

Unifying Robustness Analysis and System ID

John Doyle Matthew Newlin Fernando Paganini Jorge Tierno

California Institute of Technology
Pasadena, CA 91125

Abstract

A unified systems analysis framework is presented which includes conventional robustness analysis, model validation, and system identification as special cases and thus shows them to be instances of the same fundamental problem. A concrete version of this framework is developed for the linear case, based on a generalized structured singular value. This unification forms the basis for the use of common computational tools and a more natural interplay between modeling, identification, and robustness analysis.

1 Introduction

The main motivation behind this paper is the belief that control theory needs to play a broader role in technology; in fact, of the multiple technological problems involving mathematical tools of the dynamical systems theory, only a small fraction reduce to the design of a feedback system for a well defined plant. A broad class of problems in modeling, system identification, system design, simulation, and optimization are addressed with similar mathematical tools; a natural objective is therefore the development of a more unified theory, in which a common language of mathematical and computational machinery is used to perform the previous range of activities. Although this goal may seem ambitious, this paper documents some progress in this direction by exhibiting a framework that captures as special cases robustness analysis and system identification.

The connection between these two problems is that they both are special cases of the following analysis question:

Q: *Given a mathematical description in terms of equations involving uncertainty, do there exist values of the uncertainty in a given class such that the equations have a solution?*

For the case of robustness analysis, this question is specialized as follows. The mathematical description is a dynamical system with uncertainty, for example parametric or dynamic uncertainty. Loosely speaking, to solve a robust stability problem, which also encompasses many robust performance problems, is to test whether there exist uncertainty values for which loop equations admit nontrivial solutions, and is therefore a special case of answering **Q**.

Most activity in robustness analysis has focused on obtaining computable answers to this question for a rich variety of uncertainty descriptions, which can usually be fit into the Linear Fractional Transformation (LFT) and μ ([10]) paradigms. A class of perturbation structures which lead to particularly tractable computation are those which correspond to Integral Quadratic Constraints (IQCs [7]) on signals: in this case the

problem reduces to a convex feasibility problem, the solution of a Linear Matrix Inequality (LMI [1]). For tighter descriptions of uncertainty such as real parameters, worst case computation is provably NP hard ([2]), but available algorithms with good performance on typical problems ([17]) are encouraging.

A recent development in robustness analysis ([12]) which will be exploited in this paper, is that a larger class of analysis problems, involving uncertain systems and an arbitrary number of IQCs, can be formulated and solved in implicit form. This is related to the behavioral paradigm for system theory ([16]) and is described in Section 2 and references therein.

A large research field under the umbrella of control theory is the area of system identification, which obtains dynamical models from experimental data, and the related model validation problem which checks consistency of a model with data. Mathematically, this area has relied on extending methods of statistics, mainly time series analysis. A standard reference is [6]. Since these models are typically stochastic, it has been difficult to reconcile this theory with robust control, which relies on deterministic descriptions. Some recent work ([14, 8]) has attempted to bridge this gap.

The main argument to base this unification is the recognition that model validation/ID problems are special cases of **Q**: given a model and experimental data, do there exist values of the uncertainty (parameters, disturbances, etc.) that solve the equations? This is, strictly speaking, a model validation question; system identification involves additionally finding the parameter values, but this is often no harder—existence of a solution is usually shown by finding a solution.

It is shown in Section 3 that the implicit LFT formulation over constant matrices provides a natural framework in which to cast a large class of model validation/ID problems, and therefore is the unifying paradigm.

In addition to providing a conceptual framework in which to relate robustness analysis and identification, the formulation pursued here suggests that computational tools developed to deal with the implicit robustness analysis formulation could be applied to system identification. These tools are discussed in Section 4.

Finally, Section 5 contains an example where the conventional least squares identification problem is reviewed from this perspective.

2 Robustness Analysis in Implicit Form

We begin by considering an uncertain LFT system in implicit form, depicted in Figure 1.

In this formulation, $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ is in principle an arbitrary map, and Δ is a structured uncertainty operator, described below. The “manifest” variables w (signals of interest) are described implicitly by the above equations, as

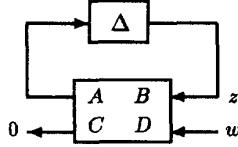


Figure 1: An Implicit LFT system.

$(\Delta \star M)w = 0$, where $\Delta \star M$ is the LFT

$$\Delta \star M = D + C\Delta(I - A\Delta)^{-1}B \quad (1)$$

and it is assumed that $(I - A\Delta)^{-1}$ is well defined. Note that there is no partition of the manifest variables into inputs and outputs. In this respect, this formulation is consistent with the behavioral framework for system theory, introduced by Willems ([16]). References [3, 4] contain motivation and introductory material on these representations.

An "internal" description of the same system in terms of the generalized state z is

$$\varphi(\Delta, M) \begin{bmatrix} z \\ w \end{bmatrix} = 0, \quad \varphi(\Delta, M) = \begin{bmatrix} I - \Delta A & -\Delta B \\ C & D \end{bmatrix} \quad (2)$$

As in standard robust control, Δ varies in Δ which is the class of structured uncertainty operators of the form

$$\Delta = \text{diag}[\delta_1 I_{r_1}, \dots, \delta_L I_{r_L}, \Delta_{L+1}, \dots, \Delta_{L+F}] \quad (3)$$

where the blocks in Δ represent real parameters or dynamic (linear time invariant (LTI), linear time varying, or nonlinear) perturbations. Usually these perturbations are restricted to a normalized ball of uncertainty $\mathbf{B}_\Delta = \{\Delta : \|\Delta\| \leq 1\}$ in some operator norm.

In addition to allowing the representation of any standard input-output uncertain LFT system, the implicit representation allows the formulation of "over constrained" problems, which have more equations than free variables w . A special case of this is considered in [12], where it is shown that a finite number of IQCs on w can be given a kernel representation as in Figure 1. An example of the use of such additional constraints is the case of "whiteness" constraints to pose a Robust \mathcal{H}_2 performance problem (see [12]). Therefore a richer class of robust performance analysis problems can be formulated in this paradigm and converted to a robust stability test in the sense described below. The following definition is from [13]:

Definition 1 Consider the implicit system (2), where A, B, C, D , and Δ are linear operators in a vector valued \mathcal{L}_2 space. The system has robust \mathcal{L}_2 stability if for each $\Delta \in \mathbf{B}_\Delta$,

$$\inf \left\{ \left\| \varphi(\Delta, M) \begin{bmatrix} z \\ w \end{bmatrix} \right\| : z, w \in \mathcal{L}_2, \left\| \begin{bmatrix} z \\ w \end{bmatrix} \right\| = 1 \right\} > 0 \quad (4)$$

According to the definition (where \mathcal{L}_2 could be replaced by any Banach space of signals), \mathcal{L}_2 stability implies that there are no nontrivial signals satisfying the equations (2) for any $\Delta \in \mathbf{B}_\Delta$. This is a condition of the type expressed in the analysis question \mathbf{Q} of the introduction; \mathcal{L}_2 stability is equivalent to a negative answer to \mathbf{Q} .

Note that in this infinite dimensional case, the \mathcal{L}_2 stability definition has the technical requirement that nontrivial approximate solutions z, w to equations (2) cannot exist with an arbitrarily small amount of equation error. This is equivalent to saying that apart from being injective, the operator $\varphi(\Delta, M)$ has a left inverse which is a bounded operator in \mathcal{L}_2 .

The following proposition from [13] reduces the representation to a simpler form.

Proposition 1 The implicit system (2) has robust \mathcal{L}_2 stability if and only if both (i) and (ii) hold.

(i) D has a bounded left inverse L .

(ii) The implicit system $\begin{bmatrix} I - \Delta \hat{A} \\ \hat{C} \end{bmatrix} z = 0$ has robust \mathcal{L}_2 stability, where $\hat{A} = A - BLC$, $\hat{C} = C - DLC$.

Condition (i) is a nominal stability condition, which if not satisfied, says there are nontrivial solutions to $Dw = 0$ and the system is not stable at $\Delta = 0$. In (ii), where from now on we replace \hat{A} and \hat{C} by A and C , the left invertibility condition on $\begin{bmatrix} I - \Delta A \\ C \end{bmatrix}$ for each $\Delta \in \mathbf{B}_\Delta$ resembles a PBH test for detectability of the pair A, C . Tests for robust \mathcal{L}_2 stability of this implicit representation are given in [13].

In many important cases, the robustness analysis can be conducted in a constant matrix representation, which is essential if computational tests are to be derived. These have the form

$$\begin{bmatrix} I - \Delta A \\ C \end{bmatrix} z = 0 \quad (5)$$

where A and C are constant matrices and $\Delta \in \Delta \subset \mathbb{C}^{n \times n}$.

One such case is that of state space implicit descriptions in discrete time (see [4, 12]). Assume that M in Figure 1 is a finite dimensional LTI system. By writing a state space realization of this system we obtain a new implicit description, where M is replaced by a constant matrix, the delta structure Δ is replaced by an augmented structure $\Delta_S = \text{diag}[\lambda I, \Delta]$, where λ is the delay operator. It is shown in [12] that under mild assumptions the analysis can be reduced to a constant matrix problem such as (5).

Another constant matrix case is when M and Δ are time invariant. Then the robust stability test reduces to (see [13])

$$\text{Ker} \begin{bmatrix} I - \Delta_0 A(j\omega) \\ C(j\omega) \end{bmatrix} = 0 \quad \forall \Delta_0 \in \mathbf{B}_{\Delta_0} \quad \forall \omega, \quad (6)$$

where $A(j\omega)$ and $C(j\omega)$ are the frequency responses of the LTI systems A and C , and Δ_0 is a constant complex perturbation with the same structure as the original Δ . This is a constant matrix test at each frequency.

These conditions are reminiscent of the structured singular value μ ([5, 10]), which corresponds to the case where $C = 0$. We give the following definition from [13]:

Definition 2 The structured singular value of the matrix A with respect to the structure $\Delta \subset \mathbb{C}^{n \times n}$, subject to the implicit constraints C is defined as

$$\mu_{\Delta, C}(A) \triangleq 0 \quad \text{if } \text{Ker} \begin{bmatrix} I - \Delta A \\ C \end{bmatrix} = 0 \quad \forall \Delta \in \Delta, \text{ otherwise}$$

$$\mu_{\Delta, C}(A) \triangleq \left(\min \left\{ \bar{\sigma}(\Delta) : \Delta \in \Delta, \text{Ker} \begin{bmatrix} I - \Delta A \\ C \end{bmatrix} \neq 0 \right\} \right)^{-1}$$

Note that the structure Δ could be specified to be real, or (7) to be mixed with real and complex blocks.

When we formulate the special case of the analysis question \mathbf{Q} for these constant matrix problems as follows

\mathbf{Q}_μ : Given equations (5) where A, C are constant matrices and $\mathbf{B}_\Delta \subset \mathbb{C}^{n \times n}$ as in (3), does there exist a $\Delta \in \mathbf{B}_\Delta$ such that the equations (5) admit nontrivial solutions?

we see that \mathbf{Q}_μ can be restated as the test "Is $\mu_{\Delta, C}(A) \geq 1$?" A negative answer to \mathbf{Q}_μ is equivalent to robust stability.

Thus the class of analysis questions which can be stated in the form \mathbf{Q}_μ can all be answered if we can compute the quantity $\mu_{\Delta, C}(A)$. As in the standard case ($C = 0$), exact computation of $\mu_{\Delta, C}(A)$ is difficult in general. In Section 4 we consider upper and lower bounds for this problem and their computation.

3 Model Validation and Identification

The basic element of any quantitative approach to scientific and technological problems is a mathematical model. Typically, the model is obtained using some combination of first principles analysis and identification from experimental data. In general, one might start with a model with some a priori structure, perhaps using some first principles knowledge, and which includes some description of the a priori uncertainty (parameters, disturbances, etc.). After performing an experiment, one is faced with the mathematical problem of finding values of the uncertainty that agree with our data.

An extensive field of research pursues the answer to this problem. For the case of dynamical models, a standard reference is [6]. A canonical example of the methods of standard system identification is the fitting of a parametric model by using prediction error methods (PEMs) described next. Assume the following model structure,

$$y = G(\lambda, \theta)u + H(\lambda, \theta)d \quad (8)$$

where λ is the shift operator, θ is a vector of parameters, G and H are discrete time systems, and d is a disturbance. Given data u and y , these methods attempt to find values of θ and d which agree with the data. Since many solutions may exist, the standard approach is to search for the solution which minimizes some norm of d .

A related problem is model validation: given a model and data, is the model consistent with the data? In the PEM example, it may be that values of θ have already been chosen, and we wish to determine whether the model is consistent with a set of data with a plausible (e.g. small enough) instance of d . The model validation problem clearly fits into the analysis question **Q** stated in the introduction: given the equations, the parameters and the data, we must check whether there exist values of the uncertainty (in this case d , satisfying some constraints, e.g. $\|d\| \leq \gamma$) which verify the equations.

The identification problem is different in that the parameters θ are also unknown. Consequently in **Q** we inquire whether there exist values of d and θ verifying the equations. To minimize $\|d\|$, we can ask **Q** for various sizes of d . Additionally, in conventional identification in general and in the PEM problem in particular we want to find the values of θ that give the affirmative answer to **Q**. Traditionally, the model is said to be identified when a fixed value of θ is chosen.

More generally, we might prefer a final identified model where, for example, some parametric uncertainty is left. This entails a somewhat expanded notion of identification. Other choices will arise as we include other sources of uncertainty. Also, we may have multiple experiments, where some of the uncertainty is fixed to have a common value across experiments, and other uncertainty (e.g., noise, parametric variations due to changes in experimental conditions, unmodeled dynamics) are allowed to vary from one experiment to another.

Therefore, a general methodology for model validation and system identification (henceforth denoted MV/ID) should provide computational tools for answering the above general question **Q** for very rich uncertainty structures, including noise, unmodeled dynamics, and parameters.

3.1 MV/ID in an LFT Setting

We will now consider a general class of MV/ID problems which are described in terms of LFTs. Recent work in [14] and [8] has shown how this type of problem is strongly related to robustness analysis machinery. In this section it is shown how these problems are very naturally formulated as implicit LFT analysis problems considered in Section 2.

Figure 2 shows a generic input-output MV/ID structure, where we assume all the elements in the diagram are constant vectors and matrices. In section 3.2 we will briefly explain how dynamical models based on finite time histories may be converted to this form. d is a vector of unknown inputs (disturbances), constrained by $\|d\| \leq 1$; Δ is in \mathbf{B}_Δ ; u and y are the measured inputs and outputs. The MV/ID problem is, again, to find values of Δ and d consistent with Figure 2. The problem is assumed to be well-posed in the sense that there are no nontrivial solutions with $d = 0$ when y and u are 0.

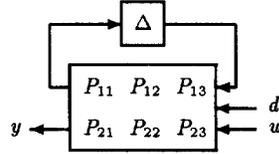


Figure 2: A standard input-output MV/ID setup.

This LFT structure captures a rich variety of linear identification problems. As a simple example, consider the standard linear regression problem

$$y = M\theta + d \quad (9)$$

where M and y are known, θ is a vector of unknown parameters and d is a vector of unknown errors. These equations are of the form of Figure 2, with $\Delta = \theta$, $u = 1$, $P = \begin{bmatrix} 0 & 0 & 1 \\ M & I & 0 \end{bmatrix}$.

In Figure 3, the equations of Figure 2 are represented in implicit form, with u and y combined into the vector v which includes all the known data:

$$v = \begin{bmatrix} u \\ y \end{bmatrix}, \quad \tilde{P}_{13} = [P_{13} \ 0], \quad \tilde{P}_{23} = [P_{23} \ -I] \quad (10)$$

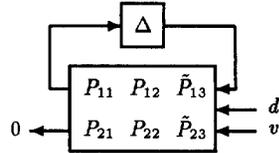


Figure 3: A standard MV/ID setup in implicit form.

Note that the input-output partition has been eliminated from the model. In fact, it could well be that we wish to validate some model based on observations of a system where this distinction is not available. Then we would arrive directly at Figure 3. For example, in the linear regression above, the “input” is an artifact of the representation.

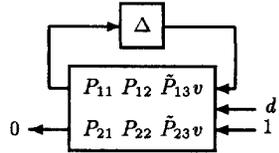


Figure 4: The MV/ID setup with data inside the matrix.

We can now incorporate the data v into the matrix by considering a fictitious scalar “input” of value 1. This results in Figure 4.

The representation has up to now two different sources of uncertainty: Δ and d . This distinction disappears and the constraint $\|d\| \leq 1$ is included in the problem when we introduce the uncertainty block $\Delta_d = d$, $\|\Delta_d\| \leq 1$, and write

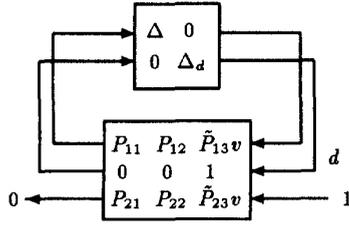


Figure 5: MV/ID as an analysis problem.

$d = \Delta_d 1$. This is shown in Figure 5. Note that Δ_d is just a new name for d which reflects its location in the diagram.

The MV/ID question is now reduced to the question “Do there exist nontrivial signals satisfying the implicit equations of Figure 5 for $\Delta \in \mathbf{B}_\Delta$?” This is an instance of a constant matrix problem of Figure 1 in Section 2. The fact that one signal is constrained to be 1 is irrelevant since everything can be normalized by linearity if the problem is well-posed.

Unless the nominal model satisfies the data, $\tilde{P}_{23}v$ is left invertible (it’s a nonzero column). Then Proposition 1 tells us the MV/ID analysis can be reduced to the question “Is $\text{Ker} \begin{bmatrix} I - \Delta A \\ C \end{bmatrix} \neq 0$ for some $\Delta \in \mathbf{B}_\Delta$?” This question is precisely \mathbf{Q}_μ posed in Section 2, and entails the same computation as in the robustness analysis problem. A positive answer to \mathbf{Q}_μ is equivalent to establishing that the model and the data are consistent.

3.2 Time Domain Data and Dynamical Models

Section 3.1 was based on static representations for models and data, with no explicit time variable. System Identification, however, deals with dynamical models and observations across time. Since the time horizon is finite, the dynamic MV/ID problem will always be represented by a finite, albeit large, number of equations which can be represented in terms of constant matrices. To illustrate, consider an autoregressive model

$$A(\lambda)y = d, \quad A(\lambda) = 1 + a_1\lambda + \dots + a_m\lambda^m \quad (11)$$

with y known and the a_i and d unknown. Over a finite horizon $[0, N]$, the equations can be written as

$$\begin{bmatrix} y(0) & \dots & \dots & \dots \\ y(1) & y(0) & \dots & \dots \\ y(2) & y(1) & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots \\ y(N) & \dots & \dots & y(N-m) \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} d(0) \\ d(1) \\ d(2) \\ \vdots \\ d(N) \end{bmatrix} \quad (12)$$

which has the form $y = M\theta + d$, a special case of (9). Thus this problem has a constant matrix LFT representation.

This example is particularly simple because it reduces to a linear regression. More generally the disturbance d may enter the equations through some dynamics which contain uncertainty. Such is the case of the ARMA model $A(\lambda)y = B(\lambda)d$. Here, the unknown parameters in $B(\lambda)$ are convolved in time with the unknown disturbance. A constant matrix representation of these equations will feature repetition of the parameters in $B(\lambda)$ which results in large δI blocks in Δ , and consequently more difficult computation. Progress has been made recently in computation with large problems involving such repetition. See [15] for more details on the general formulation of finite time horizon analysis problems and the related computation.

4 Computation for Implicit Analysis

In this section we address the issue of obtaining computational tools to answer the general question \mathbf{Q}_μ , of which both constant matrix robustness analysis and MV/ID are special cases. The idea is to extend available tools for standard μ analysis ($C = 0$ in Definition 2) which involve upper and lower bounds. Except in special cases, the upper and lower bounds must be close to know that either is close to $\mu_{\Delta, C}(A)$.

If the lower bound is ≥ 1 , then we have a solution to the equations of (5) and the system is not robustly stable. Thus a lower bound ≥ 1 is a sufficient condition for the absence of robust stability. For MV/ID, a lower bound ≥ 1 is a sufficient condition for the existence of a solution to the problem.

On the other hand, an upper bound < 1 guarantees the absence of a solution and is thus a sufficient condition for robust stability and a sufficient condition for the absence of a solution to the MV/ID problem.

4.1 Upper Bounds

The upper bounds to the structured singular value extend in a natural way to our implicit formulation. We first define the set \mathbf{X} of positive scaling matrices which commute with the elements in Δ . They have the structure

$$X = \text{diag}[X_1, \dots, X_L, x_{L+1}I_{m_1}, \dots, x_{L+F}I_{m_F}] \quad (13)$$

Now consider the two Linear Matrix Inequality feasibility conditions (LMIs)

$$\exists X \in \mathbf{X} : A^*XA - \beta^2 X - C^*C < 0 \quad (14)$$

$$\exists X \in \mathbf{X} : C_\perp(A^*XA - \beta^2 X)C_\perp^* < 0 \quad (15)$$

In (15), C_\perp^* is a matrix whose columns form a basis for the kernel of C . It is not difficult to show that these LMIs are equivalent. Testing whether an LMI is satisfied is a convex feasibility problem, for which interior point methods are available (see [1]).

Define the upper bound for μ

$$\hat{\mu}_{\Delta, C}(A) = \inf\{\beta > 0 : (14) \text{ is satisfied}\} \quad (16)$$

It is easy to show ([13]) that $\mu_{\Delta, C}(A) \leq \hat{\mu}_{\Delta, C}(A)$. Thus if (14) is satisfied for $\beta < 1$ then the answer to the question \mathbf{Q}_μ is negative.

A natural question is under which conditions the upper bound is exact ($\mu = \hat{\mu}$). From the point of view of Δ , we inquire which structures are μ -simple (i.e. which structures ensure $\mu = \hat{\mu}$ for any matrices A, C). The answer from [13] is the following:

Proposition 2 *The following structures are μ -simple in the implicit case.*

(i) $\Delta = \{\delta I : \delta \in \mathbb{C}\}$

(ii) $\Delta = C^{n \times n}$

(iii) $\Delta = \{\text{diag}[\Delta_1, \Delta_2] : \Delta_i \in \mathbb{C}^{m_i \times m_i}\}$

(iv) $\Delta = \{\text{diag}[\Delta_1, \Delta_2] : \Delta_i \in \mathbb{R}^{m_i \times m_i}\}$, for A, C real

Remarks:

- In standard μ ([10]), two additional complex structures are μ -simple ($\text{diag}[\delta_1 I, \Delta_2]$ and $\text{diag}[\Delta_1, \Delta_2, \Delta_3]$) but this does not carry through ([13]) to the implicit case.
- The upper bound is an *exact* robust stability test for the dynamic problem where the uncertainty is in a class of time varying operators; see [13].
- Even if the structure is not μ -simple, special cases on the matrices A, C can yield $\mu = \hat{\mu}$. An example of this is when A is rank 1 in the kernel of C .

In the case where Δ contains blocks of repeated real perturbations, a tighter upper bound can be defined as in the standard case ([17]) by adding a term $j(A^*G - GA)$ to the LMI (14). In this case $G = G^* = \text{diag}[G_1 \dots G_r, 0 \dots 0]$, where the G_i blocks correspond to the real δI blocks.

Although the upper bound can in principle be computed exactly by solving an LMI, MV/ID problems can have very large matrices with extensive repetition in uncertainty blocks. In these cases the LMI involves very large full blocks, which standard LMI solvers cannot handle. By taking advantage of special structure in these repeated problems ([15]), some reduction in complexity should be possible, but further research is required.

4.2 Lower Bounds

In contrast to an upper bound, which guarantees that there are no solutions to the equations in \mathbf{Q}_μ , a lower bound guarantees that there is a solution to the equations. This suggests an equivalent definition of $\mu_{\Delta, C}(A)$:

$$\mu_{\Delta, C}(A) \triangleq \max \{ \beta : (\beta, \Delta) \in S \cup \{(0, 0)\} \} \quad (17)$$

$$S \triangleq \left\{ (\beta, \Delta) : \|\Delta\| = 1, \text{Ker} \begin{bmatrix} \beta I - \Delta A \\ C \end{bmatrix} \neq 0 \right\} \quad (18)$$

If a point is found in S with $\beta \geq 1$, then the answer to the question \mathbf{Q}_μ is yes.

For a general problem, the set S may be disjoint, it may have isolated points, or it may be empty. We do not consider the general problem here; our problems are easier. The MV/ID problems are continuous in the following sense: if we have a point in S then a small change in uncertainty or unknown parameters, along with the appropriate small change in measurement noise or equation error, provides us with another point in S . We can find a point in S for a MV/ID problem by choosing any values for the uncertainty and finding the output noise that is consistent. This allows a perturbation analysis of optimality.

Proposition 3 *Let P parametrize $\text{Ker}(C)$ and let $0 = (\beta I - \Delta A)Px$ so that $(\beta, \Delta) \in S$. If $\exists \hat{\beta} > 0$, $\hat{\Delta} \leq 0$, and \hat{x} such that*

$$0 = (\hat{\beta} - \hat{\Delta}A)Px + (\beta - \Delta A)P\hat{x} \quad (19)$$

then (β, Δ) is not a local maximum of (17).

Note that this test is a linear programming feasibility test on $\hat{\beta}$, $\hat{\Delta}$, and \hat{x} .

Remark: More generally we may encounter problems that are not continuous. We then face the question of the physical or engineering interpretation of such a problem. Often the problem is poorly posed. If one chooses to proceed with a discontinuous problem, one can go through a regularization procedure completely analogous to the one in [11], where a discussion of the relevance of such problems and the interpretation of regularization may be found.

The approach to the computation of the maximization problem (17) is heavily influenced by the fact that the problem is NP hard ([2] shows that a special case of (17) is NP complete).

Roughly speaking, the fact that (17) is NP hard means that it cannot be computed exactly in the worst case without entirely unacceptable growth in computation cost with problem size. To obtain acceptable computation for the large problems we are interested in, one is forced either to consider special cases, e.g. μ -simple problems, or relax the requirement for exact computation. It is important to note that

being NP hard is a property of a problem, not a property of an algorithm.

Extensive experience with lower bound computation for the special case $C = 0$ shows that specialized algorithms are much better than standard optimization code. The successful algorithms can be separated roughly into the two classes discussed below, which look for a local maximum of β in S .

The first class of algorithms is based on conditions on β , Δ , and x consistent with a local maximum. These are called alignment conditions. The algorithms try to satisfy these conditions directly. If these local maximum conditions are not achieved, then no lower bound is provided.

The power algorithms described in [9] and [18] are efficient and usually calculate the structured singular value reasonably accurately. The principal difficulty in generalizing the algorithms of this type is that the conditions for a local maximum derived from Proposition 3 are considerably more expensive to compute than the alignment conditions of the $C = 0$ case. Since it is difficult to analyze a power algorithm mathematically, the performance of such algorithms is measured by testing the algorithm on a large set of representative problems.

The second class of algorithms moves from one point in S to another, so a lower bound is constantly being provided and improved upon. The first step is to find a point in S , which is easy in our case. Next, we check conditions (19) to decide if we are done. If we are not done, then a better point can be found nearby. This can be proved using a contraction mapping and starting in the direction of $(\hat{\beta}, \hat{\Delta})$. Although standard optimization code is in this second class of algorithms, experience suggests that specialized code incorporating elements of algorithms of the first class will be much better.

5 The Least Squares Problem Revisited

A canonical example of a computationally easy problem in system identification is the least squares problem: in the linear regression setup $y = M\theta + d$ of (9) we wish to find the values of θ which satisfy the equations and minimize the 2 norm of the vector d . Assuming M is full column rank, this problem has an explicit solution $\hat{\theta} = (M^*M)^{-1}M^*y$, which gives a minimum of $\|d\|^2$ equal to

$$\gamma_0^2 = y^*(I - M(M^*M)^{-1}M^*)y. \quad (20)$$

An important requirement to validate our approach to MV/ID is to ensure that this simple problem does not turn into a hard one when recast in our analysis framework. Of course, no solution will be more efficient than the least squares solution, and we would not expect a method which is developed to encompass a very general problem to be optimal in a special case. We do, however, want the problem to be tractable in the new formulation. Fortunately, as is shown below, the resulting implicit analysis problem is indeed tractable and we can recover the least squares solution.

We first specialize the MV/ID setup of Figure 5 to the linear regression case. This is done by following the steps in Section 3.1, and yields the diagram on Figure 6.

The only modification to the setup in Figure 5 is that scalings k , γ are added to fix the allowable sizes of θ , d respectively. In the least squares problem, we attempt to find the smallest value of γ such that the corresponding \mathbf{Q}_μ gives an affirmative answer for some value of k (we do not have constraints on parameter size in this case). So the strategy will be to let $k \rightarrow \infty$ in the analysis, and attempt to minimize γ .

For simplicity assume $\|y\| = 1$. Therefore the “ D ” matrix in the implicit representation (refer to Figure 1) is left invertible, and Proposition 1 converts it to the standard form (5),

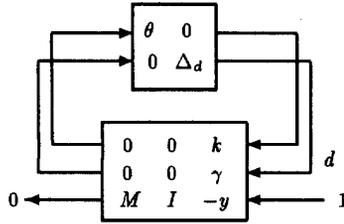


Figure 6: The least squares problem.

with

$$A = \begin{bmatrix} ky^*M & ky^* \\ \gamma y^*M & \gamma y^* \end{bmatrix}; C = (I - yy^*) \begin{bmatrix} M & I \end{bmatrix} \quad (21)$$

The structure Δ consists in this case in two real full blocks. Although they are nonsquare (they are columns) the analysis in Proposition 2 (iv) remains valid; therefore the structure is μ -simple and LMIs (14), (15) with $\beta = 1$ capture exactly the answer to \mathbf{Q}_μ . This fact is already indication that our problem remains tractable in the new formulation, although a solution based on LMIs will be less efficient than the least squares solution.

To make the point clearer, we can show explicitly that the LMI approach gives the least squares solution as $k \rightarrow \infty$. Consider the LMI (15) with $\beta = 1$; the X scaling in this case consists of two scalar parameters, one for each of the full blocks. Let us call the first one $x > 0$, and we can normalize the second one to 1. Some algebra gives $C_1^* = \begin{bmatrix} I & 0 \\ -M & y \end{bmatrix}$,

$$C_1(A^*XA - X)C_1^* = \begin{bmatrix} -(xI + M^*M) & -M^*y \\ -y^*M & k^2x + \gamma^2 - 1 \end{bmatrix} \quad (22)$$

A Schur complement operation reduces (15) to

$$k^2x + \gamma^2 - 1 + y^*M(xI + M^*M)^{-1}M^*y < 0 \quad (23)$$

As $k \rightarrow \infty$, the unknown x will have to go to zero if (23) is to be satisfied; this implies that $(xI + M^*M)^{-1} \rightarrow (M^*M)^{-1}$ and the LMI will be feasible for large k if and only if

$$\gamma^2 < 1 - y^*M(M^*M)^{-1}M^*y = \gamma_0^2 \quad (24)$$

Recapitulating,

- For $\gamma < \gamma_0$ the LMI is feasible and therefore $\hat{\mu}_{\Delta,C}(A) = \mu_{\Delta,C}(A) < 1$, which in turn implies that the answer to \mathbf{Q}_μ is negative (no solutions with $\|d\| \leq \gamma$).
- For $\gamma > \gamma_0$ and large enough k the LMI is not feasible, $\hat{\mu}_{\Delta,C}(A) = \mu_{\Delta,C}(A) \geq 1$ and therefore there exist solutions to the MV/ID problem.

So we again find that γ_0 is the minimum norm for d , as expected. We have not shown how to solve for θ and d , but this information can also be obtained from the LMI approach.

We reiterate that we are not advocating this method for a least squares problem; this is a method suitable for a large class of problems. We have simply shown that it remains tractable in this simple case.

6 Conclusion

In this paper, we have considered a general analysis question involving equations and uncertainty; when formulated in the form \mathbf{Q}_μ , it contains as special cases a large class of robustness analysis problems which admit a constant matrix formulation, and a general model validation/system identification problem.

The advantage of this formulation as a unifying scheme is its great generality. This generality implies in turn that from a computational point of view, we cannot expect guaranteed performance of algorithms: there will always be problems in such a large class that are intractable, as is already known from computational complexity results ([2]) for the standard versions of μ analysis. The hope is that moderately efficient algorithms which can handle "realistic" instances of the problem can be developed. It is unclear at this stage whether such objective can be met, and substantial research on computation is required.

The results of this research on computation should settle the question of whether the unification presented here leads to the more ambitious goal of a unified approach to model based controller design tied to experimental data. In this perspective, models are merely a tool to synthesize efficient automatic control systems, and the process of tailoring a design to a particular application must be approached as a whole, with various stages and possibly iterations of modeling, design, analysis and simulation. If these activities can be performed with the same underlying mathematics, it is expected that a more fruitful interplay between these disciplines will emerge.

Acknowledgements

The framework in this paper was influenced by many stimulating discussions with our colleagues. We would particularly like to acknowledge the input of Roy Smith, Munzer Dahleh, and Keith Glover. This work was supported by NASA, AFOSR, and the NSF PYI program.

References

- [1] Boyd S., El Ghaoui L., Feron E., Balakrishnan V., *Linear Matrix Inequalities in System and Control Theory*, SIAM, Philadelphia, Pennsylvania, 1994.
- [2] Braatz, R., Young, P., Doyle, J., Morari, M., 1993, *Computational complexity of μ calculation*, IEEE TAC May, 1993.
- [3] D'Andrea R., Paganini F., Doyle J., *Uncertain Behavior*, Proc. 1993 IEEE CDC.
- [4] D'Andrea R., Paganini F., *Interconnection of Uncertain Behavioral Systems for Robust Control*, Proc. 1993 IEEE CDC.
- [5] Doyle, J., *Analysis of feedback systems with structured uncertainty*, IEE Proceedings, 129, 242-250
- [6] Ljung, L., 1987, *System Identification Theory for the User*, Prentice-Hall, Englewood Cliffs, New Jersey.
- [7] Megretski A., Treil S., *Power Distribution inequalities in optimization and robustness of uncertain systems*, Journal of Mathematical Systems, Estimation, and Control, 1993.
- [8] Newlin, M., Smith, R., *A Generalization of the Structured Singular Value and its Application to Robust Minimum Gain Problems and Model Validation*, to appear IEEE TAC.
- [9] Packard A., Fan M., Doyle J., *A Power Method for the Structured Singular Value* Proc. 1988 IEEE CDC.
- [10] Packard A., Doyle J., *The Complex Structured Singular Value*, Automatica, Vol. 29, No. 1, pp. 71-109, 1993.
- [11] Packard A., Pandey P., *Continuity Properties of the Real/Complex Structured Singular Value*, IEEE TAC, Vol. 38, pp. 415-428, 1993.
- [12] Paganini F., D'Andrea R., Doyle J., *Behavioral Approach to Robustness Analysis*, Proc. 1994 ACC.
- [13] Paganini F., Doyle J., *Analysis of Implicitly Defined Systems*, Proc. 1994 IEEE CDC.
- [14] Smith, R., Doyle, J., *Model validation: a connection between robust control and identification*, IEEE TAC, July, 1992.
- [15] Tierno J., Doyle J., *Finite Time Horizon Robust Performance Analysis*, Proc. 1994 IEEE CDC.
- [16] Willems J.C., *Paradigms and Puzzles in the Theory of Dynamical Systems*, IEEE TAC, Vol. 36, pp. 259-294, 1991.
- [17] Young P., *Robustness with Parametric and Dynamic Uncertainty*, PhD thesis, California Institute of Technology, 1993.
- [18] Young P., Doyle J., *Computation of μ with Real and Complex Uncertainties*, Proc. 1990 IEEE CDC.