the deterministic context, the set $U(k)$ denotes the class of control policies u such that for each $j = k, \ldots, N-1$ there exists a map $\bar{u}_j$: $\mathbf{R}^{n \times (j-k+1)} \to U$ with $u_j = \bar{u}_j(x_k, \ldots, x_j)$. The value function for the dynamic game optimal control problem is defined as

$$W^{\mu}(x;k) = \inf_{u \in U(k)} \sup_{z \in l_2([k, N-1], \mathbf{R}^n)} \left\{ \sum_{j=k}^{N-1} \left(L(x_j, u_j) - \frac{1}{2\mu} |z_j|^2 \right) + \Phi(x_N) \right\} \tag{4}$$

where x_j is generated by system (3) initialized at time k with state x, and z denotes the sequence $z_k, \ldots, z_{N-1}$.

2.3. *Risk-neutral stochastic optimal control problem*

The value function for the risk-neutral stochastic optimal control problem is given by

$$W^{\varepsilon}(x;k) = \inf_{u \in U(k)} \mathbf{E}_{x,k} \left[\sum_{j=k}^{N-1} L(x_j^{\varepsilon}, u_j) + \Phi(x_N^{\varepsilon}) \right] \tag{5}$$

where $U(k)$ is defined as in the risk-sensitive case and x_j^{ε} is generated by system (2) initialized at time k with state x.

2.4. *Deterministic optimal control problem*

Consider the nonlinear discrete-time deterministic system described by

$$x_{j+1} = b(x_j, u_j) \tag{6}$$

where $x_j \in \mathbf{R}^n$ is the state vector, and $u_j \in U \subset \mathbf{R}^m$ is the control variable.

The value function for the deterministic optimal control problem is defined as

$$W(x;k) = \inf_{u \in U(k)} \left\{ \sum_{j=k}^{N-1} L(x_j, u_j) + \Phi(x_N) \right\} \tag{7}$$

where $U(k)$ is defined as in the dynamic game case and x_j is generated by system (6) initialized at time k with state x.

3. LIMIT RESULTS

The goal of the present section is to establish the limit results relating the value functions of the four optimal control problems introduced in the previous section. We start by stating the dynamic programming equations for the different control problems.

Risk-sensitive stochastic optimal control problem

$$\begin{aligned} W^{\mu,\varepsilon}(x;k) &= \inf_{u \in U} \frac{\varepsilon}{\mu} \log \mathbf{E}_{x,k} \left[\exp \frac{\mu}{\varepsilon} \{ L(x,u) + W^{\mu,\varepsilon}(b(x,u) + \sqrt{\varepsilon}\xi_k; k+1) \} \right] \\ W^{\mu,\varepsilon}(x;N) &= \Phi(x) \end{aligned} \tag{8}$$

where $W^{\mu,\varepsilon}(x;k) = (\varepsilon/\mu) \log S^{\mu,\varepsilon}(x;k)$.

Dynamic game optimal control problem

$$\begin{aligned} W^\mu(x;k) &= \inf_{u\in U} \sup_z \left\{ L(x,u) - \frac{1}{2\mu} |z|^2 + W^\mu(b(x,u)+z;k+1) \right\} \\ W^\mu(x;N) &= \Phi(x) \end{aligned} \tag{9}$$

Risk-neutral stochastic optimal control problem

$$\begin{aligned} W^\varepsilon(x;k) &= \inf_{u\in U} \mathbf{E}_{x,k}[L(x,u) + W^\varepsilon(b(x,u)+\sqrt{\varepsilon}\xi_k;k+1)] \\ W^\varepsilon(x;N) &= \Phi(x) \end{aligned} \tag{10}$$

Deterministic optimal control problem

$$\begin{aligned} W(x;k) &= \inf_{u\in U} \{L(x,u) + W(b(x,u);k+1)\} \\ W(x;N) &= \Phi(x) \end{aligned} \tag{11}$$

We now focus our attention on the link between the value function of the risk-sensitive control problem and the one of the dynamic game control problem. The relevant result is given in Theorem 3.2. First, we analyse some important characteristics of $W^\mu(\cdot;k)$ and introduce a lemma useful in the proof of Theorem 3.2.

Assumptions (i) to (iii) in Section 2 imply that $W^\mu(\cdot;k), k=0,1,\ldots,N$, are bounded and Lipschitz continuous. The boundeness follows from the boundedness of $L(\cdot,\cdot)$ and $\Phi(\cdot)$. For $k=N$, the Lipschitz continuity is an immediate consequence of assumption (iii). For $k<N$, one can inductively assume that $W^\mu(\cdot;k+1)$ is Lipschitz continuous. Then, assumptions (ii) and (iii) entail that $L(\cdot,u)-(1/2\mu)|z|^2 + W^\mu(b(\cdot,u)+z;k+1)$ is Lipschitz continuous uniformly in $(u,z)\in U\times\mathbf{R}^n$. Hence,

$$\sup_z \left\{ L(\cdot;u) - \frac{1}{2\mu}|z|^2 + W^\mu(b(\cdot,u)+z;k+1) \right\}$$

is Lipschitz continuous uniformly in $u\in U$ and

$$W^\mu(\cdot;k) = \inf_{u\in U}\sup_z \left\{ L(\cdot,u) - \frac{1}{2\mu}|z|^2 + W^\mu(b(\cdot,u)+z;k+1) \right\}$$

is Lipschitz continuous. These properties will justify the use of Lemma 3.1 in the proof of Theorem 3.2.

Lemma 3.1

Consider a function $f(\cdot,\cdot)\in C(\mathbf{R}^n\times\Theta,\mathbf{R})$, $\Theta\subset\mathbf{R}^p$, such that

(i) $\exists c_1, c_2 > 0$: $\quad f(z,\theta) \leqslant c_1 - c_2|z|^2, \quad \forall z\in\mathbf{R}^n, \quad \forall\theta\in\Theta$;
(ii) $\exists c_3$: $\quad \sup_z f(z,\theta) \geqslant c_3, \quad \forall\theta\in\Theta$;
(iii) $f(\cdot,\theta)$ is uniformly Lipschitz continuous on compact sets, uniformly in $\theta\in\Theta$.

Given a family of functions $f_\varepsilon(\cdot,\cdot)\in C(\mathbf{R}^n\times\Theta,\mathbf{R})$, $\varepsilon>0$, uniformly approaching $f(\cdot,\cdot)$:

$$\sup_{(z,\theta)\in\mathbf{R}^n\times\Theta} |f_\varepsilon(z,\theta) - f(z,\theta)| \to 0, \text{ as } \varepsilon\downarrow 0$$

we have

$$\lim_{\varepsilon \downarrow 0} \varepsilon \log \int_{\mathbf{R}^n} \exp \frac{1}{\varepsilon} f_\varepsilon(z,\theta)\, \mathrm{d}z = \sup_z f(z,\theta)$$

uniformly in $\theta \in \Theta$.

Proof. Given in the Appendix. □

The following theorem relates risk-sensitive stochastic optimal control and dynamic game optimal control.

Theorem 3.2

We have

$$\lim_{\varepsilon \downarrow 0} W^{\mu,\varepsilon}(x;k) = W^{\mu}(x;k), \qquad k = 0, 1, \ldots, N$$

uniformly in $x \in \mathbf{R}^n$, where $W^\mu(x;k)$ is the value of the dynamic game control problem (4).

Proof. For $k = N$ the result is trivially true. For $k < N$, the result will be proved by induction. Assume that

$$\lim_{\varepsilon \downarrow 0} W^{\mu,\varepsilon}(x;k+1) = W^{\mu}(x;k+1)$$

uniformly in $x \in \mathbf{R}^n$ and write

$$V^{\mu,\varepsilon}(x,u;k) = \frac{\varepsilon}{\mu} \log \int_{\mathbf{R}^n} \exp \frac{\mu}{\varepsilon} \left\{ L(x,u) - \frac{1}{2\mu} |z|^2 + W^{\mu,\varepsilon}(b(x,u)+z;k+1) \right\} (2\pi\varepsilon)^{-n/2}\, \mathrm{d}z \tag{12}$$

We want to apply Lemma 3.1 to $V^{\mu,\varepsilon}(x,u;k)$. To this end, define

$$\theta = (x,u);$$

$$f(z,\theta) = \mu\left\{ L(x,u) - \frac{1}{2\mu} |z|^2 + W^{\mu}(b(x,u)+z;k+1) \right\}$$

$$f_\varepsilon(z,\ \theta) = \mu\left\{ L(x,u) - \frac{1}{2\mu} |z|^2 + W^{\mu,\varepsilon}(b(x,u)+z;k+1) \right\} + \varepsilon \log(2\pi\varepsilon)^{-n/2}$$

Assumptions (i) and (ii) of Lemma 3.1 are satisfied in view of the boundedness of $L(\cdot,\cdot)$ (assumption (ii) in Section 2) and $W^\mu(\cdot;k+1)$. Assumption (iii) holds true thanks to the Lipschitz continuity of $W^\mu(\cdot;k+1)$. Then,

$$\lim_{\varepsilon \downarrow 0} V^{\mu,\varepsilon}(x,u;k) = \sup_z \left\{ L(x,u) - \frac{1}{2\mu} |z|^2 + W^{\mu}(b(x,u)+z;k+1) \right\}$$

uniformly in $(x,u) \in \mathbf{R}^n \times U$. Finally, thanks to the uniform validity of this limit,

$$\begin{aligned} \lim_{\varepsilon \downarrow 0} W^{\mu,\varepsilon}(x;k) &= \lim_{\varepsilon \downarrow 0} \inf_{u \in U} V^{\mu,\varepsilon}(x,u;k) \\ &= \inf_{u \in U} \sup_z \left\{ L(x,u) - \frac{1}{2\mu} |z|^2 + W^{\mu}(b(x,u)+z;k+1) \right\} \\ &= W^\mu(x;k) \end{aligned}$$

uniformly in $x \in \mathbf{R}^n$. □

Next, Theorem 3.3 studies the behaviour of $W^{\mu,\varepsilon}(x;k)$ when the risk-sensitivity parameter μ tends to zero.

Theorem 3.3

We have

$$\lim_{\mu\downarrow 0} W^{\mu,\varepsilon}(x;k) = W^{\varepsilon}(x;k), \qquad k=0,1,\dots,N$$

uniformly in $x\in\mathbf{R}^n$, where $W^{\varepsilon}(x;k)$ is the value of the risk-neutral stochastic control problem (5).

Proof. The theorem will be proved by induction. For $k=N$ the result is trivially true. Next, consider the function $V^{\mu,\varepsilon}(\cdot,\cdot;k)$ defined by equation (12) and assume that

$$\lim_{\mu\downarrow 0} W^{\mu,\varepsilon}(x;k+1) = W^{\varepsilon}(x;k+1) \tag{13}$$

uniformly in $x\in\mathbf{R}^n$. Since $L(\cdot,\cdot)$ and $\Phi(\cdot)$ are bounded (Assumptions (ii) and (iii) in Section 2), $W^{\varepsilon}(\cdot,k+1)$ is bounded as well. Then, the inductive assumption (13) implies

$$\begin{aligned}
\lim_{\mu\downarrow 0} V^{\mu,\varepsilon}(x,u;k) &= \lim_{\mu\downarrow 0}\frac{\varepsilon}{\mu}\log\left\{\int_{\mathbf{R}^n}\exp\frac{\mu}{\varepsilon}\{L(x,u)+W^{\varepsilon}(b(x,u)+z;k+1)\}\right.\\
&\qquad\left.\times\exp\left\{-\frac{1}{2\varepsilon}|z|^2\right\}(2\pi\varepsilon)^{-n/2}\,\mathrm{d}z+o(\mu)\right\}\\
&= \lim_{\mu\downarrow 0}\frac{\varepsilon}{\mu}\log\left\{\int_{\mathbf{R}^n}\left(1+\frac{\mu}{\varepsilon}\{L(x,u)+W^{\varepsilon}(b(x,u)+z;k+1)\}\right)\right.\\
&\qquad\left.\times\exp\left\{-\frac{1}{2\varepsilon}|z|^2\right\}(2\pi\varepsilon)^{-n/2}\,\mathrm{d}z+o(\mu)\right\}\\
&= \int_{\mathbf{R}^n}\{L(x,u)+W^{\varepsilon}(b(x,u)+z;k+1)\}\exp\left\{-\frac{1}{2\varepsilon}|z|^2\right\}(2\pi\varepsilon)^{-n/2}\,\mathrm{d}z
\end{aligned}$$

uniformly in $(x,u)\in\mathbf{R}^n\times U$. The last part of the proof is entirely similar to that of Theorem 3.2. □

The value function for the deterministic optimal control problem is obtained as limiting case both from the risk-sensitive stochastic optimal control and the dynamic game optimal control problem.

Theorem 3.4

We have

$$\lim_{\varepsilon\downarrow 0} W^{\varepsilon}(x;k) = \lim_{\mu\downarrow 0} W^{\mu}(x;k) = W(x;k), \qquad k=0,1,\dots,N$$

uniformly in $x\in\mathbf{R}^n$, where $W(x;k)$ is the value of the deterministic optimal control problem (7).

Proof. Note that the result is trivially true for $k = N$ and assume the inductive assumption

$$\lim_{\varepsilon \downarrow 0} W^{\varepsilon}(x; k+1) = \lim_{\mu \downarrow 0} W^{\mu}(x; k+1) = W(x; k+1)$$

uniformly in $x \in \mathbf{R}^n$. Define

$$V^{\mu}(x, u; k) = \sup_{z} \left\{ L(x, u) - \frac{1}{2\mu} |z|^2 + W^{\mu}(b(x, u) + z; k+1) \right\}$$

and

$$V^{\varepsilon}(x, u; k) = = \int_{\mathbf{R}^n} \{L(x, u) + W^{\varepsilon}(b(x, u) + z; k+1)\} \exp\left\{ -\frac{1}{2\varepsilon} |z|^2 \right\} (2\pi\varepsilon)^{-n/2} \, dz$$

Similarly to $W^{\mu}(\cdot; k+1)$, also $W(\cdot; k+1)$ is Lipschitz continuous. Then, the inductive assumption implies

$$\begin{aligned} \lim_{\mu \downarrow 0} V^{\mu}(x, u; k) &= \lim_{\varepsilon \downarrow 0} V^{\varepsilon}(x, u; k) \\ &= L(x, u) + W(b(x, u); k+1) \end{aligned}$$

uniformly in $(x, u) \in \mathbf{R}^n \times U$. The last part of the proof is entirely analogous to that of Theorem 3.2. □

Remark 3.5

In full analogy with the proof that $\lim_{\varepsilon \downarrow 0} W^{\varepsilon}(x; k) = \lim_{\mu \downarrow 0} W^{\mu}(x; k) = W(x; k)$, $k = 0, 1, \ldots, N$, one can also prove that $\lim_{\rho \downarrow 0} W^{a\rho, b\rho}(x; k) = W(x; k)$, uniformly in $x \in \mathbf{R}^n$.

In this section, we have shown that the value functions of the four control problems introduced in Section 2 are linked to each other when certain parameters tend to zero. This result enlightens the profound links among these techniques in the general nonlinear context. In the next section, a first-order expansion supplying more details on these connections is provided.

4. FIRST-ORDER EXPANSION

In this section we provide an explicit first-order expansion formula for $W^{\mu,\varepsilon}(\cdot; k)$. This result illuminates how risk-neutral stochastic optimal control and dynamic game optimal control combined in determining the value function for the risk-sensitive stochastic optimal control problem. The expansion will be worked out under the following additional conditions (throughout the section, the symbols D and Δ will denote the gradient and the Laplacian operators, respectively).

(iv) $b(x, u) = f(x) + g(x)u$, $u \in U$, U compact and convex.
(v) $f(\cdot)$, $g(\cdot)$, $L(\cdot, \cdot)$ and $\Phi(\cdot)$ are of class C^3.
(vi) $D^2\Phi(x) \geqslant 0$, $\forall x \in \mathbf{R}^n$.
$D_u^2 L(x, u) > 0$, uniformly in $(x, u) \in \mathbf{R}^n \times U$.
(vii) $W(\cdot; k)$ is of class C^2 and $D^2 W(x; k) \geqslant 0$, $\forall x \in \mathbf{R}^n$, $k = 0, 1, \ldots, N-1$.

Under Assumptions (iv) to (vii)

$$\begin{aligned} D_u^2\{L(x, u) + W(b(x, u); k+1)\} &= D_u^2 L(x, u) + g(x)' D^2 W(f(x) + g(x)u; k+1) g(x) \\ &> 0, \qquad k = 0, 1, \ldots, N-1 \end{aligned}$$

uniformly in $(x, u) \in \mathbf{R}^n \times U$. Therefore, taking also into account the convexity assumption on U, the optimal control law for the deterministic control problem exists and is unique. Denote by

$$u^*(x; k) = \operatorname*{argmin}_{u \in U} \{L(x, u) + W(b(x, u); k + 1)\}$$

the optimal feedback control law at time k. We assume that

(viii) $u^*(\cdot; k), k = 0, 1, \ldots, N-1$, are of class C^3.
(ix) $u^*(x; k) \in \text{Interior}(U)$, $\forall x \in \mathbf{R}^n$, $k = 0, 1, \ldots, N-1$.

Remark 4.1

- Assumption (viii) implies that $W(\cdot; k), k = 0, 1, \ldots, N-1$, are of class C^3.
- If $L(x, u) = \Gamma(x) + \frac{1}{2}|u|^2$, then Assumption (ix) is always satisfied provided that U is 'sufficiently large'.

We first introduce the following lemma which will be useful in the proof of the expansion result.

Lemma 4.2

Given a family of uniformly bounded functions $f_\varepsilon(\cdot)$: $\mathbf{R}^n \to \mathbf{R}$, $\varepsilon > 0$, assume that the following expansion

$$f_\varepsilon(z) = f(z) + \varepsilon g(z) + o(\varepsilon), \text{ as } \varepsilon \downarrow 0$$

holds true uniformly on compact sets, where $f(\cdot) \in C^2(\mathbf{R}^n, \mathbf{R})$ and $g(\cdot) \in C(\mathbf{R}^n, \mathbf{R})$. Then,

$$\lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \log \int_{\mathbf{R}^n} \exp\left\{f_\varepsilon(z + \theta) - f(\theta) - \frac{1}{2\varepsilon}|z|^2\right\}(2\pi\varepsilon)^{-n/2}\, dz = g(\theta) + \tfrac{1}{2}|Df(\theta)|^2 + \tfrac{1}{2}\Delta f(\theta)$$

uniformly on compact sets.

Proof. Given in the Appendix. □

Denoting by $x_k^*(x; j)$, $j = k, k+1, \ldots, N$, the optimal state trajectory for the deterministic control problem with initial state x at time k, define the following functions

$$W_g(x; k) = \frac{1}{2} \sum_{j=k+1}^{n} |DW(x_k^*(x; j); j)|^2$$

$$W_n(x; k) = \frac{1}{2} \sum_{j=k+1}^{n} \Delta W(x_k^*(x; j); j)$$

Theorem 4.3

Given $a > 0$ and $b > 0$, we have the following expansion

$$W^{a\rho, b\rho}(x; k) = W(x; k) + \rho(W_g(x; k),\ W_n(x; k))\begin{pmatrix} a \\ b \end{pmatrix} + o(\rho), \qquad \text{as } \rho \downarrow 0$$

uniformly on compact subsets of $\mathbf{R}^n$.

Proof. The expansion is trivially true for $k = N$. Define

$$V(x, u; k) = L(x, u) + W(b(x, u); k+1), \qquad k = 0, 1, \ldots, N-1 \tag{14}$$

and $V^{a\rho, b\rho}(x, u; k)$ by equation (12) and consider the following preliminary result

$$V^{a\rho, b\rho}(x, u; k) = V(x, u; k) + \rho(W_g(b(x, u); k+1),\ W_n(b(x, u); k+1))\begin{pmatrix} a \\ b \end{pmatrix}$$

$$+ \rho(\tfrac{1}{2} | DW(b(x, u); k+1) |^2,\ \tfrac{1}{2}\Delta W(b(x, u); k+1))\begin{pmatrix} a \\ b \end{pmatrix}$$

$$+ o(\rho), \text{ as } \rho \downarrow 0$$

uniformly in $(x, u) \in \mathbf{R}^n \times U$. The theorem will be proved by showing that the preliminary result at time k implies the expansion at time k and that the expansion at time $k+1$ implies the preliminary result at time k.

Preliminary result at time $k \Rightarrow$ expansion at time k. Let X be a compact subset of $\mathbf{R}^n$. Write

$$H^{a\rho, b\rho}(x, u; k) = V(x, u; k) + \rho Q(x, u; k)$$

where

$$Q(x, u; k) = (W_g(b(x, u); k+1),\ W_n(b(x, u); k+1))\begin{pmatrix} a \\ b \end{pmatrix}$$

$$+ (\tfrac{1}{2} | DW(b(x, u); k+1) |^2,\ \tfrac{1}{2}\Delta W(b(x, u); k+1))\begin{pmatrix} a \\ b \end{pmatrix}$$

Name $u_H^{\rho *}(x; k)$ a minimum point of $H^{a\rho, b\rho}(x, \cdot\,; k)$. Since $Q(x, u; k)$ is bounded on $\mathbf{R}^n \times U$ and $D_u^2 V(x, u; k) > 0$ uniformly in $(x, u) \in \mathbf{R}^n \times U$,

$$\lim_{\rho \downarrow 0} | u_H^{\rho *}(x; k) - u^*(x; k) | = 0$$

uniformly in $\mathbf{R}^n$. Hence, with suitable $\bar{\rho}$,

$$u_H^{\rho *}(x; k) \in \text{Interior}(U), \qquad \forall x \in X, \qquad \rho \in (0, \bar{\rho})$$

For any $x \in X$, $v \in \mathbf{R}^m$, and $\rho \in (0, \bar{\rho})$, we then have

$$\begin{aligned} 0 &= D_u H^{a\rho, b\rho}(x, u_H^{\rho *}(x; k);\ k)v \\ &= D_u V(x, u^*(x; k); k)v + v' D_u^2 V(x, \bar{u}^\rho(x; k); k)(u_H^{\rho *}(x; k) - u^*(x; k)) \\ &\quad + \rho D_u Q(x, u_H^{\rho *}(x; k); k)v \end{aligned}$$

where $\bar{u}^\rho(x; k) \in \overline{u_H^{\rho *}(x; k)\ u^*(x; k)}$. Setting $v = (u_H^{\rho *}(x; k) - u^*(x; k))$ this implies

$$\begin{aligned} &(u_H^{\rho *}(x; k) - u^*(x; k))' D_u^2 V(x, \bar{u}^\rho(x; k); k)(u_H^{\rho *}(x; k) - u^*(x; k)) \\ &\qquad = -\rho D_u Q(x, u_H^{\rho *}(x; k); k)(u_H^{\rho *}(x; k) - u^*(x; k) \end{aligned}$$

Since $D_u^2 V(x, \bar{u}^\rho(x; k); k) > 0$ uniformly in $x \in X$ and $D_u Q(x, u_H^{\rho *}(x; k); k)$ is bounded, we obtain the uniform estimate

$$| u_H^{\rho *}(x; k) - u^*(x; k) | \leqslant c\rho, \quad \forall x \in X, \quad \rho \in (0, \bar{\rho}) \tag{15}$$

with c a suitable constant. Next, consider the following expansion

$$
\begin{aligned}
&H^{a\rho,b\rho}(x, u_H^{\rho,*}(x;k);k) = \\
&\quad V(x, u^*(x;k);k) + D_u V(x, u^*(x;k);k)(u_H^{\rho,*}(x;k) - u^*(x;k)) \\
&\quad + \tfrac{1}{2}(u_H^{\rho,*}(x;k) - u^*(x;k))' D_u^2 V(x, \bar{u}^\rho(x;k);k)(u_H^{\rho,*}(x;k) - u^*(x;k)) \\
&\quad + \rho Q(x, u^*(x;k);k) + \rho D_u Q(x, \bar{\bar{u}}^\rho(x;k);k)(u_H^{\rho,*}(x;k) - u^*(x;k))
\end{aligned}
$$

where $\bar{u}^\rho(x;k), \bar{\bar{u}}^\rho(x;k) \in \overline{u_H^{\rho,*}(x;k)u^*(x;k)}$. Taking into account that $D_u^2 V(x, \bar{u}^\rho(x;k);k)$ and $D_u Q(x, \bar{\bar{u}}^\rho(x;k);k)$ are bounded in X, the uniform estimate (15) used in this last equation leads to

$$\inf_{u \in U} H^{a\rho,b\rho}(x,u;k) = \inf_{u \in U} V(x,u;k) + \rho Q(x, u^*(x;k);k) + o(\rho), \text{ as } \rho \downarrow 0 \tag{16}$$

uniformly on X. On the other hand, the preliminary result at time k implies

$$\inf_{u \in U} V^{a\rho,b\rho}(x,u;k) = \inf_{u \in U} H^{a\rho,b\rho}(x,u;k) + o(\rho), \text{ as } \rho \downarrow 0 \tag{17}$$

uniformly on X. Equaitons (16) and (17) give

$$\inf_{u \in U} V^{a\rho,b\rho}(x,u;k) = \inf_{u \in U} V(x,u;k) + \rho Q(x, u^*(x;k);k) + o(\rho), \qquad \text{as } \rho \downarrow 0$$

uniformly on X. Since,

$$Q(x, u^*(x;k);k) = (W_g(x;k),\ W_n(x;k)) \begin{pmatrix} a \\ b \end{pmatrix}$$

the expansion at time k is proved.

Expansion at time $k+1 \Rightarrow$ preliminary result at time k. Set

$$
\begin{aligned}
\varepsilon &= b\rho \\
\theta &= b(x,u) \\
f_\varepsilon(z) &= \frac{a}{b}\, W^{a\rho,b\rho}(z;k+1) \\
f(z) &= \frac{a}{b}\, W(z;k+1) \\
g(z) &= (W_g(z;k+1),\ W_n(z;k+1)) \begin{pmatrix} (a/b)^2 \\ a/b \end{pmatrix}
\end{aligned}
$$

The regularity assumptions on $f(\cdot)$ and $g(\cdot)$ assumed in Lemma 4.2 hold true in view of the regularity of $W(\cdot;k+1)$. Moreover, the expansion for $f_\varepsilon(\cdot)$ is satisfied thanks to the inductive assumption. Then, the result stated in Lemma 4.2 gives

$$\lim_{\rho \downarrow 0} \frac{1}{\rho} \{V^{a\rho,b\rho}(x,u;k) - V(x,u;k)\}$$

$$= \lim_{\rho \downarrow 0} \frac{b^2}{a} \frac{1}{b\rho} \log \int_{\mathbf{R}^n} \exp\left\{\frac{a}{b} W^{a\rho,b\rho}(b(x,u)+z;k+1) - \frac{a}{b} W(b(x,u);k+1)\right.$$

$$\left. - \frac{1}{2b\rho} |z|^2 \right\} (2\pi b\rho)^{-n/2} \, dz$$

$$= (W_g(b(x,u);k+1), W_n(b(x,u);k+1)) \binom{a}{b}$$

$$+ (\tfrac{1}{2} |DW(b(x,u);k+1)|^2, \tfrac{1}{2} \Delta W(b(x,u);k+1)) \binom{a}{b}$$

uniformly in $(x,u) \in \mathbf{R}^n \times U$. □

Remark 4.4

In the expansion stated in Theorem 4.3, the term $a\rho W_g(x;k)$ represents the first-order variation of the value function due to dynamic game features of risk-sensitive control. As a matter of fact, such a term can be easily interpreted as the incremental cost due to the noncooperative action of the opposing player in the dynamic game control problem. At time j, the opposing player choose z so as to drift the state away from the optimal trajectory $x_k^*(x;\cdot)$ with the goal of maximizing the dynamic programming cost function

$$2a\rho L(x,u) - |z|^2 + 2a\rho W^{\mu}(b(x,u)+z;k+1)$$
$$= [2a\rho L(x,u) - |z|^2 + 2a\rho W(b(x,u)+z;k+1)] + o(\rho)$$

(See Section 3). Then, up to terms vanishing as $o(\rho)$, the best choice for z will be that value which minimizes the term in the square bracket above. Consequently, z will be proportional to $DW(x_k^*(x;j);j)$ and the corresponding increment in the value function will be proportional to the square of this expression. The term $b\rho W_n(x;k)$ can be interpreted as the incremental cost due to the diffusion effect of noise around the optimal trajectory $x_k^*(x,\cdot)$, since the noise action tends to spread out the state in all the directions.

Remark 4.5

For clarity of exposition, in Theorem 4.3 the expansion for $W^{a\rho,b\rho}(x;k)$ has been worked out under boundedness conditions on $b(\cdot,\cdot)$, $L(\cdot,\cdot)$ and $\Phi(\cdot)$. However, the result is valid even under milder growth conditions on such functions. For instance, it is possible to see that the following set of assumptions suffices for the expansion to hold true, provided $\rho > 0$ is sufficiently small:

(i) $b(x,u) = f(x) + g(x)u$, $U = \mathbf{R}^m$.
(ii) $L(x,u) = \frac{1}{2}u'Ru + \Gamma(x)$, $R > 0$.
(iii) $f(\cdot)$, $g(\cdot)$, $\Gamma(\cdot)$ and $\Phi(\cdot)$ are of class C^3.
(iv) $Df(x)$, $D^2\Gamma(x)$ and $D^2\Phi(x)$ bounded.
(v) $D^2\Phi(x) \geqslant 0$; $D^2W(x;k) \geqslant 0$, $k = 0,1,\ldots,N-1$, $\forall x \in \mathbf{R}^n$.
(vi) $u^*(\cdot;k)$, $k = 0,1,\ldots,N-1$, are of class C^3.

Remark 4.6

Under the same assumptions as in Remark 4.5, it is possible to work out the following first-order expansion for the risk-sensitive control law:

$$\begin{aligned} u^{a\rho,b\rho*}(x;k) &= u^*(x;k) \\ &\quad - \{D_{uu}V(x,u^*(x;k);k)\}^{-1} \\ &\quad \times \rho(D_u\overline{W_g}(x,u^*(x;k);k+1), D_u\overline{W_n}(x,u^*(x;k);k+1))\begin{pmatrix} a \\ b \end{pmatrix} \\ &\quad + o(\rho), \text{ as } \rho \downarrow 0 \end{aligned}$$

uniformly on compact sets, where $u^*(x;k)$ is the control law at time k for the deterministic control problem, $V(\cdot,\cdot;k+1)$ is defined in Equation (14) and

$$\overline{W_g}(x,u;k) = \frac{1}{2}\sum_{j=k}^{N} |DW(x_k^*(b(x,u);j);j)|^2$$

$$\overline{W_n}(x,u;k) = \frac{1}{2}\sum_{j=k}^{N} \Delta W(x_k^*(b(x,u);j);j)$$

The second term in the right-hand-side of the expansion for the control law can be easily interpreted as a first-order variation due to dynamic game and diffusion effects (see the expansion stated in Theorem 4.3).

5. THE LINEAR QUADRATIC CASE

In this section, we consider the linear quadratic case obtained with the following choices

$$b(x,u) = Ax + Bu, \qquad U = \mathbf{R}^m$$

$$L(x,u) = \tfrac{1}{2}u'Ru + \tfrac{1}{2}x'Qx, \qquad R > 0, \qquad Q \geqslant 0$$

$$\Phi(x) = \tfrac{1}{2}x'\bar{Q}x, \qquad \bar{Q} \geqslant 0$$

This situation has been widely analysed in the literature. It is well known that the solution of the optimal control problems can be stated in terms of the solution of the following Riccati equations

$$P_k = Q + A'P_{k+1}A - A'P_{k+1}B[R + B'P_{k+1}B]^{-1}B'P_{k+1}A, \qquad P_N = \bar{Q}$$

$$\begin{cases} P_k^\mu = Q + A'\tilde{P}_{k+1}^\mu A - A'\tilde{P}_{k+1}^\mu B[R + B'\tilde{P}_{k+1}^\mu B]^{-1}B'\tilde{P}_{k+1}^\mu A, & P_N^\mu = \bar{Q} \\ \tilde{P}_{k+1}^\mu = P_{k+1}^\mu + P_{k+1}^\mu\left[\dfrac{1}{\mu}I - P_{k+1}^\mu\right]^{-1}P_{k+1}^\mu \end{cases}$$

In the linear quadratic case, we have the following well-known results.

Risk-sensitive stochastic optimal control[11]

$$W^{\mu,\varepsilon}(x;k) = \tfrac{1}{2}x'P_k^\mu x + \frac{\varepsilon}{\mu}\log F_k, \qquad \frac{1}{\mu}I - P_{k+1}^\mu > 0, \qquad k = 0,1,\ldots,N-1$$

where F_k is given by

$$F_k = F_{k+1}\sqrt{|(I - \mu P_{k+1}^\mu)^{-1}|}, \qquad F_N = 1$$

and the optimal control law writes

$$u^{\mu,\varepsilon *}[x;k) = -[R + B'\tilde{P}^{\mu}_{k+1}B]^{-1}B'\tilde{P}^{\mu}_{k+1}Ax$$

Dynamic game optimal control[4]

$$W^{\mu}(x;k) = \tfrac{1}{2}x'P^{\mu}_{k}x, \qquad \frac{1}{\mu}I - P^{\mu}_{k+1} > 0, \qquad k = 0, 1, \ldots, N-1$$

with the control law

$$u^{\mu *}(x;k) = u^{\mu,\varepsilon *}(x;k)$$

Risk-neutral stochastic optimal control[1]

$$W^{\varepsilon}(x;k) = \tfrac{1}{2}x'P_kx + \tfrac{1}{2}\varepsilon \sum_{j=k+1}^{N} \operatorname{tr} P_j$$

with the control law

$$u^{\varepsilon *}(x;k) = -[R + B'P_{k+1}B]^{-1}B'P_{k+1}Ax$$

Deterministic optimal control[1]

$$W(x;k) = \tfrac{1}{2}x'P_kx$$

and

$$u^{*}(x;k) = u^{\varepsilon *}(x;k)$$

By applying Theorem 4.3 one immediately obtains the following first-order expansion for $W^{a\rho,b\rho}(x;k)$

$$W^{a\rho,b\rho}(x;k) = \tfrac{1}{2}x'P_kx + \rho\left(\frac{1}{2}\sum_{j=k+1}^{N} |P_jx_k^{*}(x;j)|^2, \frac{1}{2}\sum_{j=k+1}^{N} \operatorname{tr} P_j\right)\binom{a}{b} + o(\rho), \text{ as } \rho \downarrow 0$$

uniformly on compact subsets of $\mathbf{R}^n$. This expansion shows explicitely the form taken in the linear quadratic case by the dynamic game and the diffusion contributions.

Further, the feedback control law for the risk-sensitive optimal control problem is given by

$$\begin{aligned} u^{a\rho,b\rho *}&(x;k) \\ &= -[R + B'P_{k+1}B]^{-1}B'P_{k+1}Ax \\ &\quad + \rho([R + B'P_{k+1}B]^{-1}B'(X_{k+1} + P^2_{k+1})(B[R + B'P_{k+1}B]^{-1}B'P_{k+1} - I)Ax, \ 0)\binom{a}{b} \\ &\quad + o(\rho), \qquad \text{as } \rho \downarrow 0 \end{aligned}$$

where X_k is recursively obtained as follows

$$\begin{aligned} X_k = A'(B[R + B'P_{k+1}B]^{-1}B'P_{k+1} - I)'(X_{k+1} + P^2_{k+1}) \\ \times (B[R + B'P_{k+1}B]^{-1}B'P_{k+1} - I)A, \qquad X_N = 0 \end{aligned}$$

Remark 5.1

Note that, unlike in the general case (see Remark 4.6), here the term in the expansion of $u^{a\rho,b\rho *}(x;k)$ due to noise effects is *zero*.

6. CONCLUSIONS

In this paper, we have studied the connections among risk-sensitive control, dynamic game, risk-neutral control and deterministic optimal control in a general nonlinear context. We have shown that:

(i) Dynamic game, risk-neutral control and deterministic optimal control can be recovered from risk-sensitive control by letting certain parameters go to zero.
(ii) A first order expansion has been worked out which shows how features of stochastic control and dynamic game combine in the risk-sensitive technique.

These results shed new light on the robust characteristics of risk-sensitive control which stem from the compensation for two kinds of disturbances, namely the diffusion effect of the noise and the action of the opposing player of the dynamic game. Moreover, these achievements show the generality of risk-sensitive control which, in a certain sense, comprises both min-max and stochastic techniques.

APPENDIX

Proof of Lemma 3.1

The proof is based on large deviations arguments (see, for example, Deuschel and Strook[6] and Varadhan.[18])

Given $\gamma > 0$, define

$$A_\gamma(\theta) \equiv \left\{ z' \colon f(z',\theta) \in \left[\sup_z f(z,\theta) - \gamma,\ \sup_z f(z,\theta) \right] \right\}$$

Assumptions (i) and (ii) imply that the set $\{\operatorname{argmin}_z f(z,\theta), \theta \in \Theta\}$ is bounded. Then, Assumption (iii) entails that the Lebesgue measure of $A_\gamma(\theta)$ is uniformly bounded from below, uniformly in $\theta \in \Theta$: $\mu(A_\gamma(\theta)) \geqslant \alpha(\gamma) > 0$, $\forall \theta \in \Theta$, where μ denotes Lebesgue measure in $\mathbf{R}^n$. On the other hand, Assumptions (i) and (ii) imply that such a measure is also bounded from above:

$$\mu(A_\gamma(\theta)) \leqslant \beta(\gamma), \qquad \forall \theta \in \Theta$$

Lower bound. With the notation

$$l(\varepsilon) = \sup_{(z,\theta) \in \mathbf{R}^n \times \Theta} | f_\varepsilon(z,\ \theta) - f(z,\theta) |$$

we have

$$\begin{aligned}
&\liminf_{\varepsilon \downarrow 0} \varepsilon \log \int_{\mathbf{R}^n} \exp \frac{1}{\varepsilon} f_\varepsilon(z,\ \theta)\, dz \\
&\quad \geqslant \liminf_{\varepsilon \downarrow 0} \varepsilon \log \int_{A_\gamma(\theta)} \exp \frac{1}{\varepsilon} \left\{ \sup_z f(z,\theta) - \gamma - l(\varepsilon) \right\} dz \\
&\quad \geqslant \liminf_{\varepsilon \downarrow 0} \left\{ \sup_z f(z,\theta) - \gamma - l(\varepsilon) + \varepsilon \log \alpha(\gamma) \right\} \\
&\quad = \sup_z f(z,\theta) - \gamma
\end{aligned}$$

uniformly in $\theta \in \Theta$.

Upper bound. Note first that ($\bar{A}$ denotes complement of set A)

$$\lim_{\varepsilon \downarrow 0} \int_{\bar{A}_\gamma(\theta)} \exp \frac{1}{\varepsilon} \left\{ f_\varepsilon(z,\ \theta) - \sup_z f(z,\theta) - l(\varepsilon) \right\} dz = 0 \tag{18}$$

uniformly in $\theta \in \Theta$. Indeed, from the elementary inequality

$$e^{\alpha x} \leqslant e^{(1-\alpha)\delta} e^{x}, \qquad \forall x \leqslant -\delta, \qquad \forall \alpha > 1,$$

it follows that

$$\begin{aligned}
&\int_{\bar{A}_\gamma(\theta)} \exp \frac{1}{\varepsilon} \left\{ f_\varepsilon(z,\ \theta) - \sup_z f(z,\theta) - l(\varepsilon) \right\} \mathrm{d}z \\
&\quad \leqslant \exp\left\{ \left(1 - \frac{\varepsilon'}{\varepsilon}\right) \frac{\gamma}{\varepsilon'} \right\} \int_{\bar{A}_\gamma(\theta)} \exp \frac{1}{\varepsilon'} \left\{ f_\varepsilon(z,\ \theta) - \sup_z f(z,\theta) - l(\varepsilon) \right\} \mathrm{d}z \\
&\quad \leqslant \exp\left\{ \left(1 - \frac{\varepsilon'}{\varepsilon}\right) \frac{\gamma}{\varepsilon'} \right\} \int_{\bar{A}_\gamma(\theta)} \exp \frac{1}{\varepsilon'} \left[f(z,\theta) - c_3 \right\} \mathrm{d}z, \qquad \forall \varepsilon < \varepsilon'
\end{aligned}$$

This implies (18).

Then,

$$\begin{aligned}
&\limsup_{\varepsilon \downarrow 0} \varepsilon \log \int_{\mathbf{R}^n} \exp \frac{1}{\varepsilon} f_\varepsilon(z,\ \theta)\, \mathrm{d}z \\
&\quad \leqslant \limsup_{\varepsilon \downarrow 0} \varepsilon \log \left\{ \int_{A_\gamma(\theta)} \exp \frac{1}{\varepsilon} \left\{ \sup_z f(z,\theta) + l(\varepsilon) \right\} \mathrm{d}z \right. \\
&\quad\quad \left. \times \left(1 + \frac{\int_{\bar{A}_\gamma(\theta)} \exp \frac{1}{\varepsilon} f_\varepsilon(z,\ \theta)\, \mathrm{d}z}{\int_{A_\gamma(\theta)} \exp \frac{1}{\varepsilon} \left\{ \sup_z f(z,\theta) + l(\varepsilon) \right\} \mathrm{d}z} \right) \right\} \\
&\quad \leqslant \limsup_{\varepsilon \downarrow 0} \left\{ \sup_z f(z,\theta) + l(\varepsilon) + \varepsilon \log \beta(\gamma) \right. \\
&\quad\quad \left. + \varepsilon \log \left(1 + \frac{1}{\alpha(\gamma)} \int_{\bar{A}_\gamma(\theta)} \exp \frac{1}{\varepsilon} \left\{ f_\varepsilon(z,\theta) - \sup_z f(z,\theta) - l(\varepsilon) \right\} \mathrm{d}z \right) \right\} \\
&\quad = \sup_z f(z,\theta)
\end{aligned}$$

uniformly in $\theta \in \Theta$.

Proof of Lemma 4.2

Fix $c > 0$. Thanks to the regularity assumptions on $f(\cdot)$ and $g(\cdot)$ and the expansion assumed for $f_\varepsilon(\cdot)$, the following holds true uniformly in $\theta \in \Theta$, where Θ is any compact set in $\mathbf{R}^n$,

$$\begin{aligned}
&\int_{|z| \leqslant c} \exp \left\{ f_\varepsilon(z+\theta) - f(\theta) - \frac{1}{2\varepsilon} |z|^2 \right\} (2\pi\varepsilon)^{-n/2}\, \mathrm{d}z \\
&\quad = \int_{|z| \leqslant c} \exp \left\{ f(z+\theta) + \varepsilon g(z+\theta) + o(\varepsilon) - f(\theta) - \frac{1}{2\varepsilon} |z|^2 \right\} (2\pi\varepsilon)^{-n/2}\, \mathrm{d}z \\
&\quad = \int_{|z| \leqslant c} \left\{ 1 + \frac{D \exp f(\theta)}{\exp f(\theta)} z + \tfrac{1}{2} z' \frac{D^2 \exp f(\bar{z}+\theta)}{\exp f(\theta)} z \right\} \\
&\quad\quad \times \{ 1 + \varepsilon g(z+\theta) \} \exp \left(-\frac{1}{2\varepsilon} |z|^2 \right) (2\pi\varepsilon)^{-n/2}\, \mathrm{d}z + o(\varepsilon)
\end{aligned}$$

where $\bar{z} \in \overline{0\ z}$. Let us compute separately the three integral terms. Thanks to the continuity of $g(\cdot)$,

$$\int_{|z| \leqslant c} \{ 1 + \varepsilon g(z+\theta) \} \exp \left(-\frac{1}{2\varepsilon} |z|^2 \right) (2\pi\varepsilon)^{-n/2}\, \mathrm{d}z = 1 + \varepsilon g(\theta) + o(\varepsilon)$$

uniformly in $\theta \in \Theta$. Thanks to the boundedness of $g(z+\theta)$ and $D \exp f(\theta)/\exp f(\theta)$, $|z| \leqslant c$, $\theta \in \Theta$,

$$\int_{|z| \leqslant c} \frac{D \exp f(\theta)}{\exp f(\theta)} z\{1 + \varepsilon g(z+\theta)\} \exp\left(-\frac{1}{2\varepsilon}|z|^2\right)(2\pi\varepsilon)^{-n/2} \, dz = o(\varepsilon)$$

uniformly in $\theta \in \Theta$. Thanks to the regularity assumptions on $f(\cdot)$ and $g(\cdot)$,

$$\int_{|z| \leqslant c} \tfrac{1}{2} z' \frac{D^2 \exp f(\tilde{z}+\theta)}{\exp f(\theta)} z\{1 + \varepsilon g(z+\theta)\} \exp\left(-\frac{1}{2\varepsilon}|z|^2\right)(2\pi\varepsilon)^{-n/2} \, dz$$

$$= \int_{|z| \leqslant c} \tfrac{1}{2} z' \frac{D^2 \exp f(\theta)}{\exp f(\theta)} z \exp\left(-\frac{1}{2\varepsilon}|z|^2\right)(2\pi\varepsilon)^{-n/2} \, dz + o(\varepsilon)$$

$$= \int_{|z| \leqslant c} \tfrac{1}{2}\{|Df(\theta)|^2 + \Delta f(\theta)\} \frac{1}{n} |z|^2 \exp\left(-\frac{1}{2\varepsilon}|z|^2\right)(2\pi\varepsilon)^{-n/2} \, dz + o(\varepsilon)$$

$$= \tfrac{1}{2}\varepsilon |Df(\theta)|^2 + \tfrac{1}{2}\varepsilon \Delta f(\theta) + o(\varepsilon)$$

uniformly in $\theta \in \Theta$. Then,

$$\int_{|z| \leqslant c} \exp\left\{f_\varepsilon(z,\ \theta) - f(\theta) - \frac{1}{2\varepsilon}|z|^2\right\}(2\pi\varepsilon)^{-n/2} = 1 + \varepsilon g(\theta) + \tfrac{1}{2}\varepsilon |Df(\theta)|^2 + \tfrac{1}{2}\Delta f(\theta) + o(\varepsilon) \quad (19)$$

uniformly in $\theta \in \Theta$. On the other hand, from the uniform boundedness of the sequence $f_\varepsilon(\cdot)$, it follows that

$$\int_{|z| \leqslant c} \exp\left\{f_\varepsilon(z+\theta) - f(\theta) - \frac{1}{2\varepsilon}|z|^2\right\}(2\pi\varepsilon)^{-n/2} \, dz = o(\varepsilon) \quad (20)$$

uniformly in $\theta \in \Theta$. The result readily follows from equations (19) and (20).

ACKNOWLEDGEMENTS

M. C. Campi wishes to acknowledge the support of the Consiglio Nationale delle Ricerche and M. R. James the support of the Cooperative Research Centre for Robust and Adaptive Systems.

REFERENCES

1. Anderson, B. D. O., and J. B. Moore, *Optimal Control. Linear Quadratic Methods*, Prentice Hall, Englewood Cliffs, NJ, 1989.
2. Ball, J. A., and J. W. Helton, 'H_∞ control for nonlinear plants: connections with differential games', Proc. 28th IEEE CDC, Tampa, FL, 1989, pp. 956–962.
3. James, M. R., and J. S. Baras, 'Robust H_∞ output feedback control for nonlinear systems', *IEEE Trans. Automatic Control.*
4. Basar, T., and P. Bernhard, *H_∞-Optimal Control and Related Minimax Design Problems: a Dynamic Game Approach*, Birkhauser, Boston, MA, 1991.
5. Bernstein, D., and P. Haddad, 'LQG control with an H_∞ performance bound: a Riccati equation approach, *IEEE Trans. Automatic Control*, **AC-34**, 293–305 (1989).
6. Deuschel, J. D., and D. W. Stroock, *Large Deviations*, Academic Press, New York, 1989.
7. Doyle, J. C., K. Glover, P. P. Khargonekar and B. A. Francis, 'State-space solutions to standard H_2 and H_∞ control problems', *IEEE Trans. Automatic Control*, **AC-34**, 831–874 (1989).
8. Fleming, W. H., and W. M. McEneaney, 'Risk sensitive optimal control and differential games', Proc. Conf. on Adaptive and Stochastic Control, University of Kansas, Springer Verlag, 1991.
9. Glover, K., and J. C. Doyle, 'State space formulae for all stabilizing controllers that satisfy and H_∞ norm bound and relations to risk sensitivity, *Systems and Control Letters*, **11**, 167–172 (1988).
10. Isidori, A., and A. Astolfi, 'Disturbance attenuation and H_∞ control via measurement feedback in nonlinear systems', *IEEE Trans. Automatic Control.*, **AC-37** 1283–1293 (1992).
11. Jacobson, D. H., 'Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games', *IEEE Trans. Automatic Control*, **AC-18**, 124–131 (1973).
12. Jacobson, D. H., *Extension of Linear/Quadratic Control, Optimization and Matrix Theory*, Academic Press, New York, 1977.
13. James, M. R., 'Asymptotic analysis of nonlinear stochastic risk-sensitive control and differential games, *Math. Control, Signals, and Systems*, **5**, 401–417 (1992).

14. James, M. R., J. S. Baras and R. J. Elliott, 'Risk-sensitive control and dynamic games for partially observed discrete-time nonlinear systems', *IEEE Trans. Automatic Control*, **AC-39**, 780–792 (1994).
15. James, M. R., J. S. Baras and R. J. Elliott, 'Output feedback risk-sensitive control and differential games for continuous-time nonlinear systems', *32nd IEEE CDC*, San Antonio, 1993.
16. Limebeer, D. J. N., B. D. O. Anderson, P. P. Khargonekar and M. Green, 'A game theoretic approach to H_∞ control for time varying systems', *SIAM J. Control Optim.*, **30**, 262–283 (1992).
17. van der Schaft, A. J., 'On a state space approach to nonlinear H_∞ control', *Systems and Control Letters*, **16**, 1–8 (1991).
18. Varadhan, S. R. S., *Large Deviations and Applications*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1984.
19. Whittle, P. 'Risk-sensitive linear/quadratic/gaussian control', *Adv. Appl. Prob.*, **13**, 764–777 (1981).
20. Whittle, P., *Risk-Sensitive Optimal Control*, John Wiley, New York, 1990.