Design Criteria for Uncertain Models with Structured and Unstructured Uncertainties*

Ali H. Sayed and Vitor H. Nascimento

Adaptive and Nonlinear Systems Laboratory, Electrical Engineering Department, University of California, Los Angeles, CA 90024

Abstract. This paper introduces and solves a weighted game-type cost criterion for estimation and control purposes that allows for a general class of uncertainties in the model or data. Both structured and unstructured uncertainties are allowed, including some special cases that have been used in the literature. The optimal solution is shown to satisfy an orthogonality condition similar to least-squares designs, except that the weighting matrices need to be modified in a certain optimal manner. One particular application in the context of state regulation for uncertain state-space models is considered. It is shown that in this case, the solution leads to a control law with design equations that are similar in nature to LQR designs. The gain matrix, however, as well as the Riccati variable, turn out to be state-dependent in a certain way. Further applications of these game-type formulations to image processing, estimation, and communications are discussed in [1–3].

1 INTRODUCTION

This paper develops a technique for estimation and control purposes that is suitable for models with bounded data uncertainties. The technique will be referred to as a BDU design method for brevity, and it expands on earlier works in the companion articles [1–4]. It is based on a constrained game-type formulation that allows the designer to explicitly incorporate into the problem statement a-priori information about bounds on the sizes of the uncertainties in the model. A key feature of the BDU formulation is that geometric insights (such as orthogonality conditions and projections), which are widely appreciated for classical quadratic-cost designs, can be pursued in this new framework. This geometric viewpoint was discussed at some length in the article [1] for a special case of the new cost function that we introduce in this paper.

The optimization problem (1) that we pose and solve here is of independent interest in its own right and it can be applied in several contexts.

* This material was based on work supported in part by the National Science Foundation under Award No. CCR-9732376. The work of V. H. Nascimento was also supported by a fellowship from CNPq - Brazil, while on leave from Escola Politécnica da Universidade de São Paulo.

Examples to this effect can be found in [1] where similar costs were applied to problems in image restoration, image separation, array signal processing, and estimation. Later in this paper we shall discuss one additional application in the context of state regulation for state-space models with parametric uncertainties.

2 FORMULATION OF THE BDU PROBLEM

We start by formulating a general optimization problem with uncertainties in the data. Thus consider the cost function

$$J(x,y) = x^T Q x + R(x,y) \ ,$$

where $x^T Q x$ is a regularization term, while the residual cost $R(x,y)$ is defined by

$$R(x,y) \stackrel{\Delta}{=} \left(Ax - b + Hy\right)^T W \left(Ax - b + Hy\right) .$$

Here, $Q > 0$ and $W \geq 0$ are given Hermitian weighting matrices, x is an $n-$dimensional column vector, A is an $N \times n$ known or nominal matrix, b is an $N \times 1$ known or nominal vector, H is an $N \times m$ known matrix, and y denotes an $m \times 1$ unknown perturbation vector. We now consider the problem of solving:

$$\hat{x} = \arg\min_{x} \max_{\|y\| \leq \phi(x)} \ J(x,y) \ , \tag{1}$$

where the notation $\|\cdot\|$ stands for the Euclidean norm of its vector argument or the maximum singular value of its matrix argument. The non-negative function $\phi(x)$ is a known bound on the perturbation y and it is only a function of x (it can be linear or nonlinear).

Problem (1) can be regarded as a constrained two-player game problem, with the designer trying to pick an $\hat{x}$ that minimizes the cost while the opponent $\{y\}$ tries to maximize the cost. The game problem is constrained since it imposes a limit (through $\phi(x)$) on how large (or how damaging) the opponent can be. Observe further that the strength of the opponent can vary with the choice of x.

2.1 Special Cases

The formulation (1) allows for both structured and unstructured uncertainties in the data. Before proceeding to its solution, let us exhibit two special cases. Consider first the problem

$$\min_{x} \max_{\substack{\|\delta A\| \leq \eta \\ \|\delta b\| \leq \eta_b}} \left[x^T Q x + \Big((A + \delta A)x - (b + \delta b) \Big)^T W \Big((A + \delta A)x - (b + \delta b) \Big) \right]$$

where $\{\delta A\}$ denotes an $N \times n$ perturbation matrix to the nominal matrix A, and δb denotes an $N \times 1$ perturbation vector to the nominal vector b. We showed in the companion article [3] that the above problem is equivalent to one of the following form:

$$\min_{x} \max_{\|y\| \leq \eta \|x\| + \eta_b} \left[x^T Q x + \Big(Ax - b + y) \Big)^T W \Big(Ax - b + y \Big) \right] ,$$

which is a special case of (1), with $H = I$ and $\phi(x) = \eta \|x\| + \eta_b$. In this example, the uncertainties $\{\delta A, \delta b\}$ are not related in any way and we shall say that they are unstructured. The special case $Q = 0$ and $W = I$ was treated in [1,4,5]. In particular, a geometric framework was developed in [1] for such problems that is similar in nature to the geometry of least-squares problems. We shall comment briefly on this aspect further ahead. On the other hand, reference [5] solves the case $Q = 0$ and $W = I$ by using LMI techniques, which for this particular problem turn out to be more costly than the direct solution methods proposed in [1,4]. When W is non-unity, the problem becomes more rich, and also more involved, even when $Q = 0$.

Consider now the alternative problem

$$\min_{x} \max_{\substack{\delta A \\ \delta b}} \left[x^T Q x + \Big((A + \delta A)x - (b + \delta b) \Big)^T W \Big((A + \delta A)x - (b + \delta b) \Big) \right]$$

where the perturbations $\{\delta A, \delta b\}$ are now assumed to be generated by a model of the form

$$\begin{bmatrix} \delta A & \delta b \end{bmatrix} = HS \begin{bmatrix} E_a & E_b \end{bmatrix} , \tag{2}$$

where S is a contraction, $\|S\| \leq 1$, and $\{H, E_a, E_b\}$ are known. Then it can be easily seen that this problem is equivalent to the following

$$\min_{x} \max_{\|y\| \leq \|E_a x - E_b\|} \left[x^T Q x + \Big(Ax - b + Hy) \Big)^T W \Big(Ax - b + Hy \Big) \right] ,$$

which is again a special case of (1) with $\phi(x) = \|E_a x - E_b\|$. Here, the perturbations $\{\delta A, \delta b\}$ are related (for example, they both lie in the range space of H). We shall say that they are structured. Such structured perturbations have been used in robust control design (see, e.g., [6]).

The formulation (1) that we consider in this paper is more general in that it allows for other classes of perturbations through the choice of the function $\phi(x)$.

3 SOLUTION OF THE BDU PROBLEM

We now proceed to the solution of (1). It turns out that the derivation given in [3] for a special case of (1) extends to this more general scenario with the appropriate modifications.

First we note that for any given y, the residual cost $R(x,y)$ is convex in x. Therefore, the maximum

$$C(x) \stackrel{\Delta}{=} \max_{\|y\| \leq \phi(x)} R(x,y) \ , \tag{3}$$

is a convex function in x. Now since $x^T Q x$ is strictly convex in x when $Q > 0$, we conclude that $x^T Q x + C(x)$ is strictly convex in x, which shows that problem (1) has a unique global minimum $\hat{x}$.[1] To determine $\hat{x}$ we proceed in steps.

3.1 The Maximization Problem

We now solve (3) for any fixed x. Note first that the cost $R(x,y)$ is convex in y, so that the maximum over y is achieved at the boundary, $\|y\| = \phi(x)$. We can therefore replace the inequality constraint in (3) by an equality. Introducing a Lagrange multiplier λ, the solution to (3) can then be found from the unconstrained problem:

$$\max_{y,\lambda} \left[\big(Ax - b + Hy\big)^T W \big(Ax - b + Hy\big) - \lambda\big(\|y\|^2 - \phi^2(x)\big) \right] . \tag{4}$$

Note that since the original problem has an inequality constraint, the Lagrange multiplier must be nonnegative: $\lambda \geq 0$ [7]. Differentiating (4) with respect to y and λ, and denoting the optimal solutions by $\{y^o, \lambda^o\}$, we obtain the equations

$$(\lambda^o I - H^T W H) y^o = H^T W (Ax - b) \ , \qquad \|y^o\| = \phi(x) \ . \tag{5}$$

It turns out that the solution λ^o should satisfy $\lambda^o \geq \|H^T W H\|$. This is because the Hessian of the cost in (4) w.r.t y must be nonpositive-definite [7].[2] We should further stress that the solutions $\{y^o, \lambda^o\}$ are functions of x and we shall therefore sometimes write $\{y^o(x), \lambda^o(x)\}$.

At this stage, we do not need to solve the equations (5) for $\{y^o, \lambda^o\}$. It is enough to know that the optimal $\{y^o, \lambda^o\}$ satisfy (5).[3] Using this fact, we can verify that the maximum cost in (4) is equal to

$$\begin{aligned} C(x) = (Ax - b)^T &\left[W + WH\big(\lambda^o(x) I - H^T W H\big)^\dagger H^T W \right] (Ax - b) \\ &+ \lambda^o(x) \phi^2(x) \ , \end{aligned} \tag{6}$$

where $X^\dagger$ denotes the pseudo-inverse of X.

[1] It can be easily seen that in the special case $\phi(0) = 0$ and $Wb = 0$, the unique solution of (1) is $\hat{x} = 0$. In the sequel we shall therefore assume that $\phi(0)$ and Wb are not zero simultaneously.

[2] We refer to the case $\lambda^o = \|H^T W H\|$ as the singular case, while $\lambda^o > \|H^T W H\|$ is called the regular case. Both cases are handled simultaneously in our framework through the use of the pseudo-inverse notation.

[3] In fact, we can show that the solution λ^o is always unique while there might be several y^o.

3.2 The Minimization Problem

The original problem (1) is therefore equivalent to:

$$\min_x \left[x^T Q x \; + \; C(x) \right] \; . \tag{7}$$

However, rather than minimizing the above cost over n variables, which are the entries of the vector x, we shall instead show how to reduce the problem to one of minimizing a certain cost function over a single scalar variable (see (14) further ahead). For this purpose, we introduce the following function of two independent variables x and λ,

$$C(x, \lambda) = (Ax - b)^T \left[W + WH \left(\lambda I - H^T W H \right)^\dagger H^T W \right] (Ax - b) + \lambda \phi^2(x) \; .$$

Then it can be verified, by direct differentiation with respect to λ and by using the expression for $\lambda^o(x)$ from (5), that

$$\lambda^o(x) \; = \; \arg \min_{\lambda \geq \|H^T W H\|} \; C(x, \lambda) \; .$$

This means that problem (1) is equivalent to

$$\min_{\lambda \geq \|H^T W H\|} \; \min_x \left[x^T Q x \; + \; C(x, \lambda) \right] \; . \tag{8}$$

The cost function in the above expression, viz., $J(x, \lambda) = x^T Q x + C(x, \lambda)$, is now a function of two independent variables $\{x, \lambda\}$. This should be contrasted with the cost function in (7). Now define, for compactness of notation, the quantities $M(\lambda) = Q + A^T W(\lambda) A$ and $d(\lambda) = A^T W(\lambda) b$, where

$$W(\lambda) = W + WH \left(\lambda I - H^T W H \right)^\dagger H^T W \; .$$

To solve problem (8), we first search for the minimum over x for every fixed value of λ, which can be done by setting the derivative of $J(x, \lambda)$ w.r.t. x equal to zero. This shows that any minimum x must satisfy the equality

$$M(\lambda) x + \frac{1}{2} \lambda \nabla \phi^2(x) = d(\lambda) \; , \tag{9}$$

where $\nabla \phi^2(x)$ is the gradient of $\phi^2(x)$ w.r.t. x.

Special Cases

Let us reconsider the special cases $\phi(x) = \|E_a x - E_b\|$ and $\phi(x) = \eta \|x\| + \eta_b$. For the first choice we obtain $\nabla \phi^2(x) = 2 E_a^T \left(E_a x - E_b \right)$ so that the solution of Eq. (9), which is dependent on λ, becomes

$$x^o(\lambda) = \left[M(\lambda) + \lambda E_a^T E_a \right]^{-1} \left(d(\lambda) + \lambda E_a^T E_b \right) . \tag{10}$$

The second choice, $\phi(x) = \eta\|x\| + \eta_b$, was studied in the companion article [3]. In this case, solving for x^o is not so immediate since Eq. (9) now becomes, for any nonzero x,

$$x = \left[M(\lambda) + \lambda\eta\left(\eta + \frac{\eta_b}{\|x\|}\right)\right]^{-1} d(\lambda)\ . \tag{11}$$

Note that x appears on both sides of the equality (except when $\eta_b = 0$, in which case the expression for x is complete in terms of $\{M, \lambda, \eta, d\}$). To solve for x in the general case we define $\alpha = \|x\|^2$ and square the above equation to obtain the scalar equation in α:

$$\alpha^2 - d^T(\lambda)\left[M(\lambda) + \lambda\eta\left(\eta + \frac{\eta_b}{\alpha}\right)\right]^{-2} d(\lambda)\ = 0\ . \tag{12}$$

It can be shown that a unique solution $\alpha^o(\lambda) > 0$ exists for this equation if, and only if, $\lambda\eta\eta_b\ < \|d(\lambda)\|^2$. Otherwise, $\alpha^o(\lambda) = 0$. In the former case, the expression for x^o, which is a function of λ, becomes

$$x^o(\lambda) = \left[M(\lambda) + \lambda\eta\left(\eta + \frac{\eta_b}{\alpha^o(\lambda)}\right)\right]^{-1} d(\lambda)\ . \tag{13}$$

In the latter case we clearly have $x^o(\lambda) = 0$.

The General Case

Let us assume that (9) has a unique solution $x^o(\lambda)$, as was the case with the above two special cases. This will also be always the case whenever $\phi(x)$ is a differentiable and strictly convex function (since then $J(x, \lambda)$ will be differentiable and strictly convex in x). We thus have a procedure that allows us to determine the minimizing x^o for every λ. This in turn allows us to re-express the resulting cost $J(x^o(\lambda), \lambda)$ as a function of λ alone, say $G(\lambda) = J(x^o(\lambda), \lambda)$. In this way, we conclude that the solution $\hat{x}$ of the original optimization problem (1) can be solved by determining the $\hat{\lambda}$ that solves

$$\min_{\lambda \geq \|H^T W H\|} G(\lambda)\ , \tag{14}$$

and by taking the corresponding $x^o(\hat{\lambda})$ as $\hat{x}$. That is, $\hat{x}$ solves (9) when $\lambda = \hat{\lambda}$. We summarize the solution in the following statement.

Theorem 1 (Solution). *The unique global minimum of (1) can be determined as follows. Introduce the cost function*

$$G(\lambda) = x^{oT}(\lambda) Q x^o(\lambda) + C[x^o(\lambda), \lambda]\ , \tag{15}$$

where $x^o(\lambda)$ is the unique solution of (9)*. Let $\hat{\lambda}$ denote the minimum of $G(\lambda)$ over the interval $\lambda \geq \|H^T W H\|$. Then the optimum solution of* (1) *is $\hat{x} = x^o(\hat{\lambda})$.*

$\diamondsuit$

We thus see that the solution of (1) requires that we determine an optimal scalar parameter $\hat{\lambda}$, which corresponds to the minimizing argument of a certain nonlinear function $G(\lambda)$ (or, equivalently, to the root of its derivative function). This step can be carried out very efficiently by any root finding routine, especially since the function $G(\lambda)$ is well defined and, moreover, $\hat{\lambda}$ is unique. We obtain as corollaries the following two special cases.

Corollary 1 (Structured Uncertainties). *When $\phi(x) = \|E_a x - E_b\|$, $x^o(\lambda)$ is given by* (10) *and the global minimum of (1) becomes*

$$\hat{x} = \left[\hat{Q} + A^T \hat{W} A\right]^{-1} A^T \hat{W} b \ \stackrel{\Delta}{=}\ Kb \tag{16}$$

where $\hat{Q} = Q + \hat{\lambda} E_a^T E_a$ and $\hat{W} = W + WH(\hat{\lambda} I - H^T W H)^{\dagger} H^T W$.

$\diamondsuit$

Corollary 2 (Unstructured Uncertainties). *When $\phi(x) = \eta\|x\| + \eta_b$, $x^o(\lambda)$ is given by (13) if $\lambda\eta\eta_b < \|d(\lambda)\|^2$ (otherwise it is zero). Moreover, $\alpha^o(\lambda)$ in (13) is the unique positive root of (12). Let $\hat{\lambda}$ denote the minimum of $G(\lambda)$ over the interval $\lambda \geq \|W\|$. Then*

$$\hat{x} = \left[\hat{Q} + A^T \hat{W} A\right]^{-1} A^T \hat{W} b \ \stackrel{\Delta}{=}\ Kb\,, \tag{17}$$

if $\hat{\lambda}\eta\eta_b < \|d(\hat{\lambda})\|^2$ (otherwise $\hat{x} = 0$), where

$$\hat{Q} = Q + \hat{\lambda}\eta\left(\eta + \frac{\eta_b}{\alpha^o(\hat{\lambda})}\right) I\,, \quad \hat{W} = W + W(\hat{\lambda} I - W)^{\dagger} W\,.$$

$\diamondsuit$

3.3 The Orthogonality Condition

Observe that the optimal solution $\hat{x}$ in the above cases satisfies an orthogonality condition of the form $\hat{Q}\hat{x} + A^T \hat{W}(A\hat{x} - b) = 0$, for some $\{\hat{Q}, \hat{W}\}$. Compared with the solution to the standard regularized least-squares problem,

$$\min_x \left[x^T Q x + (Ax - b)^T W (Ax - b)\right]\,,$$

whose unique solution satisfies $Q\hat{x} + A^T W(A\hat{x} - b) = 0$, we see that the solution to the BDU problem satisfies a similar orthogonality condition, with the given weighting matrices $\{Q, W\}$ replaced by new matrices $\{\hat{Q}, \hat{W}\}$! To determine the necessary corrections to $\{Q, W\}$, one determines the optimal scalar $\hat{\lambda}$ from the minimization (14). The convenience of such a geometric viewpoint is discussed in [1] for the special case $Q = 0$ and $W = I$.

4 APPLICATION TO STATE REGULATION

As mentioned earlier, the BDU cost functions can be useful in different contexts, including image restoration, image separation, array signal processing, and estimation (see [1] for some examples). Here we discuss another application for the weighted BDU problem in the context of state regulation for state-space models with parametric uncertainties.

Thus consider the linear state-space model $x_{i+1} = F_i x_i + G_i u_i$, where x_0 denotes the value of the initial state, and the $\{u_i\}$ denote the control (input) sequence. The classical linear quadratic regulator (LQR) problem seeks a control sequence $\{u_i\}$ that regulates the state vector towards zero while keeping the control cost low. This is achieved as follows. Introduce, for compactness of notation, the local cost

$$V_i(x_{i+1}, u_i) \stackrel{\Delta}{=} \left(x_{i+1}^T R_{i+1} x_{i+1} + u_i^T Q_i u_i\right) \ , \quad R_{N+1} = P_{N+1} \ .$$

Then the optimal control is determined by solving

$$\min_{\{u_0, u_1, \ldots, u_N\}} \left(x_{N+1}^T P_{N+1} x_{N+1} \ + \ \sum_{j=0}^{N} \left[u_j^T Q_j u_j + x_j^T R_j x_j\right] \right) \ ,$$

with $Q_j > 0$, $R_j \geq 0$, and $P_{N+1} \geq 0$. We shall write the above problem more compactly as (note that x_0 does not really affect the solution):

$$x_0^T R_0 x_0 \ + \ \min_{\{u_0, u_1, \ldots, u_N\}} \ (V_0 + V_1 + \ldots + V_N) \ , \tag{18}$$

It is well known that the LQR problem can be solved recursively by re-expressing the LQR cost as nested minimizations of the form:

$$x_0^T R_0 x_0 \ + \ \min_{u_0} \left\{ V_0 + \min_{u_1} \left\{ V_1 + \ldots + \min_{u_N} \{V_N\} \right\} \right\}, \tag{19}$$

where only the last term, through the state-equation for x_{N+1}, is dependent on u_N. Hence we can determine $\hat{u}_N$ by solving

$$\min_{u_N} V_N \ , \quad \text{given} \ \ x_N \ , \tag{20}$$

and then progress backwards in time to determine the other control values. By carrying out this argument one finds the well-known state-feedback solution:

$$\begin{cases} \hat{u}_i = -K_i x_i \ , \\ K_i = (Q_i + G_i^T P_{i+1} G_i)^{-1} G_i^T P_{i+1} F_i \ , \\ P_i = R_i + K_i^T Q_i K_i + (F_i - G_i K_i)^T P_{i+1} (F_i - G_i K_i) \ . \end{cases} \tag{21}$$

It is well known that the above LQR controller is sensitive to modeling errors. Robust design methods to ameliorate these sensitivity problems include the $\mathcal{H}_\infty$ design methodology (e.g., [8–11]) and the so-called guaranteed-cost designs (e.g., [12–14]). We suggest below a procedure that is based on the BDU problem solved above. At the end of this exposition, we shall compare our result with a guaranteed-cost design. [A comparison with an $\mathcal{H}_\infty$ design is given in [1] for a special first-order problem.]

4.1 State Regulation

Consider now the state-equation with parametric uncertainties:

$$x_{i+1} = (F_i + \delta F_i) x_i + (G_i + \delta G_i) u_i \ , \tag{22}$$

with known x_0, and where the uncertainties $\{\delta F_i, \delta G_i\}$ are assumed to be generated via

$$\begin{bmatrix} \delta F_i \ \delta G_i \end{bmatrix} = HS \begin{bmatrix} E_f \ E_g \end{bmatrix} \ , \tag{23}$$

for known H, E_f, E_g, and for any contraction $\|S\| \leq 1$. The solution of the case with unstructured uncertainties $\{\delta F_i, \delta G_i\}$, say $\|\delta F_i\| \leq \eta_{f,i}$ and $\|\delta G_i\| \leq \eta_{g,i}$, is very similar and is treated in [3]. We focus here on the above structured case (23) for the sake of demonstration. Still, we should mention that by choosing different $\phi(x)$, the approach described in the earlier sections can handle other classes of uncertainties as well.

Consider the problem of determining a control sequence $\{\hat{u}_j, 0 \leq j \leq N\}$ that solves the nested min-max optimizations:

$$x_0^T R_0 x_0 \ + \ \min_{u_0} \max_{\substack{\delta F_0 \\ \delta G_0}} \left\{ V_0 + \min_{u_1} \max_{\substack{\delta F_1 \\ \delta G_1}} \left\{ V_1 + \ldots + \min_{u_N} \max_{\substack{\delta F_N \\ \delta G_N}} \left\{ V_N \right\} \right\} \right. \tag{24}$$

where we are writing, for compactness of notation, $\{\delta F_i, \delta G_i\}$ under the max symbols instead of the complete notation.

In order to illustrate the structure of the solution of (24), let us consider the simple case $N = 1$, viz.,

$$x_0^T R_0 x_0 \ + \ \min_{u_0} \max_{\substack{\delta F_0 \\ \delta G_0}} \left\{ V_0 + \min_{u_1} \max_{\substack{\delta F_1 \\ \delta G_1}} \left\{ V_1 \right\} \right\} \ . \tag{25}$$

To be even more explicit, recall that V_0 is a function of $\{x_1, u_0\}$ while V_1 is a function of $\{x_2, u_1\}$. Hence, V_0 is a function of $\{x_0, u_0, \delta F_0, \delta G_0\}$ and we shall denote this explicitly as $V_0(x_0, u_0, \delta F_0, \delta G_0)$. Likewise, we shall write $V_1(x_0, u_0, \delta F_0, \delta G_0, u_1, \delta F_1, \delta G_1)$. If we now solve the inner-most min-max problem in (25), for any $\{u_0, \delta F_0, \delta G_0\}$, i.e.,

$$\min_{u_1} \max_{\substack{\delta F_1 \\ \delta G_1}} \; V_1(x_0, u_0, \delta F_0, \delta G_0, u_1, \delta F_1, \delta G_1) \; , \tag{26}$$

we obtain a representation for the solution $\{\hat{u}_1, \widehat{\delta F}_1, \widehat{\delta G}_1\}$ in terms of the unknowns $\{x_0, u_0, \delta F_0, \delta G_0\}$. That is, we find

$$\begin{aligned} \hat{u}_1 &= f_1(x_0, u_0, \delta F_0, \delta G_0) \; , \\ \widehat{\delta F}_1 &= g_1(x_0, u_0, \delta F_0, \delta G_0) \; , \\ \widehat{\delta G}_1 &= h_1(x_0, u_0, \delta F_0, \delta G_0) \; , \end{aligned}$$

for some functions $\{f_1(\cdot), g_1(\cdot), h_1(\cdot)\}$. The resulting cost in (26) will also be a function of $\{x_0, u_0, \delta F_0, \delta G_0\}$, say

$$V_1^*(x_0, u_0, \delta F_0, \delta G_0\} \;=\; \min_{u_1} \max_{\substack{\delta F_1 \\ \delta G_1}} \; V_1(x_0, u_0, \delta F_0, \delta G_0, u_1, \delta F_1, \delta G_1) \; .$$

Returning to (25), we now solve the outer-most min-max problem over $\{u_0, \delta F_0, \delta G_0\}$,

$$x_0^T R_0 x_0 \;+\; \min_{u_0} \max_{\substack{\delta F_0 \\ \delta G_0}} \{V_0 + V_1^*\} \; ,$$

which would then lead to a representation for $\{\hat{u}_0, \widehat{\delta F}_0, \widehat{\delta G}_0\}$ in terms of x_0,

$$\hat{u}_0 = f_0(x_0) \; , \quad \widehat{\delta F}_0 = g_0(x_0) \; , \quad \widehat{\delta G}_0 = h_0(x_0) \; .$$

Therefore, the optimal control values that solve (25) will be

$$\hat{u}_0 = f_0(x_0) \quad \text{and} \quad \hat{u}_1 = f_1(x_0, \hat{u}_0, \widehat{\delta F}_0, \widehat{\delta G}_0) \; ,$$

where the arguments of $f_1(\cdot)$ are now defined in terms of $\{\hat{u}_0, \widehat{\delta F}_0, \widehat{\delta G}_0\}$.

If we thus reconsider the original problem (24), and focus first on the inner-most optimization, say

$$\min_{u_N} \max_{\substack{\delta F_N, \delta G_N \\ \text{subject to (23)}}} \left[u_N^T Q_N u_N + x_{N+1}^T P_{N+1} x_{N+1} \right] ,$$

then the above argument shows that in order to determine an expression for $\hat{u}_N$ from the above, the state vector x_N has to be taken as $\hat{x}_N$, which is the value that would result had the earlier optimal control signals $\{\hat{u}_j, 0 \leq$

$j \leq N-1\}$ been determined already and using the worst-case disturbances. Then expanding the term $x_{N+1}^T P_{N+1} x_{N+1}$ by using the state equation for x_{N+1},

$$x_{N+1} = (F_N + \delta F_N)\hat{x}_N + (G_N + \delta G_N)u_N \ ,$$

the above problem reduces to a problem of the same form as the structured BDU problem (2) that we considered before with the identifications:

$$A \leftarrow G_N \ , \ \ W \leftarrow P_{N+1} \ , \ \ Q \leftarrow Q_N \ , \ \ H \leftarrow \ H \ , \quad b \leftarrow -F_N\hat{x}_N \ ,$$

$$x \leftarrow u_N, \ \ E_a \leftarrow E_g, \ \ E_b \leftarrow E_f\hat{x}_N \ , \ \ \delta A \leftarrow \delta G_N, \ \delta b \leftarrow -\delta F_N\hat{x}_N.$$

Using (16), and the above identifications, we conclude that the optimal control value $\hat{u}_N$ is given by (compare with the LQR recursions)

$$\begin{cases} \hat{u}_N = -K_N\hat{x}_N \ , \\ K_N = \left(\hat{Q}_N + G_N^T\hat{W}_{N+1}G_N\right)^{-1} G_N^T\hat{W}_{N+1}F_N \ , \\ \hat{Q}_N = Q_N + \hat{\lambda}_N E_g^T E_g \ , \\ \hat{W}_{N+1} = P_{N+1} + P_{N+1}H\left(\hat{\lambda}_N I - H^T P_{N+1} H\right)^{\dagger} H^T P_{N+1} \ , \end{cases}$$

where $\hat{\lambda}_N$ is the optimal parameter that corresponds to the above data $\{A, b, W, Q, H, E_a, E_b\}$, and which can be found as explained in Thm. 1 (or Cor. 1).

Moreover, using (6)–(7) and the above identifications again, we find that

$$\hat{x}_N^T R_N \hat{x}_N + \left(\min_{u_N} \max_{\substack{\delta F_N \\ \delta G_N}} V_N \right) \ = \ \hat{x}_N^T P_N \hat{x}_N \ ,$$

where P_N is given by (compare with the Riccati recursion in (21)):

$$\begin{aligned} P_N = R_N + K_N^T Q_N K_N + (F_N - G_N K_N)^T \hat{W}_{N+1} (F_N - G_N K_N) + \\ + \hat{\lambda}_N \left[K_N^T E_g^T E_g K_N - K_N^T E_g^T E_f - E_f^T E_g K_N + E_f^T E_f\right] \ . \end{aligned} \quad (27)$$

We now proceed to determine an approximation for the optimal control value at time $N-1$ by solving

$$\min_{u_{N-1}} \max_{\delta F_{N-1}, \delta G_{N-1}} \left[u_{N-1}^T Q_{N-1} u_{N-1} + x_N^T P_N x_N \right] ,$$

where we assume that $\hat{x}_{N-1}$ is available. We take the solution as $\hat{u}_{N-1}$, and so on. Note that this step is an approximation because we are employing the P_N found above, which is a function of $\hat{x}_N$. For optimality, we would need to determine the functional form $P_N(x_N)$ — this form is defined by the same equations as above with x_N replacing $\hat{x}_N$. It turns out that for single-state models, the value of P_N is independent of the state and therefore the above $\hat{u}_{N-1}$ agrees with the optimal value — see further ahead.

Remarks

Several remarks are due now.

1. The control values $\{\hat{u}_i\}$ found above are in terms of the worst-case state $\hat{x}_i$, which we show how to evaluate in the next section.

2. Compared with the solution to the LQR problem we see that there are three main differences in the recursions. First, the gain matrix K_N is not defined directly in terms of the original quantities $\{Q_N, P_{N+1}\}$ but in terms of modified quantities $\{\hat{Q}_N, \hat{W}_{N+1}\}$. Secondly, the term P_{N+1} in the LQR Riccati recursion is replaced by $\hat{W}_{N+1}$ in (27), in addition to a new correction term that is equal to $\hat{\lambda}_N \phi^2(\hat{u}_N)$. Finally, the above solution in fact has the form of a two-point boundary value problem (TPBVP). This is because the expressions for $\{K_N, P_N\}$ are dependent on the worst-case state $\hat{x}_N$ (through $\hat{\lambda}_N$). We can denote this dependency more explicitly by writing, for any i,

$$\hat{u}_i = -K_i(\hat{x}_i)\hat{x}_i \ . \tag{28}$$

   A reasonable state-feedback implementation would be to choose $\hat{u}_i = -K_i(\hat{x}_i)x_i$ (see, e.g., [3] for a simulation in this case). We should mention that for single-state models, the state-dependency disappears (as we show in a later section).

3. Similar recursions and remarks are valid for the solution of problems with other kinds of uncertainties, e.g., unstructured uncertainties [3].

4.2 An Iterative Solution to the TPBVP

We are currently studying the TPBVP more closely. An iterative solution that we found performs reasonably well is the following.

<u>I. Initialization</u>. Choose initial values for all variables P_0 to P_N (for example, by running the LQR Riccati recursion or by using a suboptimal guaranteed-cost design). Choose also initial values for all $\hat{\lambda}_i$, say $\hat{\lambda}_i > \|H^T P_{i+1} H\|_{\mathrm{F}}$.

<u>II. Forwards Iteration</u>. Given values $\{x_0, P_{i+1}, \hat{\lambda}_i\}$, we evaluate the quantities $\{\hat{W}_{i+1}, \hat{Q}_i, K_i, \hat{y}_i, \hat{u}_i\}$ by using the recursions derived above, as well as propagate the state-vectors $\{\hat{x}_i\}$ by using $\hat{x}_{i+1} = F_i\hat{x}_i + G_i\hat{u}_i + \hat{y}_i$ where, from (5), $\hat{y}_i$ is found by solving the equation

$$\left(\hat{\lambda}_i I - H^T P_{i+1} H\right)\hat{y}_i = H^T P_{i+1}(F_i\hat{x}_i + G_i\hat{u}_i) \ .$$

If the matrix $\left(\hat{\lambda}_i I - H^T P_{i+1} H\right)$ is singular, then among all possible solutions we choose one that satisfies $\|\hat{y}_i\|^2 = \|E_f \hat{x}_i + E_g \hat{u}_i\|^2$.

III. Backwards Iteration. Given values $\{P_{N+1}, \hat{u}_i, \hat{x}_i\}$ we find new approximations for $\{P_i, \hat{\lambda}_i\}$ by using the recursions derived above for the state regulation problem.

IV. Recursion. Repeat steps II and III.

$\diamondsuit$

4.3 The One-Dimensional Case

Several simplifications occur for one-dimensional systems. In particular, there is no need to solve a two-point boundary value problem. This is because for such models, the state-dependency in the recursions disappears and we can therefore *explicitly* describe the optimal control law.

To show this, let us verify that the value of $\hat{\lambda}_N$ (and more generally $\hat{\lambda}_i$) becomes independent of $\hat{x}_N$ ($\hat{x}_i$). Indeed, recall from (15) that $\hat{\lambda}_N$ is the argument that minimizes

$$G(\lambda) = \left[K_N^2 Q_N + (F_N - G_N K_N)^2 \left(W_{N+1} + E_g^2 K_N^2 + E_f^2\right)\right] \hat{x}_N^2 \ ,$$

over $\lambda \geq \|H^2 P_{N+1}\|$. Here

$$W_{N+1} = P_{N+1} + P_{N+1}^2 H^2 (\lambda - H^2 P_{N+1})^\dagger \ .$$

Therefore, the minimum of $G(\lambda)$ is independent of $\hat{x}_N$. It then follows that $\hat{\lambda}_N$ and K_N are independent of $\hat{x}_N$ and we can iterate the recursion for P_N backwards in time.

4.4 A Simulation

We compare below in Fig. 1 the performance of this design with a guaranteed-cost design. The example presented here is of a 2-state system. The nominal model is stable with only one control variable. Moreover, $H = I$, $E_f = 0$, $E_g = \begin{bmatrix} 0 & 0 & 0.4 \end{bmatrix}^T$, $G = \begin{bmatrix} 1 & -0.5 \end{bmatrix}^T$, and $N = 20$. The lower horizontal line is the worst-case cost that is predicted by our BDU construction. The upper horizontal line is an upper bound on the optimal cost. It is never exceeded by the guaranteed-cost design. The situation at the right-most end of the graph corresponds to the worst-case scenario. Observe (at the right-end of the graph) the improvement in performance in the worst-case.

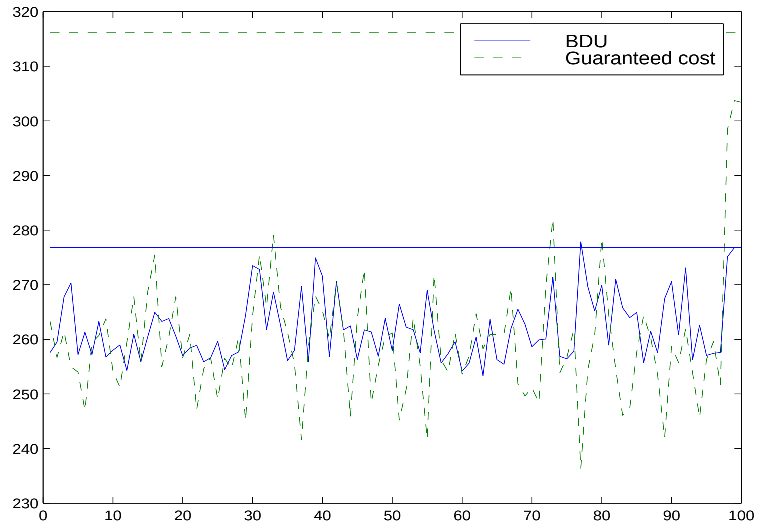

Fig. 1. 100 random runs with a stable 2-dimensional nominal model

5 CONCLUDING REMARKS

Regarding the state-regulator application, earlier work in the literature on guaranteed-cost designs found either sub-optimal steady-state and finite-horizon controllers (e.g., [13]), or optimal steady-state controllers *over the class of linear control laws* [12]. Our solution has the following properties: i) It has a geometric interpretation in terms of an orthogonality condition with modified weighting matrices, ii) it does not restrict the control law to linear controllers, iii) it also allows for unstructured and other classes of uncertainties (see [3]), and iv) it handles both regular and degenerate situations. We are currently studying these connections more closely, as well as the TPBVP.

In this paper we illustrated one application of the BDU formulation in the context of state regulation. Other applications are possible [1].

Acknowledgment. The authors would like to thank Prof. Jeff Shamma of the Mechanical and Aerospace Engineering Department, UCLA, for his careful comments and feedback on the topic of this article.

References

1. A. H. Sayed, V. H. Nascimento, and S. Chandrasekaran. Estimation and control with bounded data uncertainties. *Linear Algebra and Its Applications*, vol. 284, pp. 259–306, Nov. 1998.

2. A. H. Sayed and S. Chandrasekaran. Estimation in the presence of multiple sources of uncertainties with applications. *Proc. Asilomar Conference*, vol. 2, pp. 1811-1815, Pacific Grove, CA, Nov. 1998.
3. V. H. Nascimento and A. H. Sayed. Optimal state regulation for uncertain state-space models. In *Proc. ACC*, San Diego, CA, June 1999.
4. S. Chandrasekaran, G. Golub, M. Gu, and A. H. Sayed. Parameter estimation in the presence of bounded data uncertainties. *SIAM J. Matrix Analysis and Applications*, 19(1):235–252, Jan. 1998.
5. L. E. Ghaoui and H. Hebret. Robust solutions to least-squares problems with uncertain data, *SIAM J. Matrix Anal. Appl.*, vol. 18, pp. 1035–1064, 1997.
6. M. Fu, C. E. de Souza, and L. Xie. $\mathcal{H}_\infty$ estimation for uncertain systems. *Int. J. Robust and Nonlinear Contr.*, vol. 2, pp. 87–105, 1992.
7. R. Fletcher. *Practical Methods of Optimization*. Wiley, 1987.
8. M. Green and D. J. N. Limebeer. *Linear Robust Control*. Prentice-Hall, Englewood Cliffs, NJ, 1995.
9. K. Zhou, J. C. Doyle, and K. Glover. *Robust and Optimal Control*. Prentice-Hall, NJ, 1996.
10. P. P. Khargonekar, I. R. Petersen, and K. Zhou. Robust stabilization of uncertain linear systems: Quadratic stabilizability and $\mathcal{H}_\infty$ control theory. *IEEE Transactions on Automatic Control*, vol. 35, no. 3, 1990.
11. B. Hassibi, A. H. Sayed, and T. Kailath. *Indefinite Quadratic Estimation and Control: A Unified Approach to $\mathcal{H}_2$ and $\mathcal{H}_\infty$ Theories.* SIAM, PA, 1999.
12. S. O. R. Moheimani, A. V. Savkin, and I. R. Petersen. Minimax optimal control of discrete-time uncertain systems with structured uncertainty. *Dynamics and Control*, vol. 7, no. 1, pp. 5–24, Jan. 1997.
13. L. Xie and Y. C. Soh. Guaranteed-cost control of uncertain discrete-time systems. *Control Theory and Advanced Technology*, vo. 10, no. 4, pp. 1235–1251, June 1995.
14. A. V. Savkin and I. R. Petersen. Optimal guaranteed-cost control of discrete-time nonlinear uncertain systems. *IMA Journal of Mathematical Control and Information*, vol. 14, no. 4, pp. 319–332, Dec. 1997.
15. D. G. Luenberger. *Optimization by Vector Space Methods.* Wiley, 1969.
16. T. Basar and G. J. Olsder. *Dynamic Noncooperative Game Theory.* Academic Press, 1982.