Journal of Mathematical Systems, Estimation, and Control
Vol. 6, No. 3, 1996, pp. 1–24

An Information-State Approach to Risk-Sensitive Tracking Problems*

Iain B. Collings† Matthew R. James† John B. Moore†

Abstract

In this paper we use the information-state approach to obtain solutions to risk-sensitive quadratic control problems. Specifically we consider the case of tracking a desired trajectory. Results are presented for linear discrete-time models with Gaussian noise, and also for finite-discrete state, discrete-time hidden Markov models with continuous-range observations. These results give insight to more general information-state methods for nonlinear systems. Using such methods the tracking solution is obtained without appealing to a certainty equivalence principle. Limit results are presented which demonstrate the link to standard linear quadratic Gaussian control. Also presented is a discussion on achieving zero steady state error with risk-sensitive control policies. Simulation studies are presented to show some advantages of using the risk-sensitive approach.

Key words: risk sensitive control, LQG control

AMS Subject Classifications: 49A60, 60J05, 93C55, 93E20

1 Introduction

Recently there has been much interest in risk-sensitive control techniques. Such control policies lead to an optimal solution, similar to the case for linear quadratic Gaussian (LQG) control, however, with a risk-sensitive policy the controller's sensitivity to risk can be varied. One application area for risk-sensitive control has been in economics where risk-sensitivity is termed *hedging* or *risk-aversion*, for example Karp [9] and Caravani [5]. In these

*Received February 1994; received in final form September 22, 1994. Summary appeared in Volume 6, Number 3, 1996.
†Partially supported by the Australian Telecommunications and Electronics Research Board (ATERB) and the Co-operative Research Centre for Robust and Adaptive Systems ((CR)^{2}ASys).

papers it is seen that advantages can be gained from the risk-sensitive approach, for problems such as dynamic trading and futures market prediction.

The particular risk-sensitive control policy we consider in this paper involves an exponential in the cost function. This approach was first taken by Jacobson [6], when considering the risk-sensitive LQG problem with state feedback. Jacobson demonstrated a link between exponential performance criteria and deterministic differential games. He showed that the risk-sensitive approach provides a method for varying the the robustness of the controller and noted that in the case of no risk, or *risk-neutral*, the well known LQG solution [3] would result.

The risk-sensitive linear quadratic Gaussian (LQG) output feedback control problem was first solved by Whittle [12], where use was made of a risk-sensitive version of the certainty equivalence principle. This allowed the state estimation and control optimisation to be decoupled, solved separately and then re-coupled. The continuous-time case was solved by Bensoussan and van Schuppen [4] using a different technique, one which generalises to the nonlinear case. Recent developments in risk-sensitive control have included a solution to the output feedback control problem for nonlinear systems using information-state techniques (James, Baras and Elliott [7]). The solution is of course infinite dimensional, but does not require the use of a certainty equivalence principle.

In this paper we present the output feedback risk-sensitive LQG solution derived via the methods in [4, 7]. Specifically, we consider the case of tracking a desired trajectory. We show that the equations are consistent with those presented in [12] (although the tracking results here are in a much more intuitive form), and that in the "risk-neutral" case, the standard LQG solution results. The solution to the discrete-time hidden Markov model (HMM) risk-sensitive tracking problem is also presented. This system results in a finite-dimensional information-state, with an infinite dimensional dynamic programming task. However, it is possible to discretize that information-state space and thus obtain approximate solutions. Also discussed are methods for achieving zero steady state error for tracking with risk-sensitive control policies.

The key to the technique used in this paper is that an information- state is chosen in such a way that it represents both a state estimate and the cost incurred to the time of the estimate. A change of reference probability measure is used to arrive at a linear recursive update equation for the information-state. Then dynamic programming methods are employed to obtain the solution to the control problem, having been re-formulated in terms of the information-state. This derivation is fundamentally different to Whittle's approach [12], being more closely linked to Bensoussan and van Schuppen [4].

An important feature of this paper is that it presents a finite dimensional solution to the risk-sensitive output feedback control problem in the LQG case. It therefore provides a finite-dimensional example of the quite general infinite-dimensional controllers derived in [7], and gives insight to the nonlinear control solution. The presentation of results for tracking with hidden Markov model systems demonstrates a nonlinear situation where a finite dimensional information state can be derived. The dynamic programming solution is infinite dimensional, but can be approximated, as discussed later in Section 3.4.

Simulation studies are presented in an effort to demonstrate the effect of variations in the controller's sensitivity to risk. Various tracking problems are considered to show the advantages of the risk-sensitive approach.

2 Linear Systems

In this section we consider the risk-sensitive tracking problem for discrete-time linear systems. The case of time-invariant systems is presented, however in this finite-time framework the result is equally applicable to time-varying systems.

2.1 State space model

Consider the following discrete-time system on the probability space $(\Omega, \mathcal{F}, P)$ with complete filtration $\{\mathcal{F}_k\}$

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + v_k \\ y_{k+1} &= Cx_k + w_k \\ z_{k+1} &= Dx_k \end{aligned} \tag{2.1}$$

over the finite time interval $k = 0, 1, \ldots, T$. The state of the system is represented by the process x. The observable part of the system is represented by the process y. In this paper we will consider the problem of output tracking, and denote the desired trajectory by $\tilde{z}$. The process which is to follow $\tilde{z}$ is defined by z. The random variables v_k and w_k have normal densities $\psi \sim N(0, \Sigma)$ and $\phi \sim N(0, ,)$ respectively, where Σ and , are $n \times n$ and $p \times p$ positive definite matrices. The control, u, takes values in $\mathbb{R}^m$. The complete filtration generated by $(y_0, \ldots, y_k)$ is denoted by $\mathcal{Y}_k$, and the admissible controls u are the set of $\mathbb{R}^m$-valued $\{\mathcal{Y}_k\}$ adapted processes. We write $U_{k,l}$ for the set of such control processes defined on the interval $k, \ldots, l$.

In order to reformulate the system model (2.1), a new probability measure, $\overline{P}$, can be defined by setting

$$\Lambda_{0,k} = \left.\frac{dP}{d\overline{P}}\right|_{\mathcal{F}_k} = \prod_{\ell=1}^{k} \lambda_\ell \ , \tag{2.2}$$

where

$$\lambda_k = \frac{\psi(x_k - Ax_{k-1} - Bu_{k-1})\phi(y_k - Cx_{k-1})}{\psi(x_k)\phi(y_k)} \ . \tag{2.3}$$

Here, $\Lambda_{0,k}$ is an $\mathcal{F}_k$ martingale and $E[\Lambda_{0,k}] = 1$. Now under $\overline{P}$, x_k and y_k are two sequences of independent, normally distributed random variables with densities ψ and ϕ respectively. This reformulated model results in a linear recursion for the un-normalised information-state, as in Section 2.3.

2.2 Cost

The cost function for the risk-sensitive control problem is given, for any admissible control $u \in U_{0,T-1}$, by

$$\begin{aligned} J(u) &= E\left[\exp\theta\left\{\Psi_{0,T-1} + \frac{1}{2}x_T' M_T x_T\right\}\right] && (2.4) \\ &= \overline{E}\left[\Lambda_{0,T}\exp\theta\left\{\Psi_{0,T-1} + \frac{1}{2}x_T' M_T x_T\right\}\right] \ , && (2.5) \end{aligned}$$

where

$$\Psi_{j,k} \triangleq \sum_{\ell=j}^{k} \frac{1}{2}\left[x_\ell' M x_\ell + u_\ell' N u_\ell + (\tilde{z}_{\ell+1} - Dx_\ell)' Q (\tilde{z}_{\ell+1} - Dx_\ell)\right] \ . \tag{2.6}$$

Here, $\theta > 0$ is a real number and represents the amount of risk in the control policy. For small values of θ, approaching zero, the effect is to make control decisions assuming the stochastic disturbances are acting in an average manner. For larger values of θ, the control is effectively more conservative, or in other words, has a higher sensitivity to risk.

2.3 Information-state

In this section we present finite dimensional recursions for the *information-state* which, as the name suggests, provides information about the state of the system [10] p. 81. In the case of risk-sensitive control, it is convenient to also include a component of the cost in the information-state. For the formulation presented here, the information-state is a probability distribution (it can be compared to the 'past stress' in [12]). For small values of θ, approaching zero, the mean and variance of the information-state become the state and covariance estimates for the linear Kalman filter.

For any admissible control u, consider the measure

$$\alpha_k(x)dx \overset{\triangle}{=} \overline{E}[\Lambda_{0,k}\exp(\theta\Psi_{0,k-1})I(x_k \in dx)|\mathcal{Y}_k] \tag{2.7}$$

where $I(.)$ is the indicator function.

Lemma 2.1 *The information-state $\alpha_k(x)$ as defined in (2.7) obeys the following recursion*

$$\alpha_{k+1}(x) = \phi^{-1}(y_{k+1})\int_{\mathbb{R}^n}\phi(y_{k+1}-C\xi)\exp(\theta\Psi_{k,k})\psi(x-A\xi-Bu_k)\alpha_k(\xi)d\xi \tag{2.8}$$

Proof:

$$\begin{aligned}
\alpha_{k+1}(x)dx &= \overline{E}[\Lambda_{0,k+1}\exp(\theta\Psi_{0,k})I(x_{k+1}\in dx)|\mathcal{Y}_{k+1}] \\
&= \overline{E}[\lambda_{k+1}\Lambda_{0,k}\exp(\theta\Psi_{k,k})\exp(\theta\Psi_{0,k-1})I(x_{k+1}\in dx)|\mathcal{Y}_{k+1}] \\
\alpha_{k+1}(x) &= \int_{\mathbb{R}^n}\cdots\int_{\mathbb{R}^n}\frac{\phi(y_{k+1}-Cx_k)}{\phi(y_{k+1})}\exp(\theta\Psi_{k,k})\psi(x-Ax_k-Bu_k) \\
&\qquad \Lambda_{0,k}\exp(\theta\Psi_{0,k-1})d\overline{P}(x_0,\ldots,x_k) \\
&= \phi^{-1}(y_{k+1})\int_{\mathbb{R}^n}\phi(y_{k+1}-C\xi)\exp(\theta\Psi_{k,k}) \\
&\qquad \psi(x-A\xi-Bu_k)\alpha_k(\xi)d\xi
\end{aligned}$$

■

Theorem 2.1 *The information-state $\alpha_k(x)$ is an un-normalised Gaussian density given by*

$$\alpha_k(x) = \alpha_k(x,\chi_k) = Z_k\exp(-1/2)[(x-\mu_k)'R_k^{-1}(x-\mu_k)] \tag{2.9}$$

where $\chi_k = (\mu_k, R_k, Z_k)$, and μ_k, R_k^{-1} and Z_k are given by the following algebraic recursions

$$\boxed{\begin{aligned}
\mu_{k+1} &= R_{k+1}\big[\Sigma^{-1}Bu_k \\
&\quad +\Sigma^{-1}Aa_k^{-1}(R_k^{-1}\mu_k - A'\Sigma^{-1}Bu_k + C',\ ^{-1}y_{k+1} - \theta D'Q\tilde{z}_{k+1})\big] \\
R_{k+1}^{-1} &= \Sigma^{-1} - \Sigma^{-1}Aa_k^{-1}A'\Sigma^{-1} \\
Z_{k+1} &= Z_k|\Sigma|^{-\frac{1}{2}}|a_k|^{-\frac{1}{2}}\exp\left(-\tfrac{1}{2}\right)\left[\gamma_k - \mu_{k+1}'R_{k+1}^{-1}\mu_{k+1}\right]
\end{aligned}} \tag{2.10}$$

where

$$a_k = C',\ ^{-1}C - \theta(M + D'QD) + A'\Sigma^{-1}A + R_k^{-1} \tag{2.11}$$

$$\begin{aligned}
\gamma_k &= u_k'(-\theta N + B'\Sigma^{-1}B)u_k + \mu_k'R_k^{-1}\mu_k - \theta\tilde{z}_{k+1}'Q\tilde{z}_{k+1} \\
&\qquad -(\mu_k'R_k^{-1} - u_k'B'\Sigma^{-1}A + y_{k+1}',\ ^{-1}C - \theta\tilde{z}_{k+1}'QD)a_k^{-1} \\
&\qquad (R_k^{-1}\mu_k - A'\Sigma^{-1}Bu_k + C',\ ^{-1}y_{k+1} - \theta D'Q\tilde{z}_{k+1})
\end{aligned} \tag{2.12}$$

under the condition that a_k and R_k be positive definite for all k.

Proof: Due to the linearity of the dynamics, and the fact that v_k and w_k are independent and normally distributed, we know that $\alpha_k(x)$ is an un-normalised Gaussian density. The recursions for μ_k, R_k^{-1} and Z_k are obtained by evaluating the integral in (2.8). The details are omitted. ■

Further matrix manipulations yield the following, more familiar, expressions

$$\boxed{\begin{array}{rcl} \mu_{k+1} & = & A\mu_k + Bu_k + A\tilde{K}_k \left[C' , ^{-1}(y_{k+1} - C\mu_k \right. \\ & & \qquad \left. -\theta , (C')^{-1}D'Q\tilde{z}_{k+1}) + \theta(M + D'QD)\mu_k\right] \\ \tilde{K}_k & \triangleq & (R_k^{-1} + C' , ^{-1}C - \theta(M + D'QD))^{-1} \\ R_{k+1} & = & \Sigma + A\tilde{K}_k A' \end{array}} \tag{2.13}$$

which can be compared to the result presented in [12] for the case where $Q = 0$. When $Q \neq 0$ (that is, for tracking), the equations here are in a much more intuitive form than the tracking result in [12].

2.3.1 Limit result

Equations (2.13) can be re-expressed in the following form

$$\boxed{\begin{array}{rcl} \mu_{k+1} & = & A\mu_{k|k} + Bu_k \\ \mu_{k|k} & \triangleq & \mu_k + K_k[y_{k+1} - C\mu_k \\ & & \qquad -\theta(, (C')^{-1}D'Q\tilde{z}_{k+1} - , (C')^{-1}(M + D'QD)\mu_k)] \\ K_k & \triangleq & (R_k^{-1} - \theta(M + D'QD))^{-1}C' \\ & & \qquad\qquad [C(R_k^{-1} - \theta(M + D'QD))^{-1}C' + ,]^{-1} \\ R_{k+1} & = & \Sigma + AR_{k|k}A' \\ R_{k|k} & \triangleq & R_k - K_k C R_k \end{array}} \tag{2.14}$$

In the case when θ approaches zero, it can easily be seen that the equations in (2.14) reduce to the standard Kalman filter equations [2] p.40.

2.4 Alternate cost representation

In this section we show that the cost function can be expressed in terms of the information-state. This allows the optimisation problem to be solved by dynamic programming, without any appeal to a certainty equivalence principle.

Theorem 2.2 *For any admissible control u, the risk sensitive cost can be expressed in the form*

$$J(u) = \overline{E}\left[\langle \alpha_T(\cdot, \chi_T), \beta_T \rangle\right] \tag{2.15}$$

where $\langle f(\cdot), q(\cdot) \rangle = \int_{\mathbb{R}^n} f(z)q(z)dz$ *and* $\beta_T(x) \triangleq \exp(\frac{\theta}{2}x'M_T x)$.

Proof: We have from (2.5) that

$$\begin{aligned} J(u) &= \overline{E}\left[\Lambda_{0,T}\exp(\theta\Psi_{0,T-1})\exp(\tfrac{\theta}{2}x'_T M_T x_T)\right] \\ &= \overline{E}\left[\overline{E}\left[\Lambda_{0,T}\exp(\theta\Psi_{0,T-1})\beta_T(x_T)|\mathcal{Y}_T\right]\right] \\ &= \overline{E}\left[\int_{\mathbb{R}^n}\beta_T(x)\alpha_T(x)dx\right] \\ &= \overline{E}\left[\langle\alpha_T(\cdot,\chi_T),\beta_T\rangle\right] \ . \end{aligned}$$

∎

2.5 Dynamic programming

Following [7] we know that the alternative control problem can be solved using dynamic programming. Suppose that at some time k, $0 < k < T$, the information-state χ_k is $\chi = (\mu, R, Z)$.

The value function for this control problem is [1, 7]

$$V(\chi,k) = \inf_{u\in U_{k,T-1}} \overline{E}[\langle\alpha_k,\beta_k\rangle \mid \alpha_k = \alpha(\chi)] \ , \tag{2.16}$$

where β_k is an adjoint process defined by

$$\beta_k(x) = \overline{E}[\Lambda_{k+1,T}\exp(\theta\Psi_{k,T-1})\exp(\frac{\theta}{2}x'_T M_T x_T)|x_k = x, \mathcal{Y}_T] \ . \tag{2.17}$$

The adjoint process can be compared to the 'future stress' in [12].

Theorem 2.3 *[1, 7] The value function satisfies the recursion*

$$V(\chi,k) = \inf_{u\in U_{k,k}} \overline{E}[V(\chi_{k+1}(\chi_k,u,y_{k+1}),k+1)|\chi_k = \chi] \tag{2.18}$$

and $V(\chi,T) = \langle\alpha_T(.,\chi),\beta_T\rangle$.

∎

2.6 Dynamic programming solution

Theorem 2.4 *The value function is the exponential of a quadratic in* μ

$$V(\chi,k) = Z_k\exp(\theta/2)[\mu'_k S^a_k\mu_k + 2S^{b\prime}_k\mu_k + S^c_k] \ , \tag{2.19}$$

and the optimal control is linear in μ

$$\boxed{u^{min}_k = -(N + B'\tilde{S}_{k+1}B)^{-1}B'[\tilde{S}_{k+1}\tilde{A}\mu_k + S^b_{k+1} + \theta K^b_k]} \ , \tag{2.20}$$

where

$$\begin{aligned} \tilde{S}_{k+1} &= ((S^a_{k+1})^{-1} - \theta\tilde{\Gamma}_k\tilde{\Gamma}'_k)^{-1} \ , & \tilde{A} &= A\rho^{-1} \\ K^a_k &= (N + B'\tilde{S}_{k+1}B)^{-1}B'\tilde{S}_{k+1}\tilde{A} \ , & \tilde{M} &= (M + D'QD)\rho^{-1} \\ K^b_k &= S^a_{k+1}\tilde{\Gamma}_k\delta^{-1}(\tilde{\Gamma}'_k S^b_{k+1} - \Theta D'Q\tilde{z}_{k+1}) \ , & \sigma &= N + B'\tilde{S}_{k+1}B \\ \tilde{\Gamma}_k &= A\tilde{K}_k C'\Gamma^{-1}C\Theta \ , & \delta &= I - \theta\tilde{\Gamma}'_k S^a_{k+1}\tilde{\Gamma}_k \\ \Theta &= [(C'\Gamma C)^{-1} + \rho^{-1}R_k]^{1/2} \ , & \rho &= I - \theta R_k(M + D'QD) \end{aligned}$$

Also S_k^a and S_k^b are given by the following recursions

$$\boxed{\begin{array}{rcl} S_k^a & = & \tilde{M} + \tilde{A}' S_{k+1}^a (I + BN^{-1}B'S_{k+1}^a - \theta \tilde{\Gamma}_k \tilde{\Gamma}_k' S_{k+1}^a)^{-1} \tilde{A} \\ S_k^b & = & \tilde{A}' S_{k+1}^b - (I + \theta \tilde{M} R_k) D' Q \tilde{z}_{k+1} \\ & & \qquad + \tilde{A}' \tilde{S}_{k+1} B \sigma^{-1} B' (S_{k+1}^b + \theta K_k^b) + \theta \tilde{A}' K_k^b \end{array}} \tag{2.21}$$

under the condition that $(I - \theta \tilde{\Gamma}_k' S_{k+1}^a \tilde{\Gamma}_k)$ is positive definite for all k, and C is positive definite except in the cases where $C = D$ or $D = 0$.

Proof: By evaluating the dynamic programming equation (2.18) for $V(\chi, T-1)$ it can be seen that the value function is the exponential of a quadratic in μ. The remainder of the proof is too long for presentation in this paper, but is essentially an evaluation of the dynamic programming equation (2.18), with appropriate variable transformations. ■

Remark: The condition that C be positive definite, is a manifestation of the variable transformation used in order to present the results in a form which more readily demonstrates the link to standard LQG results. As can be seen from the exclusion when $D = 0$, the condition only applies to the tracking part of the solution, (i.e., S_k^b and K_k^b recursions). It is possible to solve the dynamic programming problem without such a variable transformation and thus remove the condition on C. □

In order to demonstrate consistency with the results presented in [12], where an appeal was made to a certainty equivalence principle, and $Q \equiv 0$, we now set

$$\Pi_k = S_k^a [I + \theta R_k S_k^a]^{-1} \tag{2.22}$$

and $Q = 0$, which results in the following recursion for Π_k

$$\boxed{\Pi_k = M + A'[\Pi_{k+1}^{-1} + BN^{-1}B' - \theta \Sigma]^{-1} A} \tag{2.23}$$

under the condition that $(I - \theta R_k \Pi_k)$ is positive definite for all k.

Substitution of (2.23) into (2.20), yields

$$\boxed{u_k^{min} = -N^{-1}B'(\Pi_{k+1}^{-1} + BN^{-1}B' - \theta\Sigma)^{-1} A [I - \theta R_k \Pi_k]^{-1} \mu_k} \tag{2.24}$$

where the term $[I - \theta R_k \Pi_k]^{-1} \mu_k$ is sometimes referred to as the minimum stress estimate.

2.6.1 Limit result

In the case where θ approaches zero, it can easily be seen that $\tilde{S}_{k+1}$, $\tilde{A}$ and $\tilde{M}$ approach S_{k+1}^a, A and $M + D'QD$ respectively, and the following equations result from manipulations to (2.20) and (2.21)

$$\boxed{\begin{array}{rcl} u_k^{min} & = & -(N + B'S^a_{k+1}B)^{-1}B'[S^a_{k+1}A\mu_k + S^b_{k+1}] \\ S^a_k & = & M + D'QD + A'[S^a_{k+1} - S^a_{k+1}B(N + B'S^a_{k+1}B)^{-1}B'S^a_{k+1}]A \\ S^b_k & = & (A - B(N + B'S^a_{k+1}B)^{-1}B'S^a_{k+1}A)'S^b_{k+1} - D'Q\tilde{z}_{k+1} \end{array}} \tag{2.25}$$

These are the standard LQG equations, as presented for example in ([3], pages 32 and 81.)

3 Hidden Markov Models

In this section we present the risk-sensitive tracking result for Hidden Markov Models. Such systems are discrete time and have finite-discrete states. By *finite-discrete* we mean they have a finite number of discrete states. We will consider the case of continuous valued observations.

3.1 State space model

Let X_k be a discrete-time homogeneous, first order Markov process belonging to a finite-discrete set. The state space of X, without loss of generality, can be identified with the set of unit vectors $\mathbf{S} = \{e_1, e_2, ..., e_n\}$, $e_i = (0, ..., 0, 1, 0, ..., 0)' \in \mathbb{R}^n$ with 1 in the i^{th} position [11]. We consider that the process is defined on the probability space $(\Omega, \mathcal{F}, P)$ with complete filtration $\{\mathcal{F}_k\}$. Suppose there is a family of generators $\mathbf{A}(u) = (a_{ij}(u)), 1 \leq i, j \leq n$ where $a_{ij}(u) = P(X_{k+1} = e_j \mid X_k = e_i, u)$ so that $E[X_{k+1} \mid X_k, u] = \mathbf{A}'(u)X_k$. These generators depend on the admissible controls, u. Of course $a_{ij}(u) \geq 0, \sum_{j=1}^n a_{ij}(u) = 1$, for each i. In this paper we consider the case of continuous valued observations y_k, and desired trajectories $\tilde{z}_k$. The state space model for the HMM is given by

$$\begin{array}{rcl} X_{k+1} & = & \mathbf{A}'(u)\,X_k + m_{k+1} \\ y_k & = & c(X_k) + w_k \\ z_k & = & d(X_k) \end{array} \tag{3.26}$$

where m_{k+1} is a $(\mathbf{A}(u), \mathcal{F}_k)$ martingale increment, in that $E[m_{k+1} \mid \mathcal{F}_k] = 0$. Also, the random variable w_k has normal density $\phi \sim N(0, ,)$, where , is a $p \times p$ positive definite matrix.

In order to reformulate the system model (3.26), a new probability measure, $\overline{P}$, can be defined by setting

$$\Lambda_{0,k} = \left.\frac{dP}{d\overline{P}}\right|_{\mathcal{F}_k} = \prod_{\ell=1}^{k} \lambda_\ell\ , \quad \text{where} \quad \lambda_k = \frac{\phi(y_k - c(X_k))}{\phi(y_k)}\ . \tag{3.27}$$

Here, $\Lambda_{0,k}$ is an $\mathcal{F}_k$ martingale and $E[\Lambda_{0,k}] = 1$. Now under $\overline{P}$, y_k is a sequence of independent, normally distributed random variables with density ϕ. This reformulated model results in a linear recursion for the un-normalised information-state, as in Section 3.3.

3.2 Cost

The cost function for the risk-sensitive control problem for HMMs is given, for any admissible control $u \in U_{0,T-1}$, by

$$J(u) = E\left[\exp\theta\left\{\Psi_{0,T-1} + \frac{1}{2}X_T' M_T X_T\right\}\right] \tag{3.28}$$

$$= \overline{E}\left[\Lambda_{0,T}\exp\theta\left\{\Psi_{0,T-1} + \frac{1}{2}X_T' M_T X_T\right\}\right] , \tag{3.29}$$

where

$$\Psi_{j,k} \triangleq \sum_{\ell=j}^{k}\frac{1}{2}\left[X_\ell' M X_\ell + u_\ell' N u_\ell + (\tilde{z}_\ell - d(X_\ell))' Q(\tilde{z}_\ell - d(X_\ell))\right] . \tag{3.30}$$

3.3 Information-state

As in Section 2.3, we again present an information-state which includes a component of the cost. Unlike the linear case, however, for HMMs the information-state is a finite-dimensional probability distribution vector.

For any admissible control u, consider the measure

$$\alpha_k(e_i) \triangleq \overline{E}[\Lambda_{0,k}\exp(\theta\Psi_{0,k-1})\langle X_k, e_i\rangle|\mathcal{Y}_k]. \tag{3.31}$$

Theorem 3.1 *The information-state $\alpha_k = (\alpha_k(e_1), \ldots, \alpha_k(e_n))'$, as defined in (3.31), obeys the following recursion*

$$\boxed{\alpha_{k+1} = \mathcal{B}_k \mathbf{A}'(u)\mathcal{D}_k\alpha_k} \tag{3.32}$$

where

$$\mathcal{B}_k = \operatorname{diag}\left(\frac{\psi(y_{k+1} - c(e_1))}{\psi(y_{k+1})}, \ldots, \frac{\psi(y_{k+1} - c(e_n))}{\psi(y_{k+1})}\right) \tag{3.33}$$

$$\mathcal{D}_k = \operatorname{diag}\left(\exp\frac{\theta}{2}[e_1' M e_1 + u_k' N u_k + (\tilde{z}_k - d(e_1))' Q(\tilde{z}_k - d(e_1))], \ldots, \exp\frac{\theta}{2}[e_n' M e_n + u_k' N u_k + (\tilde{z}_k - d(e_n))' Q(\tilde{z}_k - d(e_n))]\right) \tag{3.34}$$

Proof:

$$\begin{aligned}
\alpha_{k+1}(e_i) &= \overline{E}[\Lambda_{0,k+1}\exp(\theta\Psi_{0,k})\langle X_{k+1}, e_i\rangle|\mathcal{Y}_{k+1}] \\
&= \overline{E}[\lambda_{k+1}\Lambda_{0,k}\exp(\theta\Psi_{k,k})\exp(\theta\Psi_{0,k-1})X_k'\mathbf{A}(u)e_i|\mathcal{Y}_{k+1}] \\
&= \overline{E}\Big[\tfrac{\psi(y_{k+1}-c(e_i))}{\psi(y_{k+1})} \\
&\quad \Big(\textstyle\sum_{j=1}^n\langle X_k, e_j\rangle \exp\tfrac{\theta}{2}[e_j'Me_j + u_k'Nu_k + (\tilde{z}_k - d(e_j))'Q(\tilde{z}_k - d(e_j))]\Big) \\
&\quad \Big(\textstyle\sum_{j=1}^n a_{ji}(u)\langle X_k, e_j\rangle\Big)\Lambda_{0,k}\exp(\theta\Psi_{0,k-1})|\mathcal{Y}_{k+1}\Big] \\
&= \tfrac{\psi(y_{k+1}-c(e_i))}{\psi(y_{k+1})}\textstyle\sum_{j=1}^n \exp\tfrac{\theta}{2}[e_j'Me_j + u_k'Nu_k + (\tilde{z}_k - d(e_j))'Q(\tilde{z}_k - d(e_j))] \\
&\quad a_{ji}(u)\alpha_k(e_j)
\end{aligned}$$

Writing this in matrix notation gives the result. ■

3.4 Alternate cost and dynamic programming

For the HMM case, the cost (3.29) can be expressed in a separated form, as in Theorem 2.2, with appropriate notational changes. The dynamic programming solution is likewise obtained from Theorem 2.3. Unfortunately in this case the solution to the dynamic programming equation is not able to be evaluated in terms of Riccati equations, as in Section 2.6. The solution for the HMM system requires a search over all possible control values for each backwards step and for each possible value of the information-state. Therefore the HMM case results in a finite dimensional information-state, but unfortunately has an infinite dimensional solution to the dynamic programming problem. It is possible however to make practical approximations by quantising the information-state space and solving the approximate dynamic programming problem. This can be computationally feasible in some cases since the information state is known to have positive elements. Also, a normalised version of the information-state can be used in the Dynamic programming problem since the following property is known to hold in the HMM case [8]

$$V(c\alpha, k) = cV(\alpha, k). \tag{3.35}$$

In Section 5 we present an example of such an approximate dynamic programming solution for this risk-sensitive HMM case.

4 Constant Reference Input Case

In this section we discuss the case where $\tilde{z}_k$ is a constant value. Under such conditions it is possible to design an optimal controller with zero steady state error. Here we consider the discrete-time linear system of section 2.

Consider the cost function given in (2.5). We note that for this general function there exist some trade-offs which do not allow zero steady state error to be achieved. For example, (2.6) penalises deviations of x_k from zero, while at the same time penalising deviations of Dx_k from $\tilde{z}_k$, these are conflicting objectives. Also, the control u_k is penalised for deviations from zero when we know that, in steady state, it must be a constant non-zero value for this constant reference input case. These considerations indicate that the tracking problem must be reformulated.

4.1 Control integrator approach

A standard method for obtaining zero steady state error, is to introduce an integrator in the forward path of the control loop. This technique can be used in this risk-sensitive case with a few minor adjustments to the control policy. Figure 1 shows the block diagram for the control system presented in the preceding sections of this paper. By introducing an integrator and augmenting the state, as in Figure 2, it is possible to obtain a more appropriate cost function.

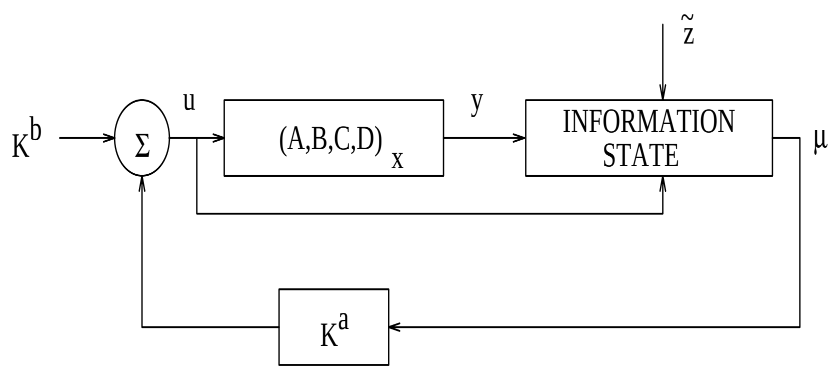

Figure 1: Block diagram for standard control policy

In this section we choose the control u_k to be an extra state, and define a new control $\tilde{u}_k$. The state of the augmented system is then given by

$$\tilde{x}_k = \begin{pmatrix} x_k \\ x_k^c \end{pmatrix} \tag{4.36}$$

where $x_k^c = u_k = \sum_{i=1}^{k} \tilde{u}_k$. This augmented state is an un-normalised Gaussian density, and is given from equations (2.10) to (2.12), with

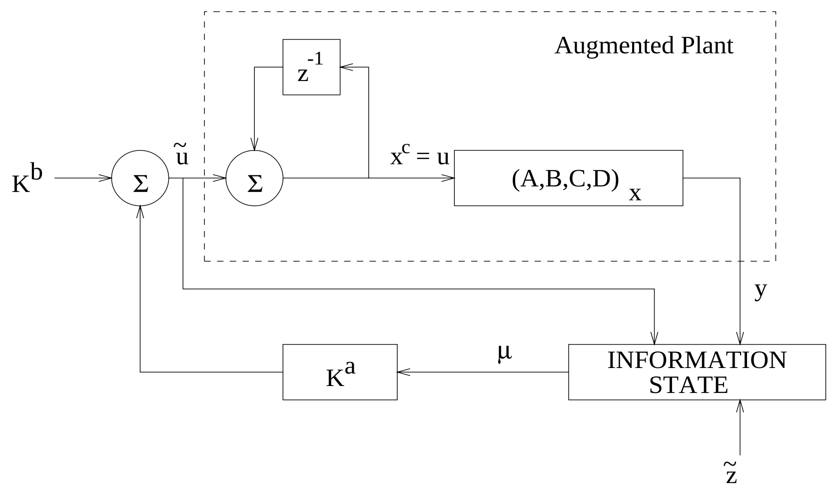

Figure 2: Block diagram for constant reference input case

appropriate re-definitions for the augmented system, by

$$\begin{aligned} \tilde{x}_{k+1} &= \begin{pmatrix} A & B \\ 0 & I \end{pmatrix} \tilde{x}_k + \begin{pmatrix} 0 \\ I \end{pmatrix} \tilde{u}_k + \begin{pmatrix} I \\ 0 \end{pmatrix} v_k \\ y_{k+1} &= \begin{pmatrix} C & 0 \end{pmatrix} \tilde{x}_k + w_k \\ z_{k+1} &= \begin{pmatrix} D & 0 \end{pmatrix} \tilde{x}_k \end{aligned} \tag{4.37}$$

The cost function to be considered is now given, for $\tilde{z}_k = \tilde{z}$, by (2.5) where $\Psi_{j,k}$ is re-defined as

$$\Psi_{j,k} \triangleq \sum_{\ell=j}^{k} \frac{1}{2} \left[\tilde{u}'_\ell N \tilde{u}_\ell + (\tilde{z} - (D \;\; 0)\tilde{x}_k)' Q (\tilde{z} - (D \;\; 0)\tilde{x}_k) \right] . \tag{4.38}$$

It can easily be seen that for this cost function there are no conflicting objectives, and as such zero steady state error can be achieved.

Unfortunately, however, there exist some hidden problems. The first is that the new state $\tilde{x}_k$ has zero state noise and as such results in a singular filtering problem. This can be overcome by assuming there exists some noise of variance ϵ and then taking the limit of the information-state, as ϵ approaches zero. As can be seen from (2.13), the limit exits with $\tilde{K}_k$ re-defined as

$$\tilde{K}_k \triangleq R_k (I + C', {}^{-1} C R_k - \theta (M + D'QD) R_k)^{-1}. \tag{4.39}$$

The second problem is that in the case of modelling errors, even with the augmented system, zero steady state error is not necessarily achieved. This is due to the fact that there exists a term in the optimal control law (2.20), which is not proportional to the state estimate μ_k. If this term is not calculated correctly, as would be the case with modelling errors, then the control $\tilde{u}_k$ would drive the output z_k to an incorrect steady state value.

Although zero steady state error may not be achieved in certain cases, there is still an advantage to applying the integrator approach. In the risk-sensitive case, when modelling errors are present, it is possible to achieve a lower minimum variance cost than for the case of LQG control (as can be seen in Section 5). One problem however is that the step response can be undesirable for values of θ which are too large. By augmenting the system with an integrator, the step response will be smoothed out, resulting in a risk-sensitive policy which has both a lower minimum variance cost and an acceptable step response.

4.2 Reference model integrator approach

An extra point to note is that in the scheme presented so far, it is necessary to have prior knowledge of the constant reference input signal, $\tilde{z}$. An approach for removing this assumption, commonly used in LQG tracking systems, is to model the reference $\tilde{z}$ by a first order integrator,

$$\begin{aligned} x^r_{k+1} &= x^r_k + v^r_k \\ \tilde{z}_{k+1} &= \tilde{D} x^r_k \end{aligned} \quad . \tag{4.40}$$

This would of course slow the response of the system, but would have the advantage of zero-steady state error in conditions of uncertain models. The augmented state vector is given by

$$\tilde{x}_k = \begin{pmatrix} x_k \\ x^c_k \\ x^r_k \end{pmatrix} . \tag{4.41}$$

This new augmented state is again an un-normalised Gaussian density, and is given from equations (2.10) to (2.12), with appropriate re-definitions for the augmented system, by

$$\begin{aligned} \tilde{x}_{k+1} &= \begin{pmatrix} A & B & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{pmatrix} \tilde{x}_k + \begin{pmatrix} 0 \\ I \\ 0 \end{pmatrix} \tilde{u}_k + \begin{pmatrix} v_k \\ 0 \\ v^r_k \end{pmatrix} \\ y_{k+1} &= \begin{pmatrix} C & 0 & -\tilde{D} \end{pmatrix} \tilde{x}_k + w_k \\ z_{k+1} &= \begin{pmatrix} D & 0 & -\tilde{D} \end{pmatrix} \tilde{x}_k \end{aligned} \tag{4.42}$$

The cost function to be considered is given by (2.5) where $\Psi_{j,k}$ is re-defined as

$$\Psi_{j,k} \triangleq \sum_{\ell=j}^{k} \frac{1}{2} \left[\tilde{u}'_\ell N \tilde{u}_\ell + \tilde{x}'_k \begin{pmatrix} D & 0 & -\tilde{D} \end{pmatrix}' Q \begin{pmatrix} D & 0 & -\tilde{D} \end{pmatrix} \tilde{x}_k \right] . \tag{4.43}$$

In this case there are no terms in (4.43) which are linear in $\tilde{x}_k$, and as such the optimal control will be proportional to the state estimate μ_k and have no extra terms (i.e., S_k^b and K_k^b will not appear). In fact the solution to the dynamic programming problem for this augmented system is given in Theorem 2.4 with the following substitutions (in addition to those for the augmented system representation (4.42))

$$\tilde{M} = \begin{pmatrix} D & 0 & -\tilde{D} \end{pmatrix}' Q \begin{pmatrix} D & 0 & -\tilde{D} \end{pmatrix} \ , \ \tilde{Q} = 0 \tag{4.44}$$

Due to the purely proportional feedback nature of this solution, it can now be seen that it is possible to obtain zero steady state error even in the case of modelling errors, as there is no longer a constant offset term contributed by K_k^b. Unfortunately, however, the initial transient will suffer due to the fact that the controller is no longer able to anticipate the step in the reference input, as it is now assumed to be unknown.

One final point to note is that this second augmentation can be used without the first augmentation, and zero steady state error will result for the case where $N = 0$, (this is termed *cheap control*). However, undesirable oscillations in the transient response will increase, compared to the situation where an integrator is present in the forward path.

5 Simulation Studies

We now present simulation studies to demonstrate the effect of variations to the risk-sensitive parameter θ.

One practical motivation for a risk-sensitive approach is when high order systems are to be optimally controlled by standard PID controllers. In this case, higher order controllers can be represented by low order controllers (by appropriate model reduction) while the risk-sensitive parameter can be used to, in effect, robustify the model order reduction. It is partly for this reason that we concentrate our simulation studies on second order systems. However, higher order systems are straightforward extensions in this vector framework.

An important point to mention is that classical controllers can be designed for the control problem considered in the first three examples, however the classical techniques do not perform an optimisation of any criterion (as such control energy can be restrictively high), and they rely to a large extent on the designer placing poles and zeros in 'good' locations, often via iterative techniques, rather than the optimal LQG and risk-sensitive policies which are entirely structured, once the cost criterion is specified. Therefore we focus our attention, in this section, on comparing the standard LQG results with those of the risk-sensitive tracker developed in this paper.

In Example 4, a two-input two-output system is considered, for which a classical controller is not easily derived. In contrast, the risk-sensitive controller for such a system is generated by a straight forward extension of the single-input single-output case considered in the first three examples, and therefore demonstrates another of the features of this risk-sensitive approach, namely expandability to higher order systems.

Example 1: In this example we demonstrate a case where modelling errors are present. The true system is given by the following parameters

$$A = \begin{bmatrix} -0.2 & 1 \\ -0.2 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0.9 \\ -0.6 \end{bmatrix} \quad \begin{matrix} C = [1\ 0] \\ D = [1\ 0] \end{matrix} \quad M = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$T = 100 \quad \Sigma = \begin{bmatrix} 0.01 & 0 \\ 0 & 0.01 \end{bmatrix} \quad \begin{matrix} N = 0.1 \\ Q = 100 \end{matrix} \quad , = 0.01$$

and the trajectory to be followed, $\tilde{z}_k$, is a unit step at $k = 20$. The modelling error is introduced by assuming in the design that A is given by

$$A = \begin{bmatrix} -0.8 & 1 \\ -0.8 & 0 \end{bmatrix}.$$

Table 1 gives values of the LQG, minimum variance, cost function (i.e., $\Psi_{0,T-1} + 0.5x'_T M_T x_T$) averaged over 100 simulation runs. It can be seen that in the case where no modelling error is present, of course $\theta = 0$ gives lowest cost. However, when the error is introduced, a higher value of θ gives a lower minimum variance cost. This example displays an advantage of the risk-sensitive approach in the presence of modelling errors.

Unfortunately, the sample path properties may not improve with a lower minimum variance cost, as one would wish, especially if θ is too large. Here, *too large* will depend on the type of modelling error, and will of course be unknown to the designer. Figure 3 shows a typical sample run for the case of no modelling errors. It shows that the cost function chosen for the tracking task considered, results in little difference in tracking errors between the LQG and risk-sensitive policies. Figure 4 shows a typical sample run for the case where modelling errors are present. Even though the minimum variance cost is lower for the risk-sensitive policy, the tracking performance might not be as desirable, having much greater oscillations in the transient response. The conservative risk-sensitive controller places more emphasis on correcting for large errors, therefore the initial response is faster, resulting in more oscillations. It is thereby making a trade-off between a lower variance cost, and adding more oscillations. Therefore the desirability of a risk-sensitive approach cannot be measured purely by the minimum variance cost.

Example 2: In this example we demonstrate the case of a constant reference input, where an integrator is added in the forward path of the control design. The system is the same as in Example 1, but with $M = \underline{0}_{3\times 3}$. When no modelling errors are present, zero steady state is achieved. When errors are introduced to the model it is not possible to have zero steady state error, however Figures 5 and 6 demonstrate that there are still advantages to the integrator approach. The modelling error in these figures is the same as that in Example 1. As can be seen from the figures, the addition of an integrator effectively increases the usable range of risk-sensitive parameter values, θ, by smoothing the step response.

Example 3: In this example we demonstrate the case of a constant reference input, where an integrator is added in the forward path of the control design as well as using a model for the reference. As can be seen from Figure 7 zero steady state error is achieved for both modelling errors and no modelling errors, however it is at the expense of the speed of transient response, (especially in the LQG case). The other point to note from this example is that the benefit from a risk-sensitive control policy is reduced when integrators are added. This is due to the fact that integrators have the effect of adding robustness to the controller, and as such the robustness gain from the risk-sensitive policy is seen to reduce. This can also be seen by the fact that the LQG cost function is much smaller when zero steady state error is achieved, compared to when it is not achieved. The result is that the effect of varying θ is less, as the variation is over a less steep region of the exponential curve.

One final point to note is that the LQG design is not optimal for the doubly augmented system, due to the fact that the reference is modelled by an integrator (4.40) when in fact it is a deterministic signal. This can be considered to be a modelling error in the design, and as such the LQG solution is not optimal with this augmentation. In fact, for the example considered, the cost is actually less for the risk-sensitive solution than for the LQG solution, even in the case where A is known precisely (i.e., the line labelled 'no modelling error' in Figure 7).

Example 4: In this simulation study we present an example of a two-input two-output system. The state variables, A, M, T, and Σ are the same as in Example 1. The remainder of the parameters are :

$$B = \begin{bmatrix} 0.9 & 0.8 \\ -0.6 & -0.2 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad D = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$, = \begin{bmatrix} 0.01 & 0 \\ 0 & 0.01 \end{bmatrix} \quad N = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix} \quad Q = \begin{bmatrix} 100 & 0 \\ 0 & 100 \end{bmatrix}$$

and the desired trajectory is a unit step in the first output and a zero response in the second output.

The results are presented in Figure 8, which again demonstrate the effect of varying the risk-sensitive parameter. These results are for the doubly augmented system, as in Example 3. It can be seen again, that the more conservative controller, corresponding to $\theta = 0.07$ initially makes more of an effort to follow the step-jump, and to achieve a lower cost, however it also results in more oscillations. The more risky controller, corresponding to $\theta = -0.07$, in effect starts by assuming the step-jump is just extra noise, and risks ignoring it. The overall cost is higher, but the response is smoother. Obviously a trade-off exists, and this is where the risk-sensitive controller, with its variable risk-sensitive parameter, becomes extremely useful.

Example 5: In this simulation study we present an example of an approximate solution to the risk-sensitive HMM control problem of Section 3. The system is given by the following parameters

$$A = \begin{bmatrix} 0.1 & 0.9 \\ 0.9 & 0.1 \end{bmatrix} + \begin{bmatrix} u & -u \\ -u & u \end{bmatrix} \qquad \begin{array}{l} c(X_k) = \langle \begin{bmatrix} 0 \\ 1 \end{bmatrix}, X_k \rangle \\ d(X_k) = \langle \begin{bmatrix} 0 \\ 1 \end{bmatrix}, X_k \rangle \end{array}$$

$$M = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \qquad \begin{array}{c} N = 0.1 \\ Q = 0 \end{array} \qquad \begin{array}{c} , = 2 \\ T = 100 \end{array}$$

and in this example there is no trajectory to be followed, (i.e., $\tilde{z}_k = 0$). We see from this system that in an uncontrolled situation, the output trajectory, z_k, will tend to oscillate between the values 0 and 1 at each discrete time instant. From the definition of M we see that the control objective is to force z_k to the value zero.

In this example the dynamic programming problem is solved by quantizing the normalized information-state, α_k^q, into six discrete values,

$$\alpha_k^q \in \{(\ell \times 0.2, m \times 0.2)'\} \qquad 0 \leq \ell, m \leq 5 \qquad \ell + m = 5 \tag{5.45}$$

and allowing only three possible control values

$$u_k \in \{0.1, 0.3, 0.5\}. \tag{5.46}$$

The dynamic programming problem is then solved by evaluating the cost which minimizes the value function for each possible information-state, α_k^q, at each step backwards in time, k.

In order to demonstrate the effect of the risk-sensitive parameter on the control policy we present, in Table 2, the steady state control values which result from the approximate solution to the dynamic programming problem. It can be seen that as the risk-sensitive parameter increases, the information-state must be increasingly more confident of the true state, before the controller is willing to apply a large control value. This example therefore demonstrates the robustness property gained from increasing the sensitivity to risk.

6 Conclusion

In this paper we have presented the solution to the linear risk-sensitive quadratic Gaussian control problem. Results have been derived for the case of tracking a desired trajectory. The solution to the dynamic programming problem has been achieved without the need to appeal to a certainty equivalence principle, and hence gives insight to the solution for nonlinear systems. Limit results have also been presented which demonstrate the link to standard linear quadratic Gaussian control. Also, the solution to the problem of risk-sensitive tracking for hidden Markov models has been presented, as well as a discussion on achieving zero steady state error with risk-sensitive control policies. Simulation studies were presented in order to show some advantages of the risk-sensitive approach.

$\times 10^2$	$\theta = 0$ (LQG)	$\theta = 0.1$	$\theta = 0.15$
No model error	4.714	4.715	4.716
With model error	9.363	6.076	6.593

Table 1: Error analysis for risk-sensitive control

		α^q					
		$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix}$	$\begin{pmatrix} 0.4 \\ 0.6 \end{pmatrix}$	$\begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix}$	$\begin{pmatrix} 0.8 \\ 0.2 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$
θ	0.1	0.1	0.1	0.1	0.3/0.5*	0.5	0.5
	1	0.1	0.1	0.1	0.1	0.5	0.5
	10	0.1	0.1	0.1	0.1	0.1	0.5

* control oscillates between the two values

Table 2: Risk-sensitive HMM control values

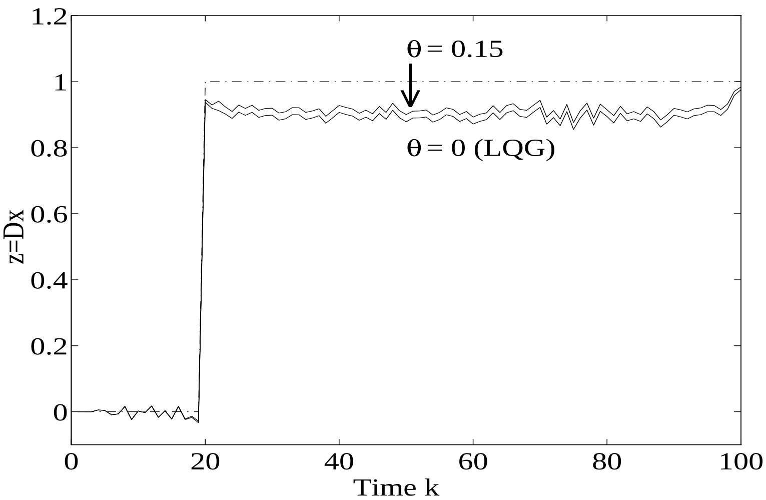

Figure 3: No modelling errors

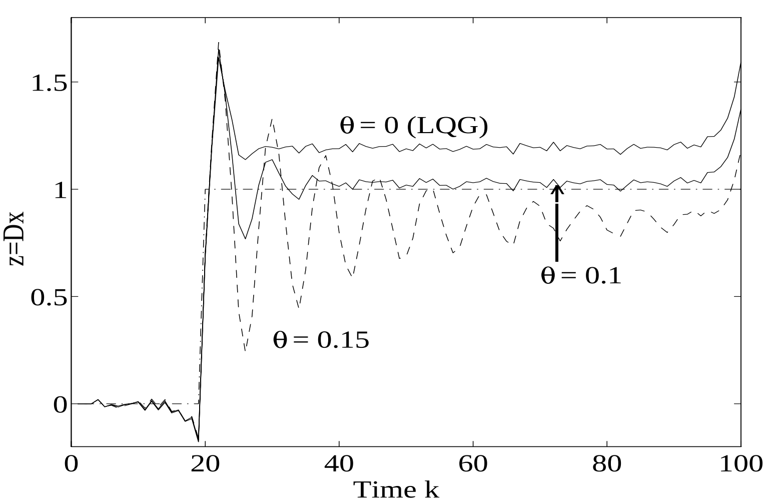

Figure 4: With modelling errors

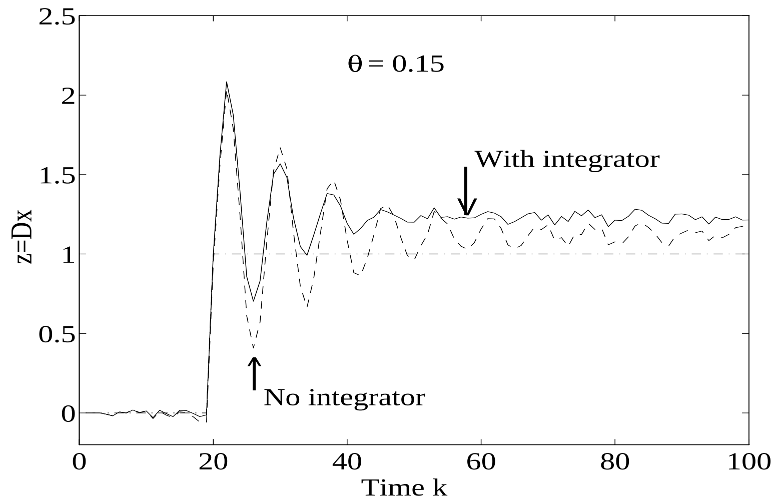

Figure 5: Augmented system with modelling errors, $\theta = 0.15$

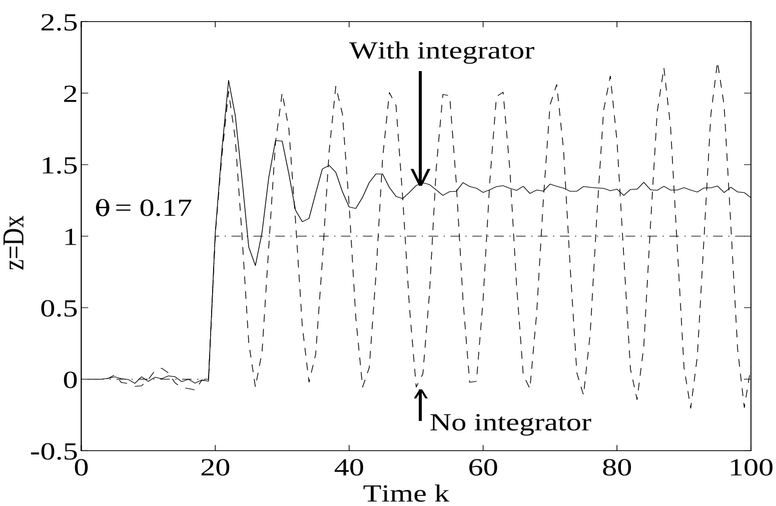

Figure 6: Augmented system with modelling errors, $\theta = 0.17$

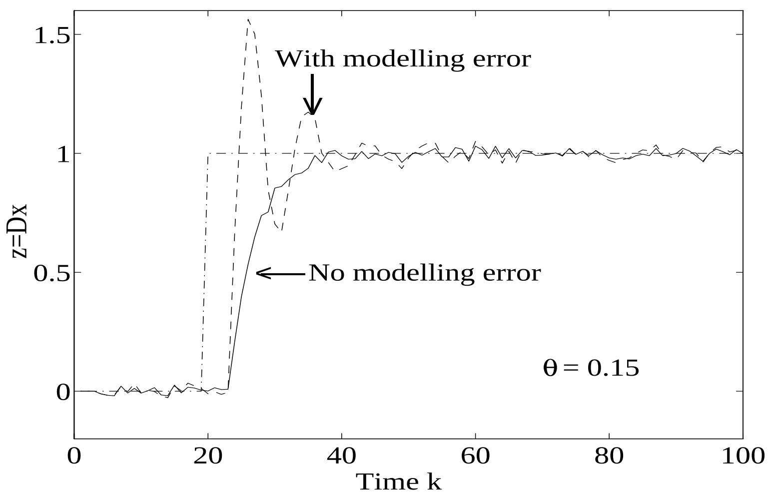

Figure 7: Doubly augmented system $\theta = 0.15$

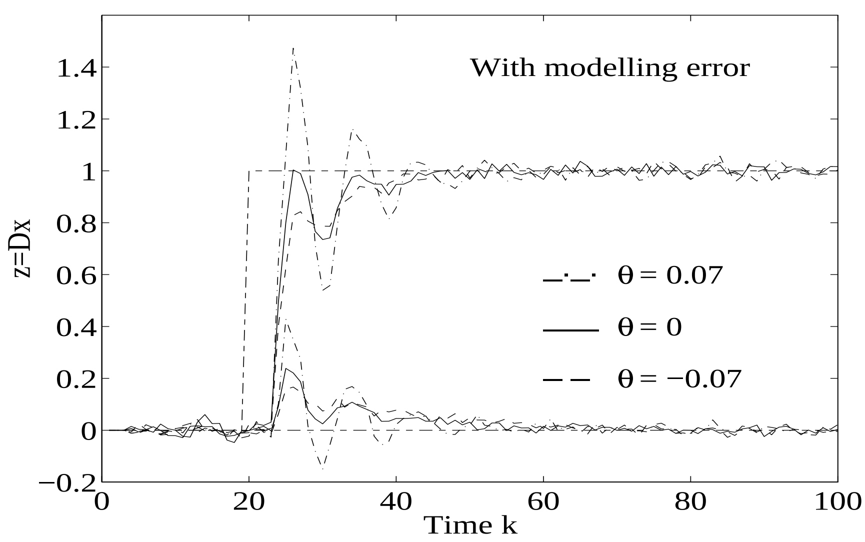

Figure 8: Two-input two-output system with modelling errors

References

[1] L. Aggoun, A. Bensoussan, R.J. Elliott, and J.B. Moore. Finite-dimensional quasi-linear risk-sensitive control, *Systems and Control Letters*, to appear, 1994.

[2] B.D.O. Anderson and J.B. Moore. *Optimal Filtering.* New Jersey: Prentice-Hall, 1979.

[3] B.D.O. Anderson and J.B. Moore. *Optimal Control: Linear Quadratic Methods.* New Jersey: Prentice-Hall, 1989.

[4] A. Bensoussan and J.H. van Schuppen. Optimal control of partially observable stochastic systems with an exponential-of-integral performance index, *SIAM Jour. on Control and Optimization* **23** (1985), 599–613.

[5] P. Caravani. On extending linear quadratic control theory to non-symmetric risky objectives, *Jour. Economic Dynamics & Control* **10** (1986), 83–88.

[6] D.H. Jacobson. Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games, *IEEE Trans. on Automatic Control* **AC-18(2)** (1973), 124–131.

[7] M.R. James, J.S. Baras, and R.J. Elliott. Risk-sensitive control and dynamic games for partially observed discrete-time systems, *IEEE Trans. on Automatic Control* **AC-39(4)** (1994), 780–792.

[8] M.R. James and R.J. Elliott. Risk-sensitive and risk-neutral control for continuous-time hidden Markov models, preprint.

[9] L.S. Karp. Dynamic hedging with uncertain production, *Int. Economic Review* **29** (1988), 621–637.

[10] P.R. Kumar and P. Varaiya. *Stochastic Systems.* New Jersey: Prentice-Hall, 1986.

[11] A. Segall. Recursive estimation from discrete-time point processes, *IEEE Trans. on Information Theory* **IT-22(4)** (1976), 422–431.

[12] P. Whittle. Risk-sensitive linear/quadratic/Gaussian control, *Advances in Applied Probability* **13** (1981), 764–777.

I.B. COLLINGS, M.R. JAMES, J.B. MOORE

Cooperative Research Centre for Sensor Signal and Info. Processing, SPRI Building, Technology Park Adelaide, The Levels, SA 5095, Australia. Currently at: Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, Victoria 3052, Australia

Department of Systems Engineering, RSISE, Australian National University, Canberra ACT 0200, Australia

Department of Systems Engineering, RSISE, Australian National University, Canberra ACT 0200, Australia

Communicated by David Hill