Convex formulations for data-based uncertainty minimization of linear uncertainty models
Häggblom, Kurt Erik

Published in:
The 11th International Conference on Control, Automation, Robotics and Vision

DOI:
10.1109/ICARCV.2010.5707895

Published: 01/01/2010

Please cite the original version:

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Convex Formulations for Data-Based Uncertainty Minimization of Linear Uncertainty Models

Kurt E. Häggblom
Process Control Laboratory, Department of Chemical Engineering
Åbo Akademi University
Åbo (Turku), Finland
khaggblo@abo.fi

Abstract—Convex formulations are derived for the minimization of uncertainty bounds with respect to a nominal model and given input-output data for general uncertainty models of LFT type. The known data give rise to data-matching conditions that have to be satisfied. It is shown how these conditions, which originally are in the form of BMIs for a number of uncertainty models, can be transformed to LMIs, thus making the optimization problem convex. These formulations make it easy to find the best uncertainty model from a number of alternatives for robust control design.

Keywords—Uncertainty modeling, LFT uncertainty, linear multivariable systems, convex optimization, linear matrix inequalities, robust control, distillation columns.

I. INTRODUCTION

A robust control design requires a model with information about the model uncertainty, i.e., information about the discrepancy between a “nominal” model and the true system. Uncertainty models can be expressed in various ways. One way is to express the uncertainty as a linear fractional transformation (LFT) with a norm-bounded uncertainty matrix. Important special cases of this type of uncertainty description are models with additive uncertainty and multiplicative input or output uncertainty. Inverse types of these uncertainties as well as various combinations of the basic types also exist.

Up till now there has been no clear consensus as to what constitutes a control-oriented uncertainty set [1]. According to [2], the various uncertainty types are all equivalent provided that the identification methods deliver optimal uncertainty sets rather than an uncertainty bound around a prefixed nominal model. Thus, the uncertainty set should be minimized over both the nominal model and the uncertainty bound, as done, e.g., in [3]. For an LFT type of uncertainty, it has recently been shown that the $\mathcal{H}_\infty$ norm of the uncertainty is a rigorous measure of the worst-case degradation of the stability margin for a system under feedback control irrespective of the particular type of uncertainty structure [4][5]. An important task is then to choose a structure, which allows a comparatively small uncertainty bound.

The uncertainty bound for a given uncertainty model can be minimized by choosing the nominal model optimally. If experimental input-output data are known, our opinion is that the nominal model should usually be determined by matching the uncertainty description to input-output data and not to a set of models fitted to the same data. It has been proved that model matching cannot provide a smaller uncertainty bound than input-output matching [6].

For choosing the best type of uncertainty description, it appears sufficient to consider frequency response samples instead of working with analytical transfer functions. This simplifies the calculations immensely since they can be done frequency by frequency using matrix algebra. Because this involves optimization, it is desirable to formulate it as a convex optimization problem. Some types of uncertainty structures, e.g., additive uncertainty, naturally yield optimization formulations that are convex. Many other types, e.g., output multiplicative uncertainty, are more difficult to formulate as convex optimizations. A major contribution of this paper is to show how this can be done. An example is given, where an iterative solution due to a non-convex formulation converges to non-global minima whereas the suggested convex formulation does not have this problem.

II. PROBLEM DESCRIPTION

A. Choosing an uncertainty model for control design

We consider linear multiple-input multiple-output (MIMO) uncertainty models of the form

$$G = G_0 + H_{21} \Delta (I - H_{11} \Delta)^{-1} H_{12},$$

where $G$ is a transfer function of the true system, $G_0$ is the transfer function of a nominal model, and $\Delta$ is the transfer function of an unknown bounded perturbation causing uncertainty about the true system. Depending on the particular type of uncertainty (additive, input or output multiplicative, inverse types of uncertainty, combinations of various types of uncertainty), $H_{11}$, $H_{12}$ and $H_{21}$ may be constant matrices or transfer functions containing combinations of weighting matrices and the nominal model. All transfer functions, including $(I - H_{11} \Delta)^{-1}$, are assumed to be proper and real-rational. In particular, this means that $\|H_{11} \Delta\|_\infty < 1$ is required.
It has been shown that
\[ \| \Delta \|_{\infty} := \sup_{\omega} \sigma(\Delta(j\omega)), \]  
where \( \sigma(A) \) denotes the largest singular value of \( A \), is a rigorous measure of the worst-case degradation of the stability margin due to a perturbation of a system under feedback control [5]. Clearly, an uncertainty description having a smaller \( \| \Delta \|_{\infty} \) than another uncertainty description is preferable from the viewpoint of feedback control. Hence, this can be used as a basis for choosing an uncertainty description for robust control design.

B. Model matching

Assume that we have information about the true system in the form of a number of possible transfer functions \( G_k \), \( k = 1, \ldots, N \). The nominal model and the perturbation \( \Delta_k \) associated with \( G_k \) are unknown, but they have to satisfy
\[ G_k = G_0 + H_{21}\Delta_k(I - H_{11}\Delta_k)^{-1}H_{12}, \quad k = 1, \ldots, N. \]  
(3)

It is clear that the size of the perturbation \( \Delta_k \) required to satisfy (3) depends not only on the type of uncertainty description (i.e., choice of \( H_{11}, H_{12} \) and \( H_{21} \)), but also on the choice of \( G_0 \). This suggests that \( G_0 \) should be determined by solving the optimization problem
\[ G_0 = \arg \min_{G_0} \max_k \| \Delta_k \|_{\infty} \]  
subject to the model matching conditions (3). Obviously, the type of uncertainty model giving the smallest minimum is the best one according to this measure.

C. Input-output matching

If information about the system is obtained through identification, input-output data are available. An attractive way of removing noise from the output is to fit a model \( G_k \) to the data and to calculate a noise-free output \( y_k \) by \( y_k = G_ku_k \), where \( u_k \) is the input of experiment \( k \), \( k = 1, \ldots, N \). Since the purpose of the set of experiments is to excite the system in various ways, the inputs do not tend to be persistently exciting in all individual experiments. This may be the case, for example, in the identification of a nonlinear system, where it may be useful to split the data into subsets with more linear behavior, each subset considered to be an experiment [7]. Thus, \( G_k \) only applies to the particular input \( u_k \), and the relevant information is input-output data \( \{u_k, y_k\} \), \( k = 1, \ldots, N \). This means that the model-matching condition (3) should be replaced by the input-output matching condition
\[ y_k = G_0u_k + H_{21}\Delta_k(I - H_{11}\Delta_k)^{-1}H_{12}u_k. \]  
(5)

It can be shown that input-output matching results in a less conservative uncertainty model than model matching [6].

D. Optimization frequency by frequency

For control design, we need the transfer function of the nominal model, but for choosing the best type of uncertainty description, we can simplify the work by considering frequency response samples and do the calculations frequency by frequency. Of course, the result of these calculations can subsequently be used for constructing the model.

In accordance with this, the input-output data are represented by their sampled frequency responses. Because of the availability of \( G_k \), these are easy to calculate for standard inputs. Thus, the available information is \( \{u_k(j\omega), y_k(j\omega), \omega \in \Omega\} \), \( k = 1, \ldots, N \), where \( \Omega \) is a chosen set of frequency points. The optimization problem now becomes
\[ G_0(j\omega) = \arg \min_{G_0(j\omega)} \max_k \sigma(\Delta_k(j\omega)) \]  
\[ \forall \omega \in \Omega. \]  
(6)

The uncertainty \( \Delta_k \) is assumed to be unstructured.

E. Convexity of data-matching constraints

It is desired to formulate the optimization problem as a convex optimization problem, where the data-matching constraints are expressed as linear matrix inequalities (LMIs). For some types of uncertainty descriptions, this is very easy to do, for other types, it requires certain transformations. There are also cases, where the problem remains non-convex.

Consider additive uncertainty, for example, which in its basic form is described by the input-output relationship
\[ y_k = G_0u_k + \Delta_ku_k, \quad k = 1, \ldots, N. \]  
(7)

The optimization can now be formulated as [3][10]
\[ \min_{G_0} \alpha \quad \text{s.t.} \quad \begin{bmatrix} \alpha I & y_k - G_0u_k \\ (y_k - G_0u_k)^* & u_k^*u_k \end{bmatrix} \succeq 0 \]  
\[ k = 1, \ldots, N, \quad \forall \omega \in \Omega \]  
(8)

which is a convex optimization problem. Here, \( A^* \) denotes the complex conjugate transpose of \( A \) and \( P \succeq 0 \) denotes that \( P \) is positive semidefinite.

Consider now output-multiplicative uncertainty, which in the absence of weights is described by the relationship
\[ y_k = G_0u_k + \Delta_kG_0u_k, \quad k = 1, \ldots, N. \]  
(9)

This results in the optimization problem
\[ \min_{G_0} \alpha \quad \text{s.t.} \quad \begin{bmatrix} \alpha I & y_k - G_0u_k \\ (y_k - G_0u_k)^* & u_k^*G_0^*G_0u_k \end{bmatrix} \succeq 0 \]  
\[ k = 1, \ldots, N, \quad \forall \omega \in \Omega \]  
(10)

which is non-convex due to the appearance of \( G_0^*G_0 \). There is evidence that an iterative solution by keeping \( G_0^*G_0 \) fixed during each iteration does not always converge to a global minimum, as shown later in this paper.

The main contribution of this paper is to show how this optimization problem, as well as other similar ones for other types of uncertainty, can be reformulated as a convex optimization problem. We also consider the more difficult case when \( H_{11} \neq 0 \) in (5).
III. MAIN RESULTS

A. Two theorems

We shall make use of the following theorems adapted from [9].

Theorem 1. Consider the matrix equation
\[ G = H_{22} + H_{21} W \Delta (I - H_1 W \Delta)^{-1} H_{12}, \]
(11)
where \( \sigma(H_1 W) < 1 \) and \( \sigma(\Delta() \leq 1 \). For every \( \Delta \), there is a \( \hat{\Delta} \) satisfying
\[ G = \hat{H}_{22} + \hat{H}_{21} \hat{\Delta} \hat{H}_{12}, \]
(12)
where
\[
\begin{align*}
\hat{H}_{22} &= H_{22} + H_{21} W H_{11}^*(I - H_1 W H_{11}^*)^{-1} H_{12} \\
\hat{H}_{21} &= -H_{21} W (I - W^* H_{11}^* H_{11} W)^{-1/2} \\
\hat{H}_{12} &= (I - H_1 W W^* H_{11}^*)^{-1/2} H_{12}
\end{align*}
\]
(13)
if and only if \( \sigma(\hat{\Delta}) \leq 1 \). For every \( \hat{\Delta} \), \( \sigma(\Delta) \leq 1 \), there is also a \( \hat{\Delta} \) satisfying (11) if and only if \( \sigma(\hat{\Delta}) \leq 1 \).

Theorem 2. Consider the matrix equation
\[ C = A \Delta B. \]
(14)
There is a solution \( \Delta \), \( \sigma(\Delta) \leq 1 \), if and only if
\[
\left[ \begin{array}{cc}
AA^* & C \\
C & B^* B
\end{array} \right] \succeq 0.
\]
(15)

B. Uncertainty modeling

We shall derive input-output relationships of a general linear uncertainty model covering a large variety of uncertainty structures. We start the modeling with the equation
\[ y = G_0(u + w_2) - w_1, \]
(16)
where \( G_0 \) is a nominal model, \( w_2 \) is a disturbance added to the input \( u \), and \( w_1 \) is a disturbance subtracted from the nominal output to produce the true output \( y \). Various combinations of disturbed and undisturbed inputs and outputs are defined by
\[
\left[ \begin{array}{c}
w_1 \\
w_2
\end{array} \right] = S_w W \Delta \left[ \begin{array}{c} z_1 \\
z_2
\end{array} \right],
\]
(17)
where \( S \) is a diagonal matrix with entries \( s_{ii} \in \{0,1\} \); \( s_{ii} = 1 \) selects the nominal and \( s_{ii} = 0 \) the true input or output (the choice of signal signs in (16) simplifies this definition). The disturbances are generated by a perturbation \( W \Delta \), \( \sigma(\Delta) \leq 1 \), according to
\[
\left[ \begin{array}{c}
w_1 \\
w_2
\end{array} \right] = S_w W \Delta \left[ \begin{array}{c} z_1 \\
z_2
\end{array} \right],
\]
(18)
where \( S_w \) selects the disturbance(s) and \( S_z \) the disturbance generating signal(s). Usually,
\[
S_w \in \left\{ \left[ \begin{array}{cc} 0 & 0 \\
I & 0
\end{array} \right], \left[ \begin{array}{cc} I & 0 \\
0 & I
\end{array} \right] \right\}, \quad S_z \in \left\{ \left[ \begin{array}{cc} I & 0 \\
I & 0
\end{array} \right], \left[ \begin{array}{cc} I & 0 \\
0 & I
\end{array} \right] \right\},
\]
(19)
but the entries can also contain weights or the inverted factor of a coprime factorized model [5]. Note that any combination of the matrices \( S_w \) and \( S_z \) may be used. Thus, \( \Delta \) is not restricted to a square matrix.

Solving \( w_1 \) from (17) and (18) followed by substitution into (16) gives
\[ y = G_0 u + [-I \quad G_0] S_w W \Delta (I - S_w W \Delta)^{-1} S_z \left[ \begin{array}{c} y \\
u
\end{array} \right], \]
(20)
where
\[ S_\Delta := S_z SS_w. \]
(21)
For \( S = 0 \), and thus \( S_\Delta = 0 \), this is equivalent to the “left 4-block interconnection” considered in [5]. The choice \( S = I \) yields the “right 4-block interconnection” and other choices give mixed variants.

It should be noted that well-posedness of (20) requires \( \sigma(S_\Delta W) < 1 \), which might be a problem if \( S \neq 0 \). For the standard choices in (19), this requires \( \sigma(W) < 1 \) if \( S_\Delta \neq 0 \). However, the formulation in (18) does not have such a restriction, which is a result of the fact the output \( y \) appears also in the right-hand side of (20). Solving (20) for \( y \) gives
\[ y = G_0 u + [-I \quad G_0] S_w W \Delta (I - S_\Delta W \Delta)^{-1} S_z \left[ \begin{array}{c} G_0 \\
I
\end{array} \right] u, \]
(22)
where
\[ \tilde{S}_\Delta = S_z \left[ S + [-I \quad 0] \right] S_w. \]
(23)

As can be seen, there is now a well-posedness restriction when \( S_{z_1} \neq 0 \) (the columns of \( S_z \) corresponding to \( z_1 \)) even if \( S = 0 \). On the other hand, if \( S_{z_1} = I \) (meaning that \( z_1 \) is the output from the nominal model) and \( S_{w_2} = 0 \) (the rows of \( S_w \) corresponding to \( w_2 \)), \( \tilde{S}_\Delta = 0 \) and the system is always well-posed.

C. The data-matching condition

By applying Theorem 1 and 2, a data-matching condition of the form (15) can be derived for (22). However, if \( S_{z_1} \neq 0 \), \( G_0 G_0^* \) will appear in the \( B^* B \) part of (15), and the minimization problem will be non-convex as shown in Section II.E. In cases where \( \sigma(S_\Delta W) < 1 \), this can be avoided by the formulation in (20). There is still a non-convexity problem if \( S_w = I \), but not if \( S_w = [I \quad 0]^T \), where the superscript \(^T\) denotes transpose. If \( S_w = [0 \quad I]^T \), and \( G_0 \) is assumed to be invertible, the problem can be avoided by multiplying (20) by \( G_0^{-1} \) from the left. In this case, \( G_0^{-1}(j\omega) \) is determined instead of \( G_0(j\omega) \).

In the sequel, minimizing the uncertainty based on (20) will be considered. Application of Theorem 1 with \( H_{22} = G_0 u \),
$$H_{21} = [-I \ G_0] S_w, \ H_{11} = S_A, \ H_{12} = S_z [y^* \ u^*]^T$$, and $G$ replaced by $y$, gives

$$y = G_0 u + H_{21} W W^* S_A (I - S_A W W^* S_A)^{-1} H_{12}$$
$$- H_{21} W (I - W^* S_A^* S_A W)^{-1/2} \hat{\Delta} (I - S_A W W^* S_A)^{-1/2} H_{12}. \quad (24)$$

According to Theorem 2, (24) is satisfied by $\hat{\Delta}, \mathcal{S}(\hat{\Delta}) \leq 1$, if and only if (15) holds. After some manipulations, we obtain

$$A A^* = H_{21} P H_{21}^*$$
$$C = y - G_0 u - H_{21} P S_A^* S_z \begin{bmatrix} y \\ u \end{bmatrix} \quad (25)$$
$$B^* B = \begin{bmatrix} y^* \\ u^* \end{bmatrix} S_z^* (I + S_A P S_A^*) S_z \begin{bmatrix} y \\ u \end{bmatrix},$$

where

$$P = (I - W W^* S_A^* S_A)^{-1} W W^*. \quad (26)$$

Since (15) must hold for every experiment $k, k = 1,..., N$, we obtain the data-matching condition

$$\begin{bmatrix} H_{21} P H_{21}^* & y_k - G_0 u_k - H_{21} P S_A^* S_z \begin{bmatrix} y_k \\ u_k \end{bmatrix} \\ \begin{bmatrix} y_k \\ u_k \end{bmatrix}^* S_z^* (I + S_A P S_A^*) S_z \begin{bmatrix} y_k \\ u_k \end{bmatrix} \end{bmatrix} \succeq 0, \quad (27)$$

where (*) indicates the complex-conjugate transpose of the elements in the symmetrical position of the full matrix (which must be Hermitian).

### D. Minimizing the uncertainty

Because the weight $W$ has been introduced in $W \hat{\Delta}$ to enable $\mathcal{S}(\hat{\Delta}) \leq 1$, (6) can be replaced by

$$G_0 = \arg \min_{G_0} \mathcal{S}(W) \quad \text{s.t.} \quad (27), \quad \forall k, \forall \omega \in \Omega. \quad (28)$$

The maximum over all $k$ in (6) is replaced by the requirement that (27), with the definition (26), must hold for $\forall k$. Because $\mathcal{S}(W) = \|W\|_2$, the minimization can also be formulated as

$$G_0 = \arg \min_{G_0} \alpha \quad \text{s.t.} \quad W W^* \preceq \alpha I$$

and (27), $\forall k, \forall \omega \in \Omega. \quad (29)$

Unfortunately, this problem is non-convex, because $WW^*$ does not appear linearly in $P$.

Let us instead consider the optimization problem

$$G_0 = \arg \min_{G_0} \beta \quad \text{s.t.} \quad P \preceq \beta I \quad \text{and} \quad (27), \quad \forall k, \forall \omega \in \Omega, \quad (30)$$

which is convex if $H_{21}$ is independent of $G_0$ because $P$ and $G_0$ then appear linearly in (27). (If $S_w = I$, an iterative solution with $H_{21}$ fixed during each iteration may be tried.) From (26) we obtain that $P \preceq \beta I$ is equivalent to

$$WW^* \preceq (\beta I + S_A S_A^*)^{-1}. \quad (31)$$

From this it follows that the minimization of $\beta$ according to (30) also minimizes $\mathcal{S}(W)$. Thus, (30) minimizes the uncertainty by means of $G_0$ subject to the data-matching constraint (27).

### IV. Example

In this example, data obtained by identification of a pilot-scale two-product distillation column are used. The identification experiments are described and the set of models used for data smoothing are given in [7]. The data have also been used in various other studies related with uncertainty modeling [3][6][8][10]. The previous studies have indicated that the input-output data can be reasonable well captured by an uncertainty model of output-multiplicative type when the nominal model determined in [7] is used. Here we use an improved version of the nominal model, which behaves better than the original, especially at high frequencies. However, it is not optimized to minimize the uncertainty.

The input-output relationship of such a model is obtained by selecting $S_z = [I \ 0], \ S_w = [I \ 0]^T$, and $S = I$ (actually $S_{11} = I$ is sufficient) in (22), resulting in

$$y_k = G_0 u_k + W \Delta_k G_0 u_k, \quad k = 1,..., N \quad (32)$$

when all $N$ experiments are considered. In this case, $N = 6$. The optimization problem can be formulated as

$$\min \alpha \quad \text{s.t.} \quad \begin{bmatrix} WW^* & y_k - G_0 u_k \\ (y_k - G_0 u_k)^* & u_k^* G_0^* G_0 u_k \end{bmatrix} \succeq 0, \quad (33)$$

$$k = 1,..., N, \quad \forall \omega \in \Omega.$$  

Since the size of the uncertainty is represented by $\mathcal{S}(W)$, the smallest uncertainty required to match all input-output data is given by

$$\min \mathcal{S}(W) = \sqrt{\min \alpha}. \quad (34)$$

Fig. 1 shows $\min \mathcal{S}(W)$ plotted against the frequency $\omega$ when the nominal model is not updated (blue dashed line). In this case, $\min \mathcal{S}(W)$ is more directly obtained from

$$\min \mathcal{S}(W) = \max_k \frac{\|y_k - G_0 u_k\|_2}{\|G_0 u_k\|_2}, \quad \forall \omega \in \Omega. \quad (35)$$

As can be seen, the uncertainty is quite large, especially at higher frequencies, where it is higher than 100%. It is interesting to see, how much this uncertainty can be reduced by optimizing with respect to $G_0$. This optimization
The problem is non-convex in the formulation (33) due to $G_0^*G_0$. If an iterative solution is attempted with $G_0^*G_0$ constant at its previous value during each iteration, the result depicted in Fig. 1 is obtained (red dotted line). At most frequencies the result is clearly better than that obtained with the original nominal model, but surprisingly, it is much worse at a few frequencies. This means that the optimization has converged to a non-global local minimum. Ten iterations were used, which was more than enough to reach a minimum. The optimization was done with the standard solver for semidefinite programming in YALMIP [11] together with Matlab.

To formulate the optimization as a convex optimization problem, we need to use the mixed input-output relationship (20), which for output multiplicative uncertainty gives the data-matching condition

$$\begin{bmatrix} P & (I + P)y_k - G_0u_k \\ (*) & y_k(I + P)y_k \end{bmatrix} = 0,$$

where

$$P = (I - WW^*)^{-1} WW^*.$$  

Since this formulation requires $\sigma(W) < 1$, there is no guarantee that it will work at all considered frequencies. However, the YALMIP software [11] provides a solution with $\min \sigma(W) < 1$ at all frequencies, as shown in Fig. 1 (green full line). Considering that the control-relevant measure defined in (2) for an unbounded uncertainty corresponds to

$$\|\mathcal{W}\|_\infty := \sup_{\omega} \sigma(W(j\omega)),$$

the improvement obtained with the convex formulation is significant.

V. CONCLUSION

Convex formulations for the minimization of uncertainty bounds with respect to a nominal model and given input-output data for general uncertainty models of LFT type have been derived. In particular, it has been shown how data-matching conditions obtained as bilinear matrix inequalities in certain cases can be transformed to linear matrix inequalities, thus making the optimization convex. The optimization procedure can effectively be used for choosing the best uncertainty model for robust control design compatible with known data.

REFERENCES