Minimax Optimal Control of Stochastic Uncertain Systems with Relative Entropy Constraints

Ian R. Petersen† Matthew R. James‡ Paul Dupuis§

Abstract

This paper considers a new class of discrete time stochastic uncertain systems in which the uncertainty is described by a constraint on the relative entropy between a nominal noise distribution and the perturbed noise distribution. This uncertainty description is a natural extension to the case of stochastic uncertain systems, of the sum quadratic constraint uncertainty description. The paper solves problems of worst case robust performance analysis and output feedback minimax optimal controller synthesis in a general nonlinear setting. Specializing these results to the linear case leads to a minimax LQG optimal controller. This controller is defined in terms of Riccati difference equations and a Kalman Filter like state equation.

1 Introduction

One of the key ideas to emerge in the field of modern control is the use of optimization and optimal control theory to give a systematic procedure for the design of feedback control

---

*This work was supported by the Australian Research Council and in part by the Office of Naval Research (ONR N00014-96-1-0276). Part of this research took place while the last author was a Visiting Fellow of the Faculty for Engineering and Information Technology at The Australian National University. Support through the ANU 1995 Faculties Research Grants Scheme is gratefully acknowledged.

†Department of Electrical Engineering, Australian Defence Force Academy, Campbell, 2600, Australia, email: irp@ee.adfa.oz.au.

‡Department of Engineering, Faculty of Engineering and Information Technology, Australian National University, Canberra, ACT 0200, Australia; e-mail: Matthew.James@anu.edu.au

§LCDS, Division of Applied Mathematics, Box F, Brown University, 192 George St., Providence, RI 02912, USA.
systems. For example, in the case of linear systems with full state measurements, the linear quadratic regulator (LQR) approach provides one of the most useful techniques for designing state feedback controllers; e.g., see [1]. Also, in the case of linear systems with partial information, the solution to the linear quadratic Gaussian (LQG) stochastic optimal control problem provides a useful technique for the design of multivariable output feedback controllers; see [1]. However, the above optimal control techniques suffer from a major disadvantage in that they do not provide a systematic means for addressing the issue of robustness. Indeed in the partial information case, an LQG controller may lead to very poor robustness properties; e.g., see [2].

The lack of robustness which may result when designing a control system via standard optimal control methods such as mentioned above has been a major motivation for research in the area of robust control. Research in this area has resulted in a large number of approaches being available to deal with the issue of robustness in control system design. These methods include those based on Kharitonov’s Theorem (see [3]), $H^\infty$ control theory (e.g., see [4]) and the quadratic stabilizability approach (e.g., see [5]). One important idea found in robust control theory is the notion of an uncertain system. This involves modelling not only the nominal dynamics of the system to be controlled but also modelling the uncertainty in the system. Many different methods of modelling uncertainty have been considered. These include $H^\infty$ norm bounded linear time invariant uncertainty blocks (e.g., see [6]), time-varying norm bounded uncertainty blocks (e.g., see [5]) and more recently, uncertainty blocks satisfying an integral quadratic constraint (e.g., see [7]). By considering the problem of controlling uncertain systems, the issue of robustness with respect to specified uncertainties can be systematically addressed. In particular, results such as [7] on minimax optimal control of uncertain systems provide a method of incorporating the issue of robustness into the optimal controller design methodologies mentioned above. However, it should be noted that [7] is a state feedback result and in the partial information case, the existing results available are much less satisfactory. One of the underlying motivations for the results to be presented in this paper is to develop an uncertain system framework whereby the standard LQG stochastic optimal controller design methodology can be extended into a minimax stochastic optimal control methodology for uncertain systems with partial information. This would allow the designer to retain all of the existing methods for choosing LQG weighting functions to achieve the desired nominal performance (e.g., see [1]) whilst the issue of robustness can be addressed by appropriate choice of the uncertainty structure in the uncertain system model.

One of the main contributions of this paper is to introduce a new class of nonlinear discrete time stochastic uncertain systems. The class of stochastic uncertain systems to be introduced involves modelling uncertainty in terms of a constraint referred to as a relative entropy constraint. This uncertainty description is a modification of the uncertainty
description which was presented in [8]. As in [8], our uncertainty description allows us to model structured uncertain dynamics in discrete time systems subject to stochastic noise processes. However, our uncertainty structure is slightly more restricted than that considered in [8] in that we must assume that all uncertainty inputs affect the system via the same channel as the stochastic noise process. Our motivation for considering uncertain systems which are subject to stochastic noise process disturbances is twofold. In the first instance, many engineering control problems involve dealing with systems which are subject to disturbances and measurement noise which can be well modelled by stochastic processes. A second motivation is that in going from the state feedback LQR optimal control problem to the measurement feedback LQG optimal control problem, a critical change in the model is the introduction of noise disturbances. Hence, it would be expected that in order to obtain a reasonable generalization of LQG control for uncertain systems, it would be necessary to consider uncertain systems which are subject to disturbances in the form of stochastic noise processes. The reason why we use a slightly different uncertain system model in this paper as opposed to that used in [8] is that in the partial information case, the uncertain system model chosen in [8] led to an apparently intractable partial information stochastic differential game problem. By using the uncertainty description proposed here, we are led to a more tractable partial information risk sensitive control problem.

The organization of the paper is as follows. In Section 2 of the paper, we consider a problem of worst case performance analysis for stochastic uncertain systems. This section begins by defining a very general class of nonlinear discrete time stochastic uncertain systems considered over a finite time interval. The uncertain systems considered are described by a nominal system which is driven by a stochastic noise process with a specified statistical description. Also considered is a perturbed system in which a general class of stochastic noise processes are allowed. These perturbed noise processes must satisfy a certain constraint on the relative entropy between the nominal noise process statistics and the perturbed noise process statistics; see Section 2 for a definition of relative entropy. This relative entropy constraint is an extension of the stochastic uncertain constraint considered in [8] and the integral quadratic constraint uncertainty description such as considered in [9]. This uncertainty description allows for uncertainty in both the nonlinear system dynamics and the statistics of the applied noise process.

The problem considered in Section 2 is a problem of worst case performance analysis with respect to a specified cost function. This is a constrained maximization problem where the maximization is with respect to the uncertainty in the system or equivalently with respect to the perturbed noise probability measure. The first main result presented in this section uses a Lagrange multiplier technique to convert the constrained maximization problem into an unconstrained maximization problem. This result is an extension of S-procedure ideas such as presented in [8]. The second main result of this section uses a
result on the duality between free energy and relative entropy (e.g., see [10]) to convert this unconstrained maximization problem into a problem of evaluating a risk sensitive cost functional. Both of the results of Section 2 are used in subsequent sections on controller synthesis.

Section 3 of the paper considers a problem of minimax optimal controller synthesis for a general class of discrete time nonlinear stochastic uncertain systems with partial information. Using the results of Section 2, it is first shown that these problems can be converted into corresponding unconstrained stochastic game problems with partial information. The stochastic game problem is then converted into an equivalent risk sensitive optimal control problem with partial information. The subsequent sections show how existing results can be applied to solve these partial information risk sensitive optimal control problems.

Section 4 of the paper applies the results of [11] and [12] to solve the above mentioned risk sensitive optimal control problems in a general nonlinear setting. In particular, the dynamic programming approach of [15] is applied in the special case of full state information and the information state approach of [12] is applied in the measurement feedback case.

Section 5 of the paper considers that special case in which the underlying nominal system is linear. In this case, the results of [13] and [14] are applied to solve the corresponding risk sensitive optimal control problems. In the state feedback case, the results of [13] lead to a Riccati equation solution to the risk sensitive control problem. Indeed, the corresponding control algorithm obtained in this case in virtually identical to that which was obtained in the previous paper [8]. However, [8] deals with a somewhat different uncertainty description than that which is considered in this paper.

As mentioned above, one of the main motivations for this paper was to consider the minimax stochastic optimal control problem in the linear partial information case. This problem is the robust generalization of the LQG problem. In Section 5 of the paper, we use the results of [14] to solve the risk sensitive control problem corresponding to this special case. The solution to the discrete time, linear, partial information risk sensitive control problem was originally obtained in [15]. However for our purposes, it was more convenient to use the results of [14] since they give an explicit characterization of the optimal risk sensitive cost. The application of the results of [14] leads to a partial information controller which is obtained via the solution of two Riccati type recursions and involves a state estimator based controller structure of the type which occurs in $H^\infty$ control. Indeed, the solution to the linear partial information risk sensitive optimal control problem is known to be closely related to the $H^\infty$ control problem; e.g., see [16].

The main achievement of this paper is the solution of a partial information minimax optimal control problem for stochastic uncertain linear systems. A number of other authors have considered related control problems in the linear partial information case. In particular, the mixed $H_2/H_\infty$ problem (e.g., see [17, 18, 19, 20]) is a closely related problem.
to that considered in this paper. However, in many cases our problem formulation which involves minimizing the worst case cost for a specific uncertain system is the natural way to formulate a robust performance problem. This is in contrast to the problems considered in $H_2/H_\infty$ control theory which typically involve minimizing nominal performance subject to a robust stability constraint. Also our solution which is given in terms of Riccati equations is a direct extension of the standard LQG method.

2 Worst Case Performance for Stochastic Uncertain Systems

The results of this section are concerned with the problem of characterizing the worst case performance for a stochastic uncertain system. In order to apply the results of this section in our subsequent consideration of minimax optimal control problems, it will be convenient to introduce a very general definition of a stochastic uncertain system. This definition will be specialized in the sequel to more concrete classes of stochastic uncertain system.

**Reference System** The stochastic uncertain systems under consideration are defined in terms of a reference or nominal system and a perturbed system. In this section, the reference system to be considered is defined by a sequence of operators mapping the state sequence and noise sequence into the next value of the state:

$$x_{k+1} = G_k(x_0, w_0).$$

Here, $x_{0N} = \{x_0, x_1, \ldots, x_N\}$ denotes the state sequence and $w_{0N} = \{w_0, w_1, \ldots, w_N\}$ denotes the noise sequence. For each $k \in \{0, 1, \ldots, N\}$, $w_k \in \mathbb{R}^r$ and $x_k \in \mathbb{R}^n$. Thus, we can write $w_{0N} \in \mathbb{R}^{r(N+1)}$ and $x_{0N+1} \in \mathbb{R}^{n(N+2)}$. It is assumed that for each $k$, the operator $G_k(\cdot)$ defines a Borel measurable mapping $\mathbb{R}^{n(k+1)} \times \mathbb{R}^{r(k+1)} \rightarrow \mathbb{R}^n$. The initial condition vector $x_0$ and the noise input sequence $w_{0N}$ are random variables with a specified joint probability measure $\mu(dw_{0N} \times dx_0)$. In the sequel, we will make use of the following decomposition of the measure $\mu$

$$\mu(dw_{0N} \times dx_0) = \mu_{x_0}(dx_0)\mu_0(dw_0|x_0)\mu_1(dw_1|x_0, w_0)\mu_2(dw_2|x_0, w_0, w_1)\ldots\mu_N(dw_N|x_0, w_0, \ldots w_{N-1}).$$
Perturbed System The corresponding perturbed system is also defined by a sequence of operators mapping the state sequence and noise sequence into the next value of the state:

\[
\begin{align*}
    x_{k+1} &= G_k(x_0k, \bar{w}_0k) \\
    z_k &= b_k(x_k)
\end{align*}
\]

where \(\bar{w}_{0N}\) is a perturbed noise sequence and \(z_{0N+1}\) is the output sequence. For each \(k\), \(z_k \in \mathbb{R}^q\) and the function \(b_k(\cdot)\) is assumed to be Borel measurable. The joint probability measure of \(\bar{w}_{0N}\) and \(x_0\) is denoted \(\nu(\cdot)\) and is contained in the set \(\mathcal{P}\) of all probability measures on the variables \(\bar{w}_{0N}\) and \(x_0\). In the sequel, we will make use of the following decomposition of the measure \(\nu\):

\[
\nu(\cdot) = \nu_{x_0}(dx_0) \nu_1(d\bar{w}_1|x_0, \bar{w}_0) \nu_2(d\bar{w}_2|x_0, \bar{w}_0, \bar{w}_1) \ldots \nu_N(d\bar{w}_N|x_0, \bar{w}_0, \ldots \bar{w}_{N-1}).
\]

It follows immediately from this definition that \(\mathcal{P}\) is a convex set of probability measures.

Remarks Note that the class of probability measures \(\mathcal{P}\) defined above is actually equivalent to the class of probability measures \(\mathcal{P}\) defined in [8].

The interpretation of equations (1) and (2) together with the corresponding set of noise probability measures \(\mathcal{P}\) as a stochastic uncertain system is as follows. The set \(\mathcal{P}\) will be further restricted in the sequel with the introduction of the relative entropy constraint on the noise probability measures. This restricted set together with equation (2) will define a family of stochastic systems. This family of stochastic systems is a stochastic uncertain system. The true system is assumed to be a member of this family.

Stochastic Uncertain System Associated with the systems (1), (2) and set of probability measures \(\mathcal{P}\), we now define a stochastic uncertain dynamical system. This is achieved by specifying the class of admissible perturbed noise processes for the perturbed system (2). The admissible perturbed noise processes will be random processes such that a certain Relative Entropy Constraint is satisfied. To define this relative entropy constraint, first let \(d_1 > 0\), \(d_2 > 0\), \ldots, \(d_s > 0\) be given positive constants. We also recall the following standard definition of relative entropy; e.g., see [10]. Given any two probability measures \(\gamma(\cdot)\) and \(\theta(\cdot)\) defined on the same measurable space, the relative entropy \(R(\gamma(\cdot)||\theta(\cdot))\) is defined by

\[
R(\gamma(\cdot)||\theta(\cdot)) = \begin{cases} 
\int \left( \log \frac{d\gamma}{d\theta} \right) d\theta & \text{if } \theta(\cdot) << \gamma(\cdot) \text{ and } \log \frac{d\gamma}{d\theta} \in L_1(d\theta) \\
+\infty & \text{otherwise.}
\end{cases}
\]
The relative entropy can be regarded as a measure of the “distance” between the probability measure \( \gamma(\cdot) \) and the probability measure \( \theta(\cdot) \). The properties of relative entropy can be found in Section 1.4 of [10].

Associated with the systems (1) and (2) is the following set of non-negative valued Borel measurable functions:

\[
q_{ij}(\cdot) : \mathbb{R}^p \to \mathbb{R} \quad i = 0, 1, \ldots, N \quad j = 1, 2, \ldots, s.
\]  

These functions will determine the admissible perturbed noise processes in the system (2). Indeed, the admissible perturbed noise random processes are defined by the elements of the set \( \mathcal{P} \) such that the following condition is satisfied:

**Relative Entropy Constraint** A probability measure \( \nu \in \mathcal{P} \) defines an admissible perturbed noise random process if

\[
E_{\nu} \left[ R(\nu_x(\cdot) \parallel \mu_x(\cdot)) - d_i 
+ \sum_{k=0}^{N} [R(\nu_k(\cdot \mid x_0, \bar{w}_0, \ldots, \bar{w}_{k-1}) \parallel \mu_k(\cdot \mid x_0, \bar{w}_0, \ldots, \bar{w}_{k-1})) - q_{ki}(z_k)] \right] 
\]

\[
= R(\nu(\cdot) \parallel \mu(\cdot)) - E_{\nu} \left[ \sum_{k=0}^{N} q_{ki}(z_k) + d_i \right] 
\]

\[
\leq 0 \quad (4)
\]

for all \( i = 1, 2, \ldots, s \). Note that the equality above follows from the chain rule for relative entropy; see [10].

The set of all probability measures \( \nu \in \mathcal{P} \) defining admissible perturbed noise random processes is denoted \( \Xi \). Note that it follows from one of the properties of relative entropy and the assumed non-negativity of the functions \( q_{ki}(\cdot) \) that that the set \( \Xi \) is non-empty. Indeed, \( \Xi \) will always contain the probability measure \( \mu(\cdot) \) as defined for the reference system.

**Remarks** The uncertainty description embodied in the above relative entropy constraint is a generalization of the stochastic sum quadratic constraint considered in [8]. Indeed, as in the stochastic sum uncertainty constraint, this uncertainty description allows for nonlinear time varying uncertainty blocks together with exogenous noise entering into the perturbed noise process. The special case of a stochastic system without uncertainty is obtained if we let

\[
q_{ij}(\cdot) \equiv 0 \quad \& \quad d_i = 0 \quad \forall i.
\]  

In this case, the relative entropy constraint (4) forces the relative entropy \( R(\nu(\cdot) \parallel \mu(\cdot)) \) to be zero. Strictly, this case does not satisfy the assumption on the constants \( d_i \). However,
it is easily to see that the case of a stochastic system without uncertainty can be obtained in the limit as we let the constants $d_i$ and the functions $q_{ij}(\cdot)$ approach zero.

Note that the relative entropy $R(\nu(\cdot)||\mu(\cdot))$ can be thought of as a measure of the difference between the reference probability measure $\mu(\cdot)$ and the perturbed probability measure $\nu(\cdot)$. In particular, the relative entropy is zero if and only if $\mu(\cdot) \equiv \nu(\cdot)$. Typical perturbations which are allowed under the above relative entropy constraint are perturbations in the mean of the probability measure $\mu$ such as illustrated in Figure 1. More specific
in [8], to the relative entropy uncertainty description considered here is that it enables us to convert problems of minimax optimal control into equivalent risk sensitive control problems. This is in contrast to the approach of [8] in which the minimax optimal control problem was converted into an equivalent stochastic game problem. The approach taken in this paper may give significant computational advantages since risk sensitive control problems are often easier to solve than stochastic game problems.

**Cost Functional** In this section, we consider the problem of characterizing, for a stochastic uncertain system defined as above, the worst case performance with respect to a cost functional defined to be:

\[ J(z_{0N+1}) = \Phi(z_{N+1}) + \sum_{k=0}^{N} L(k, z_k) \]  

where the functions \( \Phi(\cdot) \) and \( L(\cdot) \) are Borel measurable.

The problem under consideration in this section is to find

\[ \sup_{\nu \in \Xi} E_\nu \{ J(z_{0N+1}) \}. \]  

Our first step in evaluating this quantity is to use a Lagrange multiplier technique to convert the problem from a constrained optimization problem into an unconstrained optimization problem. Indeed, given a vector \( \tau = [\tau_1, \tau_2, \ldots, \tau_s] \in \mathbb{R}^s \), we define an augmented cost function as follows:

\[ J_\tau(z_{0N+1}) = \Phi(z_{N+1}) + \sum_{k=0}^{N} L(k, z_k) \]

\[ -\sum_{i=1}^{s} \tau_i \left[ R(\nu_{x_0})(\cdot)\|\mu_{x_0}(\cdot)) - d_i \right. \]

\[ \left. + \sum_{k=0}^{N} \left[ R(\nu_k(\cdot|x_0, \bar{w}_0, \ldots, \bar{w}_{k-1})||\mu_k(\cdot|x_0, \bar{w}_0, \ldots, \bar{w}_{k-1})) - q_{ki}(z_k) \right] \right] \]

Now, we define \( V_\tau \) to be the value of the corresponding unconstrained optimization problem:

\[ V_\tau \triangleq \sup_{\nu \in P} E_\nu \{ J_\tau(z_{0N+1}) \}. \]

Also, we define a set \( \Gamma \subset \mathbb{R}^s \) as

\[ \Gamma \triangleq \{ \tau = [\tau_1, \tau_2, \ldots, \tau_s] \in \mathbb{R}^s : \tau_1 \geq 0, \tau_2 \geq 0, \ldots, \tau_s \geq 0 \& V_\tau < \infty \}. \]
Using these definitions, we are now in a position to state the following theorem relating the constrained optimization problem and the unconstrained optimization problem. This theorem is an extension of the S-procedure ideas presented in [8]. The proof of this theorem will follow from a standard result on convex analysis and the convexity of the relative entropy $R(\gamma(\cdot)\|\theta(\cdot))$ with respect to the probability measure $\gamma(\cdot)$ (e.g., see [10]).

**Theorem 2.1** Consider the stochastic uncertain system (1), (2), (4) with cost functional $J(z_{0N+1})$. Then the following conditions hold:

(i) The supremum $\sup_{\nu \in \Xi} E_{\nu} [J(z_{0N+1})]$ is finite if and only if the set $\Gamma$ is non-empty.

(ii) If the set $\Gamma$ is non-empty, then

$$\sup_{\nu \in \Xi} E_{\nu} [J(z_{0N+1})] = \min_{\tau \in \Gamma} V_{\tau}. \tag{8}$$

**Proof.** We will establish this result using Theorem 1 on page 217 of [21]. In order to apply this result to the constrained optimization problem (7), we must first establish that all of the conditions required to apply this result are satisfied.

First note that the reference probability measure $\mu(\cdot)$ is contained in the set $\mathcal{P}$ and that when $\nu(\cdot) = \mu(\cdot)$, we have $R(\nu_{x_0}(\cdot)\|\mu_{x_0}(\cdot)) = 0$ and

$$R(\nu_k(\cdot|x_0, \tilde{w}_0, \ldots \tilde{w}_{k-1})\|\mu_k(\cdot|x_0, \tilde{w}_0, \ldots \tilde{w}_{k-1})) = 0 \quad \forall k$$

Hence, using the fact that $q_{ki}(\cdot) \geq 0$ and $d_i > 0$, it follows that

$$E_{\nu} \left[ \sum_{k=0}^{N} \left( -q_{ki}(z_k) \right) \right] \leq -d_i < 0$$

for all $i$. That is the relative entropy constraint (4) is strictly satisfied. This is the constraint qualification condition required in order to apply the result of [21].

In order to apply the result of [21], we will also use the fact that the function

$$E_{\nu} \sum_{i=1}^{s} \tau_i \left[ R(\nu_{x_0}(\cdot)\|\mu_{x_0}(\cdot)) - d_i \right] + \left( -q_{ki}(z_k) \right)$$

10
is a convex function of the probability measure $\nu$. This follows since
\[
E_{\nu}\left[R(\nu_{x_0}(\cdot)\|\mu_{x_0}(\cdot)) + \sum_{k=0}^{N} (R(\nu_k(\cdot|x_0, \bar{w}_0, \ldots \bar{w}_{k-1})\|\mu_k(\cdot|x_0, \bar{w}_0, \ldots \bar{w}_{k-1}))\right]
\]
is a representation of the relative entropy $R(\nu(\cdot)\|\mu(\cdot))$ which is a convex function of
the probability measure $\nu$; see Theorem C.3.1 and Theorem 1.4.3 of [10]. Furthermore $E_{\nu}J(z_{0,N+1})$ is a linear and hence concave function of the probability measure $\nu$. Moreover, the set $\mathcal{P}$ is convex. Combining all of the above facts, we have established that all of the conditions required in order to apply Theorem 1 on page 217 of [21] to the constrained optimization problem (7) are satisfied.

We now establish statement $(i)$ of the theorem. First, suppose that
\[
\sup_{\nu \in \Xi} E_{\nu}\{ J(z_{0,N+1}) \} = c < \infty
\]
Then it follows directly from the result of [21] that there exists a vector $\tau^* \geq 0$ such that
\[
V_{\tau^*} = \sup_{\nu \in \mathcal{P}} E_{\nu}\{ J_{\tau^*}(z_{0,N+1}) \} = c < \infty.
\]

Conversely, if there exists a vector $\tau^* \geq 0$ such that
\[
V_{\tau^*} = \sup_{\nu \in \mathcal{P}} E_{\nu}\{ J_{\tau^*}(z_{0,N+1}) \} = c < \infty, \tag{9}
\]
then given any $\nu \in \Xi$, it follows from (9) and the relative entropy constraint (4) that $E_{\nu}\{ J(z_{0,N+1}) \} \leq c$. Hence,
\[
\sup_{\nu \in \Xi} E_{\nu}\{ J(z_{0,N+1}) \} \leq c < \infty.
\]
This completes the proof of $(i)$.

To establish $(ii)$, we first observe that given any vector $\tau \geq 0$, it follows from the relative entropy constraint (4), that for any $\nu \in \mathcal{P}$
\[
E_{\nu}\{ J_{\tau}(z_{0,N+1}) \} \geq E_{\nu}\{ J(z_{0,N+1}) \}.
\]
Hence,
\[
\sup_{\nu \in \mathcal{P}} E_{\nu}\{ J_{\tau}(z_{0,N+1}) \} \geq \sup_{\nu \in \Xi} E_{\nu}\{ J_{\tau}(z_{0,N+1}) \} \geq \sup_{\nu \in \Xi} E_{\nu}\{ J(z_{0,N+1}) \}
\]
for all $\tau \geq 0$. Also, it follows from the result of [21] that there exists a $\tau^* \geq 0$ such that
\[
\sup_{\nu \in \mathcal{P}} E_{\nu}\{ J_{\tau^*}(z_{0,N+1}) \} = \sup_{\nu \in \Xi} E_{\nu}\{ J(z_{0,N+1}) \}
\]
Thus, part (ii) of the theorem has been established.

**Remark** In the case of no uncertainty (5), it follows from the positivity of the relative entropy $R(\nu(\cdot)||\mu(\cdot))$ that the minimum in (8) will be achieved in the limit as $\sum_{i=1}^{s} \tau_i \to \infty$. Note that this does not contradict the existence of a minimum in (8) since as mentioned above, the conditions (5) are strictly not permitted under the standing assumptions but can only be approached in the limit.

In the sequel, we will require that the system (1), satisfies the following assumption.

**Assumption 2.1**

$$\sup_{\nu \in \mathcal{P}} E_{\nu} \{ J(z_{0N+1}) \} = \infty$$

In this assumption, note that we are effectively maximizing the cost functional (6) with respect to the noise input $w_k$. Hence, this assumption amounts to a controllability type assumption with respect to the input $w_k$ and an observability type assumption with respect to the cost functional (6). For example, suppose (2) is a linear system of the form

$$x_{k+1} = Ax_k + Dw_k$$
$$z_k = Ex_k$$

and (6) is a quadratic cost functional of the form

$$J(z_{0N+1}) = \sum_{k=0}^{N} z_k^T Q z_k$$

where the pair $(A, D)$ is controllable and the pair $(A, E^T Q E)$ is observable. Then, it is straightforward to verify that this assumption will be satisfied.

This assumption has been introduced in order to rule out the case in which the minimum in (8) is achieved at $\tau = 0$. Alternatively, we could have explicitly ruled out this case in the results to follow.

**Remarks** It follows from the above assumption that the zero vector is not contained in the set $\Gamma$. Indeed, if $\tau = 0$ then $J_\tau(z_{0N+1}) = J(z_{0N+1})$ and hence $V_{\tau} = \infty$. Therefore, the zero vector is not contained in $\Gamma$.

For any non-zero vector $\tau \geq 0$, it is straightforward to verify that the $V_{\tau}$ can be re-written as

$$V_{\tau} = \sum_{i=1}^{s} \tau_i (W_{\tau} + d_i)$$

(10)
where
\[
W_\tau \triangleq \sup_{\nu \in \mathcal{P}} \mathbb{E}_\nu \left\{ \frac{\Phi(z_{N+1}) + \sum_{k=0}^{N} L(k, z_k) + \sum_{i=1}^{s} \tau_i \sum_{k=0}^{N} q_{ki}(z_k)}{\sum_{i=1}^{s} \tau_i} - R(\nu_{x_0}(\cdot)||\mu_{x_0}(\cdot)) - \sum_{k=0}^{N} (R(\nu_k(\cdot|x_0, \bar{w}_1, \ldots, \bar{w}_{k-1})||\mu_k(\cdot|x_0, \bar{w}_1, \ldots, \bar{w}_{k-1}))) \right\}.
\]

Hence, it follows from Theorem 2.1 that if \( \Gamma \neq \emptyset \), we can write
\[
\sup_{\nu \in \Xi} \mathbb{E}_\nu [J(z_{0,N+1})] = \min_{\tau \in \Gamma} \sum_{i=1}^{s} \tau_i (W_\tau + d_i). \tag{11}
\]

We now look at a risk sensitive method for evaluating the quantity \( W_\tau \). This is obtained directly using the duality between relative entropy and free energy which occurs in the theory of large deviations:

**Lemma 2.2** For each \( \tau \neq 0 \),
\[
W_\tau = \log \mathbb{E}_\mu \left\{ \exp \left[ \frac{\Phi(z_{N+1}) + \sum_{k=0}^{N} L(k, z_k) + \sum_{i=1}^{s} \tau_i \sum_{k=0}^{N} q_{ki}(z_k)}{\sum_{i=1}^{s} \tau_i} \right] \right\}
\]
where the probability measure \( \mu(\cdot) \) is as defined for the reference system (1).

**Proof.** When \( \Phi, L, \) and the \( q_{ki} \) are bounded this result follows from the relative entropy representation for logarithms of exponential integrals (Proposition 1.4.2 in [10]) and an application of the chain rule for relative entropy (Theorem C.3.1 in [10]). When one or more of these non-negative functions is unbounded, we use Proposition 4.5.1 of [10] in place of Proposition 1.4.2. \( \square \)

**Remarks** By evaluating \( W_\tau \) using this formula, the required worst case cost (7) can be found by solving the finite dimensional optimization problem corresponding to (11). Also note that it follows from (10) and the definition of \( W_\tau \) that \( W_\tau \) is a convex function of \( \tau \). Hence, the finite dimensional optimization problem in (11) is a convex optimization problem.

At this point, we recall that for the case of no uncertainty (5), the minimum in (11) is achieved in the limit as \( \sum_{i=1}^{s} \tau_i \to \infty \). Furthermore in this case, it follows from a standard result on risk sensitive cost functions that
\[
\lim_{\sum_{i=1}^{s} \tau_i \to \infty} \sum_{i=1}^{s} \tau_i W_\tau = \mathbb{E}_\mu \left\{ \Phi(z_{N+1}) + \sum_{k=0}^{N} L(k, z_k) \right\}.
\]
That is, as expected, in the case of no uncertainty, the formula (11) reduces to a simple evaluation of the cost function.

13
3 Minimax Controller Synthesis

In this section, we apply the approach developed in the previous section to the problem of synthesizing a controller which solves a minimax stochastic optimal control problem for a stochastic uncertain system with relative entropy constraints. The main result of this section shows that this problem can be converted into an equivalent problem of risk sensitive optimal control which can be solved via information state methods. The class of systems considered in this section will take a more concrete form than those considered in the previous section. As in the previous section, we will consider stochastic uncertain systems defined in terms of a known reference (nominal) system and a perturbed system.

**Reference System** The reference system to be considered in this section is defined by the state equations

\[
\begin{align*}
  x_{k+1} &= f(k, x_k, u_k) + g(k, x_k)w_k \\
  y_{k+1} &= h(k, x_k) + v_k
\end{align*}
\]  

where \( x_k \in \mathbb{R}^n \) is the system state with initial condition \( x_0 \), \( w_k \in \mathbb{R}^r \) is the process noise input, \( v_k \in \mathbb{R}^l \) is the measurement noise input, \( u_k \in \mathbb{R}^m \) is the control input and \( y_k \in \mathbb{R}^l \) is the measured output. The functions \( f(k, \cdot) \), \( g(k, \cdot) \) and \( h(k, \cdot) \) are assumed to be Borel measurable. The initial condition vector \( x_0 \), the process noise input sequence \( w_0^N \) and measurement noise sequence \( v_0^N \) are assumed to be random variables with a specified joint probability measure \( \mu(dw_0^N \times dv_0^N \times dx_0) \). This joint probability measure is assumed to have the following dependence structure:

\[
\begin{align*}
  \mu(dw_0^N \times dv_0^N \times dx_0) \\
  = \mu_{x_0}(dx_0)\mu_{w_0}(dw_0|x_0)\mu_{v_0}(dv_0|x_0)\mu_{w_1}(dw_1|x_0, w_0, v_0)\mu_{v_1}(dv_1|x_0, w_0, v_0) \\
  \times \mu_{w_2}(dw_2|x_0, w_0, w_1, v_0, v_1)\mu_{v_2}(dv_2|x_0, w_0, w_1, v_0, v_1) \ldots \\
  \times \mu_{w_N}(dw_N|x_0, w_0, \ldots w_{N-1}, v_0, \ldots v_{N-1})\mu_{v_N}(dv_N|x_0, w_0, \ldots w_{N-1}, v_0, \ldots v_{N-1}).
\end{align*}
\]

This implies the conditional independence of the system noise and observation noise.

**Perturbed System** The corresponding perturbed system is also defined by the state equations

\[
\begin{align*}
  x_{k+1} &= f(k, x_k, u_k) + g(k, x_k)\bar{w}_k \\
  z_k &= b(k, x_k, u_k) \\
  y_{k+1} &= h(k, x_k) + \bar{v}_k
\end{align*}
\]  

(13)
where \( z_k \in \mathbb{R}^q \) is the uncertainty output. In this case, \( \tilde{w}_{0N} \) and \( \tilde{v}_{0N} \) are perturbed noise sequences. Also, the functions \( b(k, \cdot) \) are assumed to be Borel measurable. Furthermore, the uncertainty output vector \( z_k \) is assumed to be partitioned as

\[
  z_k = \begin{bmatrix}
    z^1_k \\
    z^2_k \\
    \vdots \\
    z^s_k
  \end{bmatrix} = \begin{bmatrix}
    b^1(k, x_k, u_k) \\
    b^2(k, x_k, u_k) \\
    \vdots \\
    b^s(k, x_k, u_k)
  \end{bmatrix}.
\] (14)

The joint probability measure of \( \tilde{w}_{0N}, \tilde{v}_{0N} \) and \( x_0 \) is denoted \( \nu(d\tilde{w}_{0N} \times d\tilde{v}_{0N} \times dx_0) \) and is contained in the set \( \mathcal{P} \) of all probability measures on the variables \( \tilde{w}_{0N}, \tilde{v}_{0N} \) and \( x_0 \). In the sequel, we will make use of the following decomposition of the measure \( \nu \):

\[
  \nu(d\tilde{w}_{0N} \times d\tilde{v}_{0N} \times dx_0)
  = \nu_{x_0}(dx_0)\nu_0(d\tilde{w}_0 \times d\tilde{v}_0|x_0)\nu_1(d\tilde{w}_1 \times d\tilde{v}_1|x_0, w_0, v_0)\nu_2(d\tilde{w}_2 \times d\tilde{v}_2|x_0, w_0, w_1, v_0, v_1)
  \cdots \nu_N(d\tilde{w}_N \times d\tilde{v}_N|x_0, w_0, \ldots, w_{N-1}, v_0, \ldots, v_{N-1}).
\]

It follows immediately from this definition that \( \mathcal{P} \) is a convex set of probability measures.

**Admissible Controllers**

We consider causal partial information controllers of the form

\[
  u_k = \mathcal{K}(k, y_{0k+1})
\] (15)

where \( u_k \in \mathbb{R}^m \) is the control input at time \( k \) and \( y_{0k+1} \) is the output sequence over the time interval \( \{0, 1, \ldots, k + 1\} \). It is assumed that the operator \( \mathcal{K}(k, \cdot) \) defines a Borel measurable mapping \( \mathbb{R}^{l(k+1)} \to \mathbb{R}^m \). The class of all such controllers is denoted \( \Lambda \).

Associated with the stochastic uncertain system \( (12), (13) \) is the following set of non-negative valued Borel measurable functions:

\[
  q_{ij}(\cdot) : \mathbb{R}^p \to \mathbb{R} \quad i = 0, 1, \ldots, N \quad j = 1, 2, \ldots, s
\] (16)

These functions will determine the admissible perturbed noise processes in the system \( (13) \).

**Relative Entropy Constraint**

In order to complete the definition of the controlled stochastic uncertain system corresponding to the state equations \( (12), (13) \), we now specify the class of admissible perturbed noise sequences. The admissible perturbed noise processes will be random processes satisfying a Relative Entropy Constraint. To define this relative entropy constraint, first let \( d_1 > 0, d_2 > 0, \ldots, d_s > 0 \) be given positive constants.
Given any admissible controller of the form (15), a probability measure $\nu(\cdot) \in \mathcal{P}$ defines an admissible perturbed noise random process for the resulting closed loop system (13), (15) if
\[
E_{\nu} \left[ \begin{array}{l}
R(\nu_{x_0}(\cdot)||\mu_{x_0}(\cdot)) - d_i \\
+ \sum_{k=0}^{N} \left( R(\nu_k(\cdot|x_0, \bar{w}_0, \ldots, \bar{w}_{k-1}, \bar{v}_0, \ldots, \bar{v}_{k-1})||\mu_k(\cdot|\bar{x}_0, \bar{v}_0, \ldots, \bar{v}_{k-1})) \\
- q_{ki}(z^i_k) \right) \end{array} \right] \leq 0 \quad (17)
\]
for all $i = 1, 2, \ldots, s$. Here
\[
\mu_k(dw_k \times dv_k|x_0, \bar{w}_0, \ldots, \bar{w}_{k-1}, \bar{v}_0, \ldots, \bar{v}_{k-1}) \\
= \mu_{kw}(dw_k|x_0, \bar{w}_0, \ldots, \bar{w}_{k-1}, \bar{v}_0, \ldots, \bar{v}_{k-1}) \\
\times \mu_{kv}(dv_k|x_0, \bar{w}_0, \ldots, \bar{w}_{k-1}, \bar{v}_0, \ldots, \bar{v}_{k-1}).
\]
For a given controller $K$ of the form (15), the set of all probability measures $\nu \in \mathcal{P}$ defining admissible perturbed noise random processes is denoted $\Xi_K$. Note that for any controller $K \in \Lambda$, it follows from one of the properties of relative entropy and the assumed non-negativity of the functions $q_{ki}(\cdot)$, that the set $\Xi_K$ will contain the probability measure $\mu(\cdot)$ corresponding to the reference system (12).

**Remarks** To motivate the above uncertainty description and its connection to the Sum Quadratic Constraint (SQC) uncertainty description such as considered in [22], we now look at the special case in which the reference probability measure is Gaussian. That is,
\[
\mu(dw_{0N} \times dv_{0N} \times dx_0) = \prod_{k=0}^{N} \theta(dw_k) \prod_{k=0}^{N} \eta(dv_k) \psi(dx_0) \quad (18)
\]
where
\[
\theta(d\xi) = [(2\pi)^r \det(\Upsilon)]^{-\frac{1}{2}} \exp\left[-\frac{1}{2} \xi^T \Upsilon^{-1} \xi\right] d\xi \\
\eta(d\xi) = [(2\pi)^l \det(\Omega)]^{-\frac{1}{2}} \exp\left[-\frac{1}{2} \xi^T \Omega^{-1} \xi\right] d\xi \\
\psi(d\xi) = [(2\pi)^n \det(\bar{\Sigma}_0)]^{-\frac{1}{2}} \exp\left[-\frac{1}{2} (\xi - \bar{x}_0)^T \bar{\Sigma}_0^{-1} (\xi - \bar{x}_0)\right] d\xi.
\]
Hence, the reference noise sequences $w_k$ and $v_k$ are independent zero mean white Gaussian noise sequences with noise covariance matrices $\Upsilon > 0$ and $\Omega > 0$ respectively. Also, the
initial condition $x_0$ is a Gaussian random variable with mean $\bar{x}_0$ and covariance matrix $\Sigma_0 > 0$. Furthermore, we suppose that the function $q_{ki}(z)$ is of the form

$$q_{ki}(z) = \frac{1}{2}||z_i||^2.$$ 

In this case, the relative entropy constraint will allow for a perturbed system of the form

$$x_{k+1} = f(k, x_k, u_k) + g(k, x_k) (\Delta_{1k}(x_k, u_k) + w_k)$$

$$z_k = b(k, x_k, u_k)$$

$$y_{k+1} = h(k, x_k) + \Delta_{2k}(x_k, u_k) + v_k$$

where the initial condition is a Gaussian random variable with mean $\bar{x}_0 + \Delta x_0$ and covariance matrix $\Sigma_0$. Then $\bar{w}_k = w_k + \Delta_{1k}$, $\bar{v}_k = w_k + \Delta_{2k}$ and the relative entropy constraint will be satisfied provided the following sum quadratic constraint is satisfied:

$$E \left[ \frac{1}{2} \Delta_{x0}^T \Sigma_0^{-1} \Delta x_0 + \frac{1}{2} \sum_{k=0}^{N} \left( \Delta_{1k}^T \Omega_0^{-1} \Delta_{1k} + \Delta_{2k}^T \Omega_0^{-1} \Delta_{2k} - ||z_i||^2 \right) \right] \leq 0$$

for all $i = 1, 2, \ldots, s$. Note however that even in this special case of a Gaussian reference probability measure, the class of perturbed probability measures satisfying the relative entropy constraint (17) is larger than the class of perturbed probability measures defined by the above sum quadratic constraint.

**Cost Functional** The cost functional to be considered in this section is of the form

$$J(x_{0N+1}, u_{0N}) = \Phi(x_{N+1}) + \sum_{k=0}^{N} L(k, x_k, u_k)$$

where the functions $\Phi(\cdot)$ and $L(k, \cdot)$ are Borel measurable. This cost functional extends the cost functional (6) to the case of controlled stochastic uncertain systems.

The above system and cost functional is required to satisfy the following assumptions.

**Assumption 3.1** The functions $\Phi(\cdot)$ and $L(k, \cdot)$ satisfy

$$\Phi(x) \geq 0 \text{ and } L(k, x, u) \geq 0$$

for all $k = 0, 1, \ldots, N$, $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$.

**Assumption 3.2** For any admissible controller $K \in \Lambda$, the resulting closed loop system is such that

$$\sup_{\nu \in P} E_{\nu} J(x_{0N+1}, u_{0N}) = \infty.$$
Note that Assumption 3.2 is a similar controllability and observability assumption to Assumption 2.1. Although condition (21) is required to hold for all admissible controllers, the fact that the measurement noise $v_k$ can be arbitrary means that this assumption is of a similar type to Assumption 2.1.

The minimax control problem under consideration in this section involves finding an admissible controller to minimize the worst case of the expectation of the cost functional (20). That is, we are concerned with the minimax control problem

$$\inf_{K \in \Lambda} \sup_{\nu \in \Xi} E_{\nu} J(x_{0N+1}, u_{0N}).$$

In the following theorem, we show that this constrained minimax problem can be replaced by a corresponding unconstrained stochastic game problem. This stochastic game problem is defined in terms of the following augmented cost functional

$$J_{\tau}(x_{0N+1}, u_{0N})$$

$$= \Phi(x_{N+1}) + \sum_{k=0}^{N} L(k, x_k, u_k)$$

$$- \sum_{i=1}^{s} \left[ R(\nu_{x_0}(\cdot) \| \mu_{x_0}(\cdot)) - d_i + \sum_{k=0}^{N} \left( R(\nu_k(\cdot|x_0, \bar{w}_0, \ldots, \bar{w}_{k-1}, \bar{v}_0, \ldots, \bar{v}_{k-1}) \| \mu_k(\cdot|x_0, \bar{w}_0, \ldots, \bar{w}_{k-1}, \bar{v}_0, \ldots, \bar{v}_{k-1})) \right) \right]$$

where $\tau_1 \geq 0, \tau_2 \geq 0, \ldots, \tau_s \geq 0$, are given constants. In this stochastic game problem, the maximizing player input is a probability measure $\nu(\cdot) \in P$ and the minimizing player input $u_{0k}$ is assumed to have a partial information structure of the form (15). We let $\tilde{V}_{\tau}$ denote the lower value in this game problem. That is

$$\tilde{V}_{\tau} \triangleq \inf_{K \in \Lambda} \sup_{\nu \in \Xi} E [J_{\tau}(x_{0N+1}, u_{0N})].$$

Also, we define a set $\tilde{\Gamma} \subset \mathbb{R}^s$ as

$$\tilde{\Gamma} \triangleq \{ \tau = [\tau_1, \tau_2, \ldots, \tau_s] \in \mathbb{R}^s : \tau_1 \geq 0, \tau_2 \geq 0, \ldots, \tau_s \geq 0 \& \tilde{V}_{\tau} \text{ is finite} \}.$$ 

It follows from Assumption 3.2 that zero is not contained in the set $\tilde{\Gamma}$.

**Theorem 3.1** Consider the stochastic uncertain system (12), (13), (17) with cost functional $J(z_{0N+1})$. Then the following conclusions hold:
For the minimax stochastic optimal control problem
\[
\inf_{K \in \Lambda} \sup_{\nu \in \Xi} E_{\nu} [J(x_{0N+1}, u_{0N})],
\]
the value of this optimal control problem is finite if and only if the set \( \tilde{\Gamma} \) is non-empty.

(ii) If the set \( \tilde{\Gamma} \) is non-empty, then
\[
\inf_{K \in \Lambda} \sup_{\nu \in \Xi} E_{\nu} [J(x_{0N+1}, u_{0N})] = \inf_{\tau \in \tilde{\Gamma}} \tilde{V}_{\tau}.
\]

**Proof.** Using Theorem 2.1, the proof of this result follows via identical steps to the proof of Theorem 3.1 of [8].

**Remark** For any non-zero vector \( \tau \geq 0 \), it is straightforward to verify that the quantity \( \tilde{V}_{\tau} \) can be re-written as
\[
\tilde{V}_{\tau} = \sum_{i=1}^{s} \tau_i \left( \tilde{W}_{\tau} + d_i \right)
\]

where
\[
\tilde{W}_{\tau} = \inf_{K \in \Lambda} \sup_{\nu \in \mathcal{P}} \left\{ \frac{\Phi(x_{N+1}) + \sum_{k=0}^{N} L(k, x_k, u_k) + \sum_{i=1}^{s} \tau_i \sum_{k=0}^{N} q_k(z_k^i)}{\sum_{i=1}^{s} \tau_i} - R(x_0) \right. \\
\left. - \sum_{k=0}^{N} \left( R(x_0) \right) \right\}
\]

Hence, it follows from Theorem 3.1 that if \( \tilde{\Gamma} \neq \emptyset \), we can write
\[
\inf_{K \in \Lambda} \sup_{\nu \in \Xi} E_{\nu} [J(x_{0N+1}, u_{0N})] = \inf_{\tau \in \tilde{\Gamma}} \sum_{i=1}^{s} \tau_i \left( \tilde{W}_{\tau} + d_i \right).
\]

The following theorem shows that the partial information stochastic game defining the quantity \( \tilde{W}_{\tau} \) can be replaced by an equivalent partial information risk sensitive optimal control problem

**Theorem 3.2**
\[
\tilde{W}_{\tau} = \inf_{K \in \Lambda} \log E_{\mu} \left\{ \exp \left( \frac{\Phi(x_{N+1}) + \sum_{k=0}^{N} L(k, x_k, u_k) + \sum_{i=1}^{s} \tau_i \sum_{k=0}^{N} q_k(z_k^i)}{\sum_{i=1}^{s} \tau_i} \right) \right\}
\]
where the probability measure \( \mu(\cdot) \) is as defined for the reference system (12).
Proof. The proof of this theorem follows directly by applying Lemma 2.2 to the closed loop system formed by applying an arbitrary controller $K \in \Lambda$ to the uncertain system (12), (13), (17); see also [10]. \qed

Remark Theorems 3.1 and 3.2 presented above enable us to convert problems of minimax optimal control for the class of stochastic uncertain systems under consideration into corresponding risk sensitive optimal control problems dependent on the vector of Lagrange multipliers $\tau \geq 0$. In the sequel, we consider the solution of these risk sensitive optimal control problems.

4 Solution to the Risk Sensitive Optimal Control Problems

In this section, we apply some existing results on risk sensitive optimal control given in [11] and [12] to solve the risk sensitive optimal control problem derived in the previous section.

4.1 State Feedback Case

We first consider the state feedback case using the results of [11]. In order to apply the results of [11], we assume that the stochastic uncertain system under consideration is such that the reference state equation (12) is of the form

$$x_{k+1} = f(k, x_k, u_k) + g(k, x_k)w_k$$

and the perturbed state equation (13) is of the form

$$x_{k+1} = f(k, x_k, u_k) + g(k, x_k)\bar{w}_k$$

$$z_k = b(k, x_k, u_k)$$

(29)

For a given vector $\tau \geq 0$, will solve the risk sensitive optimal control problem associated with the system (28) and the risk sensitive cost functional

$$J^{RS}_\tau = \log E_{\mu} \left\{ \exp \left[ \Phi(x_{N+1}) + \sum_{k=0}^{N} L(k, x_k, u_k) + \sum_{i=1}^{s} \tau_i \sum_{k=0}^{N} q_ki(z_k^i) \right] \right\}. \quad (30)$$
**Assumptions** We assume that the reference noise probability measure $\mu(\cdot)$ is of the form

$$\mu(dw_0 x dx_0) = \prod_{k=0}^{N} \theta(dw_k)\psi(dx_0)$$

(31)

where

$$\psi(dx_0) = \delta_{\bar{x}_0}(dx_0)$$

That is, the initial condition is assumed to be known for the reference system.

It follows from Theorem 3.2 that the solution to the risk sensitive control problem (28), (30), (31) yields the quantity $\tilde{W}_\tau$ in equation (27). We now use dynamic programming to solve this state feedback risk sensitive control problem.

**Dynamic Programming Equation** In order to solve the risk sensitive control problem (28), (30), (31), we consider a function $Z_\tau(x; k)$ defined by the following dynamic programming equation:

$$Z_\tau(x; k) = \inf_u \log \mathbb{E}_\mu \left[ \exp \left( \frac{L(k, x, u) + \sum_{i=1}^{\tau} \tau_i q_{k_i}(z_{k_i})}{\sum_{i=1}^{\tau} \tau_i} + Z_\tau(f(k, x, u) + g(k, x)w_k; k + 1) \right) \right]$$

$$Z_\tau(x, N + 1) = \Phi(x)$$

(32)

Applying standard dynamic programming methods to the risk sensitive control problem (28), (30), (31), we obtain the following proposition; e.g., see [11].

**Proposition 4.1** Let the non-zero vector $\tau \geq 0$ be given and suppose $Z_\tau(x; k)$ is defined by the dynamic programming equation (32). Then $\tilde{W}_\tau$, the optimal value of the risk sensitive control problem (28), (30), (31) is given by

$$\tilde{W}_\tau = Z_\tau(\bar{x}_0; 0).$$

(33)

Furthermore, if for each $k$ and $x \in \mathbb{R}^n$, the infimum in (32) is achieved at $u(x, k)$ then this defines the optimal state feedback controller in the risk sensitive control problem (28), (30), (31).

Combining the above proposition with Theorems 3.1 and Theorem 3.2, we can then solve the minimax optimal control problem (24) as follows: For each non-zero vector $\tau \geq 0$, we solve the dynamic programming equation (32) and use equation (33) to determine the corresponding value of $\tilde{W}_\tau$. Then, as in (26) we optimize with respect to the vector $\tau \geq 0$. Assuming that the infimum in (26) is achieved at a non-zero $\tau = \tau^*$, we then solve the dynamic programming equation (32) with this value of $\tau$. Then the optimal cost can be evaluated using formulas (33) and (26). Also, if the infimum in (32) is achieved at each $k$ and $x \in \mathbb{R}^n$, this defines the corresponding minimax stochastic optimal controller.
4.2 Partial Information Case

We now apply the results of [12] to solve the risk sensitive optimal control problem (27) in the partial information case. In order to apply the results of [12], we assume that the stochastic uncertain system under consideration is such that the reference state equation (12) is of the form:

\[ x_{k+1} = f(k, x_k, u_k) + g(k, x_k)w_k \]
\[ y_{k+1} = h(k, x_k) + v_k. \]  

Also, the perturbed state equation (13) is assumed to be of the form:

\[ x_{k+1} = f(k, x_k, u_k) + g(k, x_k)\bar{w}_k \]
\[ z_k = b(k, x_k, u_k) \]
\[ y_{k+1} = h(k, x_k) + \bar{v}_k. \]  

For a given non-zero vector \( \tau \geq 0 \), will solve the partial information risk sensitive control problem associated with the system (34) and the risk sensitive cost functional (30).

**Assumptions** We assume that the reference noise probability measure \( \mu(\cdot) \) is of the form

\[ \mu(dw_0N \times dv_0N \times dx_0) = \prod_{k=0}^{N} \theta(dw_k) \prod_{k=0}^{N} \eta(dv_k) \psi(dx_0) \]  

where

\[ \theta(d\xi) = [(2\pi)^r \det(\Upsilon)]^{-\frac{1}{2}} \exp\left[-\frac{1}{2} \xi^T \Upsilon^{-1} \xi\right] d\xi \]
\[ \eta(d\xi) = [(2\pi)^l \det(\Omega)]^{-\frac{1}{2}} \exp\left[-\frac{1}{2} \xi^T \Omega^{-1} \xi\right] d\xi \]
\[ \psi(d\xi) = [(2\pi)^n \det(\bar{\Sigma}_0)]^{-\frac{1}{2}} \exp\left[-\frac{1}{2} (\xi - \bar{x}_0)^T \bar{\Sigma}_0^{-1} (\xi - \bar{x}_0)\right] d\xi. \]

That is, we assume that the reference noise probability measure is Gaussian.

It follows from Theorem 3.2 that the solution to the risk sensitive control problem (34), (30), (36) yields the quantity \( \bar{W}_\tau \) in equation (27). We now use information state approach of [12] to solve this partial information risk sensitive control problem. In order to apply the results of [12] to this partial information risk sensitive optimal control problem, we first introduce some definitions regarding the information state:
**Notation** Using the results of [12], it follows that the solution to the partial information risk sensitive optimal control problem under consideration can be given in terms of an information state process \( \sigma_k(z) \) defined recursively as follows:

\[
\sigma_{k+1} = \Sigma(k, u_k, y_{k+1})\sigma_k, \quad \sigma_0 = \psi
\]  

where the operator \( \Sigma(k, u, y) \) is defined as follows:

\[
\Sigma(k, u, y)\sigma(z)dz = \int_{\mathbb{R}^n} \frac{\zeta_k(dz; \xi)}{(2\pi)^{n/2}} \exp \left( \frac{L(k, \xi, u) + \sum_{i=1}^n \tau_i q_{ki}(b_i(k, \xi, u))}{\sum_{i=1}^n \tau_i} \right) \Psi(k, \xi, y)\sigma(z)dz.
\]

where

\[
\zeta_k(dz; \xi) = P_{\mu}(f(k, \xi, u) + g(k, \xi)w \in dz)
\]

and

\[
\Psi(k, \xi, y) = \exp \left( -\frac{1}{2} h(k, \xi)^T \Omega^{-1} h(k, \xi) + h(k, \xi)^T \Omega^{-1} y \right).
\]

**Dynamic Programming Equations** As well as the above information state recursion, the approach of [12] also involves a dynamic programming equation. In order to present this dynamic programming equation, we must first define a change of measure for the system (34). This change of measure is defined in terms of an alternative reference system:

\[
x_{k+1} = f(k, x_k, u_k) + g(k, x_k)w_k
\]

\[
y_{k+1} = v_k
\]

Given any controller of the form (15), let \( \alpha(dx_{0:N}, dy_{0:N}) \) denote the probability measure on the sequences \( x_{0:N}, y_{0:N} \) induced by the probability measure \( \mu(\cdot) \) and the equations (34), (15). Also, let \( \alpha'(dx_{0:N}, dy_{0:N}) \) denote the probability measure on the sequences \( x_{0:N}, y_{0:N} \) induced by the probability measure \( \mu(\cdot) \) and the equations (39), (15). As in [12], the Radon-Nikodym derivative between these two probability measures is given by

\[
\frac{d\alpha}{d\alpha'} \bigg|_{x_k} = \prod_{i=1}^{N+1} \Psi(i, x_{i-1}, y_i)
\]

where \( \mathcal{F}_k \) denotes the complete filtration generated by \((x_{0k}, y_{0k})\).
Using this notation, we then define quantity $S(\sigma, k)$ according to the following dynamic programming equation

$$S(\sigma, k) = \inf_u E^\dagger [S(\Sigma(k+1, u, y_{k+1}) \sigma, k+1)]$$

$$S(\sigma, N+1) = \int_{R^n} \sigma(x) \exp \Phi(x) dx \quad (40)$$

where $E^\dagger$ denotes expectation with respect to the probability measure $\alpha^\dagger(\cdot)$.

Using the above notation, we are now in a position to present the solution to the partial information risk sensitive control problem (28), (30), (31) given in [12]. Note that the results of [12] are stated under the assumption that the functions in the system (34) and cost function (30) are bounded and uniformly continuous and that the controls are restricted to a bounded set. However, it is straightforward to verify that the proof of the result in [12] which leads to the following proposition does not require these assumptions.

**Proposition 4.2** Consider the partial information risk sensitive optimal control problem (28), (30), (31) and suppose that the information state $\sigma_k$ is defined by the equation (37) and the function $S(\sigma, k)$ is defined by the dynamic programming equation (40). Then $\tilde{W}_\tau$ the optimal value in this risk sensitive control problem is given by

$$\tilde{W}_\tau = S(\psi, 0).$$

Furthermore, suppose $u_k^*$ is an admissible controller such that for each $k$,

$$u_k^* = \tilde{u}_k^*(\sigma_k)$$

where $\tilde{u}_k^*(\sigma)$ achieves the infimum in (40). Then this controller is an optimal controller for the partial information risk sensitive optimal control problem under consideration.

As in the state feedback case, we can use this result to solve the minimax optimal control problem (22) in the partial information case. This is achieved by optimizing over the non-zero vector $\tau \geq 0$ to find the infimum in (26). For the optimal value of $\tau$ (if it exists), the corresponding partial information minimax optimal controller is obtained as in the above proposition.

5 **The Linear Quadratic Gaussian Case**

In this section, we specialize the results of the previous section to the linear quadratic Gaussian case. Using the results of [13], we present the solution to the state feedback risk
sensitive optimal control problem in terms of a Riccati difference equation. Furthermore, using the results of [14] and [15], we present the solution to the partial information risk sensitive optimal control problem in terms of a pair of Riccati difference equations. These results then enable solutions to the corresponding minimax LQG problems to be presented.

5.1 State Feedback Case

In order to apply the results of [13], we assume that the stochastic uncertain system under consideration is such that the reference state equation is of the form

\[
x_{k+1} = A_k x_k + B_k u_k + D_k w_k
\]  

(41)

and the perturbed state equation is of the form

\[
\begin{align*}
x_{k+1} &= A_k x_k + B_k u_k + D_k \bar{w}_k \\
z_k &= E_{1k} x_k + E_{2k} u_k
\end{align*}
\]

(42)

**Assumptions** We assume that the vector \( z_k \) has been partitioned as in (14) and write

\[
z_k^i = E_{1k}^i x_k + E_{2k}^i u_k
\]

for \( i = 1, 2, \ldots, s \). Furthermore, we assume that

\[
\left( E_{1k}^i \right)^T E_{2k}^i = 0 \quad \forall \ i, \ k.
\]

Also, we assume that the functions \( q_{ki}(z) \) in the relative entropy constraint (4) are of the form

\[
q_{ki}(z) = \| z \|^2 \quad \forall \ i.
\]

(43)

It is assumed that the cost functional (20) is of the form

\[
J(x_{0N+1}, u_{0N}) = \frac{1}{2} x_{N+1}^T Q_{N+1} x_{N+1} + \frac{1}{2} \sum_{k=0}^{N} \left( x_k^T Q_k x_k + u_k^T R_k u_k \right)
\]

(44)

where

\[
Q_k \geq 0 \quad \& \quad R_k > 0 \quad \forall k.
\]

Furthermore, we assume that this system and cost functional satisfies Assumption 3.2.
For a given vector \( \tau \geq 0 \), we will solve the risk sensitive control problem associated with the system (41) and the risk sensitive cost functional

\[
J^{RS}_\tau = \log \mathbf{E}_\mu \left\{ \exp \left[ \frac{1}{2} x_{N+1}^T Q_{N+1} x_{N+1} + \frac{1}{2} \sum_{k=0}^{N} x_k^T \left[ Q_k + 2 \sum_{i=1}^{s} \tau_i (E_{1k}^i)^T E_{1k}^i \right] x_k \right] \right\}.
\]

We assume that the reference noise probability measure \( \mu(\cdot) \) is a Gaussian probability measure of the form

\[
\mu(dw_0 \times dx_0) = \prod_{k=0}^{N} \theta_k(dw_k)\psi(dx_0)
\]

where

\[
\theta_k(d\xi) = [(2\pi)^r \det(\Upsilon_k)]^{-\frac{1}{2}} \exp\left[-\frac{1}{2} \xi^T \Upsilon_k^{-1} \xi\right] d\xi
\]

\[
\psi(dx_0) = \delta_{\bar{x}_0}(dx_0).
\]

Note, that we have assumed that \( x_0 \) has a known value of \( \bar{x}_0 \) for the reference system.

**Riccati Equations** The solution to the state feedback linear quadratic Gaussian risk sensitive control problem (41), (45), (46) is given in terms of the following Riccati difference equation:

\[
P_k = Q_k + 2 \sum_{i=1}^{s} \tau_i (E_{1k}^i)^T E_{1k}^i
\]

\[
+ A_k^T \left[ \hat{P}_{k+1} - \tilde{P}_{k+1} B_k \left( R_k + 2 \sum_{i=1}^{s} \tau_i (E_{2k}^i)^T E_{2k}^i + B_k^T \tilde{P}_{k+1} B_k \right)^{-1} B_k^T \hat{P}_{k+1} \right] A_k
\]

\[
\tilde{P}_{k+1} = P_{k+1} + P_{k+1} D_k \left( \Upsilon_k^{-1} - D_k^T P_{k+1} D_k \right)^{-1} D_k^T P_{k+1}
\]

\[
P_{N+1} = Q_{N+1}.
\]

We will require that the solution to this Riccati equation satisfies the following conditions:

\[
\Upsilon_k^{-1} - D_k^T P_{k+1} D_k > 0 \ \forall k
\]

\[
P_k > 0 \ \forall k.
\]
Also, the following recursive equation needs to be solved:

\[
F_k = F_{k+1} \frac{\sqrt{\det \left( \left[Y_k^{-1} - D_k^T P_{k+1} D_k \right]^{-1} \right)}}{\det(Y_k)}
\]

\[
F_{N+1} = 1.
\]  (49)

Then, applying the result of [13] to the risk sensitive control problem (41), (45), (46), we obtain the following proposition:

**Proposition 5.1** Let the non-zero vector \( \tau \geq 0 \) be given and suppose \( P_k, \tilde{P}_k \) and \( F_k \) are defined as above and satisfy conditions (48). Then \( \tilde{W}_\tau \), the optimal value of the risk sensitive control problem (41), (45), (46) is given by

\[
\tilde{W}_\tau = \log \left[ F_0 \exp \left( \bar{x}_0^T P_0 \bar{x}_0 \right) \right] = \bar{x}_0^T P_0 \bar{x}_0 + \log F_0
\]  (50)

Furthermore, the corresponding state feedback optimal control law is given by

\[
u_k = - \left( R_k + 2 \sum_{i=1}^{s} \tau_i \left( E_{2k}^i \right)^T E_{2k}^i + B_k^T \tilde{P}_{k+1} B_k \right)^{-1} B_k^T \tilde{P}_{k+1} A_k x_k \]

for \( k = 0, 1, \ldots, N \).

As in the previous section, we can use this result to solve the minimax optimal control problem (22) in the state feedback linear quadratic Gaussian case. This is achieved by optimizing over the vector \( \tau \geq 0 \) to find the infimum in (26). For the optimal value of \( \tau \) (if it exists), the corresponding partial information minimax optimal controller is obtained as in the above proposition.

**Remark** The controller design algorithm defined by the above proposition and the optimization with respect to \( \tau \geq 0 \) in (26) is essentially the same as the controller design algorithm proposed in [8] in the linear state feedback case. This is in spite of the fact that a different uncertainty description is considered here. However, in the linear partial information case to follow, the results presented here are different to those of [8]. Indeed in [8], no tractable solution could be obtained in the linear partial information case.
5.2 Partial Information Case

We now apply the results of [14] (see also [15] for an equivalent result) to solve the risk
sensitive optimal control problem (27) in the partial information case. In order to apply
the results of [14], we assume that the stochastic uncertain system under consideration is
such that the reference state equation (12) is of the form:
\begin{align}
x_{k+1} &= Ax_k + Bu_k + Dw_k \\
y_{k+1} &= Cx_k + v_k.
\end{align}
(51)

Also, the perturbed state equation (13) is assumed to be of the form:
\begin{align}
x_{k+1} &= Ax_k + Bu_k + D\bar{w}_k \\
z_k &= E_1x_k + E_2u_k \\
y_{k+1} &= Cx_k + \bar{v}_k.
\end{align}
(52)

As in the state feedback case, we assume that the vector $z_k$ has been partitioned as in
(14) and write
\[
z^i_k = E^i_1x_k + E^i_2u_k
\]
for $i = 1, 2, \ldots, s$. Furthermore, we assume that
\[
(E^1_1)^T E^2_2 = 0 \quad \forall \ i, \ k.
\]

Also, for this uncertain system, we assume that the uncertainty is described by the relative
entropy constraint (17) where the functions $q_{ki}(z)$ are of the form (43).

We assume that the reference noise probability measure $\mu(\cdot)$ is a Gaussian probability
measure of the form
\[
\mu(dw_0 \times dv_0 \times dx_0) = \prod_{k=0}^N \theta(dw_k) \prod_{k=0}^N \eta(dv_k)\psi(dx_0)
\]
(53)

where
\[
\begin{align*}
\theta(\xi) &= [(2\pi)^r \det(\Upsilon)]^{-\frac{1}{2}} \exp\left[-\frac{1}{2} \xi^T\Upsilon^{-1}\xi\right]\xi^T \\
\eta(\xi) &= [(2\pi)^l \det(\Omega)]^{-\frac{1}{2}} \exp\left[-\frac{1}{2} \xi^T\Omega^{-1}\xi\right]\xi^T \\
\psi(\xi) &= [(2\pi)^n \det(\Sigma_0)]^{-\frac{1}{2}} \exp\left[-\frac{1}{2} (\xi - \bar{x}_0)^T\Sigma_0^{-1}(\xi - \bar{x}_0)\right]\xi^T.
\end{align*}
\]
Hence, the reference noise sequences $w_k$ and $v_k$ are independent zero mean white Gaussian noise sequences with noise covariance matrices $\Upsilon > 0$ and $\Omega > 0$ respectively. Also, the initial condition $x_0$ is a Gaussian random variable with mean $\bar{x}_0$ and covariance matrix $\bar{\Sigma}_0 > 0$.

For a given vector $\tau \geq 0$, will solve the partial information risk sensitive control problem associated with the system (51), the risk sensitive cost functional (45) and noise process (53). In the cost function (45) it is assumed that $Q_k \equiv Q \geq 0$ & $R_k \equiv R > 0 \ \forall k$.

**Information State Equations** The solution to the partial information linear quadratic Gaussian risk sensitive control problem (51), (45), (53) is given in terms of the following Riccati difference equation which is solved forward in time:

$$
\Sigma_{k+1} = D \Upsilon D^T + A \left( \Sigma_k^{-1} + C^T \Omega^{-1} C - Q - 2 \sum_{i=1}^{s} \tau_i \left( E_i^i \right)^T E_i^i \right)^{-1} A^T \\
\Sigma_0 = \bar{\Sigma}_0.
$$

(54)

The solution to this difference equation is required to satisfy the following conditions

$$
\Sigma_k^{-1} + C^T \Omega^{-1} C - Q - 2 \sum_{i=1}^{s} \tau_i \left( E_i^i \right)^T E_i^i > 0 \ \forall k \\
\Sigma_k > 0 \ \forall k
$$

(55)

We also consider the following filter state equation:

$$
\hat{x}_{k+1} = A \hat{x}_k + B u_k \\
+ A \left( \Sigma_k^{-1} + C^T \Omega^{-1} C - Q - 2 \sum_{i=1}^{s} \tau_i \left( E_i^i \right)^T E_i^i \right)^{-1} \\
\times \left( C^T \Omega^{-1} \left[ y_{k+1} - C \hat{x}_k \right] + \left[ Q + 2 \sum_{i=1}^{s} \tau_i \left( E_i^i \right)^T E_i^i \right] \hat{x}_k \right)
$$

(56)

The quantities $\Sigma_k, \hat{x}_k$ together with a normalizing constant $Z_k$ defined in [14] define the information state of the risk sensitive control problem (51), (45), (53).
Dynamic Programming Equations As well as the above information state equations which are solved forward in time, the solution to the risk sensitive control problem (51), (45), (53) also involves the following recursive equations which are derived via dynamic programming; see [14]. These equations are solved backwards in time.

\[
S_k^a = \tilde{M}_k + \tilde{A}_k^T S_{k+1}^a \left( I + B \left[ R + 2 \sum_{i=1}^{s} \tau_i \left( E_2^i \right)^T E_2^i \right]^{-1} B^T S_{k+1}^a - \tilde{\Gamma}_k \tilde{\Gamma}_k^T S_{k+1}^a \right)^{-1} \tilde{A}_k \\
S_{N+1}^a = \left[ Q_{N+1}^{-1} - \Sigma_{N+1} \right]^{-1} \\
S_k^c = S_{k+1}^c + \log \left[ \det(\Sigma_{k+1})^{-1} \det \left( \Sigma_k^{-1} - Q - 2 \sum_{i=1}^{s} \tau_i (E_1^i)^T E_1^i \right)^{-1} \right] \times \det(I - \tilde{\Gamma}_k^T S_{k+1}^a \tilde{\Gamma}_k)^{-1} \\
S_{N+1}^c = \log \left[ (2\pi)^n \det(\Sigma_{N+1}^{-1} - Q_{N+1})^{-1} \right]
\]

where

\[
\tilde{S}_{k+1} = \left[ \left( S_{k+1}^a \right)^{-1} - \tilde{\Gamma}_k \tilde{\Gamma}_k^T \right]^{-1} \\
K_k^a = \left( R + 2 \sum_{i=1}^{s} \tau_i \left( E_2^i \right)^T E_2^i + B^T \tilde{S}_{k+1}^a B \right)^{-1} B^T \tilde{S}_{k+1}^a \tilde{A}_k \\
\tilde{K}_k = \Sigma_k^{-1} + C^T \Omega^{-1} C - Q - 2 \sum_{i=1}^{s} \tau_i (E_1^i)^T E_1^i \\
\tilde{\Gamma}_k = A \tilde{K}_k C^T \Omega^{-1} C \Theta_k \\
\Theta_k = \left[ \left( C^T \Omega^{-1} C \right)^{-1} + \rho_k^{-1} \Sigma_k \right]^{1/2} \\
\tilde{A}_k = A \rho_k^{-1} \\
\tilde{M}_k = \left( Q + 2 \sum_{i=1}^{s} \tau_i \left( E_1^i \right)^T E_1^i \right) \rho_k^{-1} \\
\sigma = R + 2 \sum_{i=1}^{s} \tau_i \left( E_2^i \right)^T E_2^i + B^T \tilde{S}_{k+1}^a B \\
\delta_k = I - \tilde{\Gamma}_k^T S_{k+1}^a \tilde{\Gamma}_k \\
\rho_k = I - \Sigma_k \left( Q + 2 \sum_{i=1}^{s} \tau_i \left( E_1^i \right)^T E_1^i \right).
\]

The solutions to the above equations will be required to satisfy the following condition:

\[
I - \tilde{\Gamma}_k^T S_{k+1}^a \tilde{\Gamma}_k > 0 \quad (59)
\]

30
Note that the above recursion for $S^c_k$ was omitted from [14] but is straightforward to derive using the approach of [14].

Applying the result of [14] to the risk sensitive control problem (51), (45), (53), we obtain the following proposition:

**Proposition 5.2** Let the non-zero vector $\tau \geq 0$ be given and suppose $S^a_k, S^c_k, \tilde{S}_{k+1}, \tilde{A}_k$ and $\tilde{x}_k$ are defined as above and conditions (55), (59) are satisfied. Then $\tilde{W}_\tau$, the optimal value of the risk sensitive control problem (51), (45), (53) is given by

$$
\tilde{W}_\tau = \log \left[ (2\pi)^n \det(\Sigma_0) \right]^{-\frac{1}{2}} \exp \left( \frac{1}{2} \left[ \tilde{x}_0 S^a_0 \tilde{x}_0 + S^c_0 \right] \right) \frac{1}{2} \log \left[ (2\pi)^n \det(\Sigma_0) \right].
$$

(60)

Furthermore, the corresponding partial information feedback optimal control law is given by

$$
u_k = - \left( \Sigma_k + 2 \sum_{i=1}^a \tau_i \left( E_{2k}^i \right)^T E_{2k}^i + B_k^T \tilde{S}_{k+1} B_k \right)^{-1} B_k^T \tilde{S}_{k+1} \tilde{A}_k \hat{x}_k
$$

for $k = 0, 1, \ldots, N$.

As above, we can use this result to solve the minimax optimal control problem (22) in the partial information linear quadratic Gaussian case. This is achieved by optimizing over the vector $\tau \geq 0$ to minimize the optimal value $\tilde{W}_\tau$. For this optimal value of $\tau$, the corresponding state feedback minimax optimal controller is obtained as in the above proposition.

**Remarks** We recall from Section 2 that the case of no uncertainty (5) corresponds to the limit $\sum_{i=1}^a \tau_i \to \infty$. In this case, as observed in [14], the controller equations (54), (56) and (57) reduce to the equations of the standard LQG controller.

**References**


