Synthesizing mixed H2/H-infinity dynamic controller using evolutionary algorithms

Pedersen, Gerulf; Langballe, A.S.; Wisniewski, Rafal

SYNTHESIZING MIXED H_2/H_∞ DYNAMIC CONTROLLER USING EVOLUTIONARY ALGORITHMS

Gerulf K. M. Pedersen * **Anders S. Langballe** ** **Rafał Wiśniewski** *

* *Aalborg University, Department of Control Engineering, Fredrik Bajers Vej 7, DK-9220 Aalborg Ø, Denmark*
** *TDC Tele Danmark, Digtervejen 1, DK-9200 Aalborg SV, Denmark*

Abstract: This paper covers the design of an Evolutionary Algorithm (EA), which should be able to synthesize a mixed H_2/H_∞ controller. It will be shown how a system can be expressed as Matrix Inequalities (MI) and these will then be used in the design of the EA. The main objective is to examine whether a mixed H_2/H_∞ controller is feasible, and if so, how the optimal mixed controller might be found.

Keywords: Optimal control, Robust control, Genetic algorithm

1. INTRODUCTION

Over time many complicated control problems have successfully been translated into analytical or otherwise numerically solvable ones. However, the combination of robust (H_∞) and optimal (H_2) control have not yet been solved in a manner that provided useful results. By combining robust and optimal control it might be possible to obtain a controller structure that contains the same ruggedness as a robust controller and the performance of an optimal controller. If possible this would provide control engineers with previously unreachable design possibilities.

Controller synthesis for the mixed H_2/H_∞ problem has previously been attempted (Scherer and Weiland, 2000). The problem was reformulated into an analytical problem using Linear Matrix Inequalities (LMI). However, by reformulating the problem into analytical form some very restrictive constraints were applied. These constraints resulted in far from optimal solutions to the mixed controller problem.

The theory of evolution is well known in the field of biology. How evolution has proved successful in nature, have inspired computer scientists to create intelligent algorithms and programs based on the principles of evolution. This evolutionary approach requires a large amount of computations but is also both powerful and successful. This powerful method opens up for new approaches to previously unsolved or flawed solutions to existing problems. Evolutionary Algorithms (EA) have two major advantages compared to other hill-climbing techniques. First of all they are robust, which means they do not necessarily get stuck at local minima/maxima. Second of all they operate with several solutions at the same time, known as parallel computing, which enables them to cover a search area faster than other numerical methods. Using an EA in combination with a Matrix Inequality (MI) formulation of the mixed H_2/H_∞ problem might result in finding a feasible controller to this complex problem.

The H_2 and the H_∞ problem will be reformulated into MIs. These MIs will then be combined and readied for implementation as part of an EA. The EA will then be designed to synthesize a dynamic discrete-time mixed H_2/H_∞ controller.

Section 2 describes how a given system can be described using MI formulation and how this MI formulation could be written in a way that could easily be implemented in an EA. In section 3 an EA is developed and a description of how the MI constraints could be implemented is given. It is also discussed how the internal workings of the EA, such as the individuals, the fitness function and the crossover and mutation operators, have been designed. In section 4 the experiences obtained by testing the developed EA are presented and future issues are discussed.

2. MATRIX INEQUALITIES FOR H_2 AND H_∞

The system used for this study is given by equation (1).

$$\begin{bmatrix} \boldsymbol{x}(t+1) \\ \boldsymbol{z}_1(t) \\ \boldsymbol{z}_2(t) \\ \boldsymbol{y}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B}_1 & \mathbf{B}_2 & \mathbf{B} \\ \mathbf{C}_1 & \mathbf{D}_{11} & \mathbf{D}_{12} & \mathbf{E}_1 \\ \mathbf{C}_2 & \mathbf{D}_{21} & \mathbf{D}_{22} & \mathbf{E}_2 \\ \mathbf{C} & \mathbf{F}_1 & \mathbf{F}_2 & \mathbf{G} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}(t) \\ \boldsymbol{w}_1(t) \\ \boldsymbol{w}_2(t) \\ \boldsymbol{u}(t) \end{bmatrix} \qquad (1)$$

where $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{D}_{11} \in \mathbb{R}^{p_1\times m_1}$, $\mathbf{D}_{22} \in \mathbb{R}^{p_2\times m_2}$ and $\mathbf{G} \in \mathbb{R}^{q\times l}$. First some requirements for the system must be met (Gahinet and Apkarian, 1994). $(\mathbf{A},\mathbf{B})$ must be stabilizable and $(\mathbf{C},\mathbf{A})$ must be detectable. Setting $\mathbf{G}=0$ is no requirement but will be assumed to simplify calculations without any loss of generality. Then, given any discrete real-rational dynamic controller $\mathbf{K}(z)$ with the realization

$$\mathbf{K}(z) = \mathbf{D}_K + \mathbf{C}_K(z\mathbf{I}-\mathbf{A}_K)^{-1}\mathbf{B}_K \qquad (2)$$

with $\mathbf{A}_K \in \mathbb{R}^{k\times k}$, the closed-loop transfer function from $\boldsymbol{w}_i(t)$ to $\boldsymbol{z}_i(t)$ is found as

$$\boldsymbol{\mathcal{T}}_{c_i} = \mathbf{D}_{c_i} + \mathbf{C}_{c_i}(z\mathbf{I}-\mathbf{A}_c)^{-1}\mathbf{B}_c \qquad i=1,2 \quad (3)$$

where $i=1$ corresponds to the closed-loop transfer function, $\boldsymbol{\mathcal{T}}_{c_1}$, from $\boldsymbol{w}_1(t)$ to $\boldsymbol{z}_1(t)$ and $i=2$ is the corresponding $\boldsymbol{\mathcal{T}}_{c_2}$ from $\boldsymbol{w}_2(t)$ to $\boldsymbol{z}_2(t)$. The mixed H_2/H_∞ controller will be found so that the H_2 norm is minimized for $\boldsymbol{\mathcal{T}}_{c_1}$ and the H_∞ norm is minimized for $\boldsymbol{\mathcal{T}}_{c_2}$. The closed loop expressions are given as

$$\begin{aligned} \mathbf{A}_c &= \begin{bmatrix} \mathbf{A}+\mathbf{B}\mathbf{D}_K\mathbf{C} & \mathbf{B}\mathbf{C}_K \\ \mathbf{B}_K\mathbf{C} & \mathbf{A}_K \end{bmatrix}, \\ \mathbf{B}_c &= \begin{bmatrix} [\mathbf{B}_1\ \mathbf{B}_2] + \mathbf{B}\mathbf{D}_K[\mathbf{F}_1\ \mathbf{F}_2] \\ \mathbf{B}_K[\mathbf{F}_1\ \mathbf{F}_2] \end{bmatrix}, \\ \mathbf{C}_{c_i} &= [\mathbf{C}_i + \mathbf{E}_i\mathbf{D}_K\mathbf{C}\ \ \mathbf{E}_i\mathbf{C}_K], \qquad i=1,2 \\ \mathbf{D}_{c_i} &= [[\mathbf{D}_{i1}\ \mathbf{D}_{i2}] + \mathbf{E}_i\mathbf{D}_K[\mathbf{F}_1\ \mathbf{F}_2]]. \qquad i=1,2 \end{aligned} \qquad (4)$$

Gathering the control parameters into one single variable

$$\boldsymbol{\Theta} = \begin{bmatrix} \mathbf{A}_K & \mathbf{B}_K \\ \mathbf{C}_K & \mathbf{D}_K \end{bmatrix} \in \mathbb{R}^{(n+k)\times(n+k)} \qquad (5)$$

and introducing the shorthands

$$\begin{aligned} \mathbf{A}_0 &= \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{k\times k} \end{bmatrix}, & \mathbf{B}_0 &= \begin{bmatrix} \mathbf{B}_1 & \mathbf{B}_2 \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \\ \mathbf{C}_{0i} &= [\mathbf{C}_i\ \mathbf{0}], & i&=1,2 \\ \mathbf{D}_{0i} &= [\mathbf{D}_{i1}\ \mathbf{D}_{i2}], & i&=1,2 \\ \boldsymbol{\mathcal{B}} &= \begin{bmatrix} \mathbf{0} & \mathbf{B} \\ \mathbf{I}_{k\times k} & \mathbf{0} \end{bmatrix}, & \boldsymbol{\mathcal{C}} &= \begin{bmatrix} \mathbf{0} & \mathbf{I}_{k\times k} \\ \mathbf{C} & \mathbf{0} \end{bmatrix}, \\ \boldsymbol{\mathcal{E}}_i &= [\mathbf{0}\ \mathbf{E}_i], & i&=1,2 \\ \boldsymbol{\mathcal{F}} &= \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{F}_1 & \mathbf{F}_2 \end{bmatrix}. \end{aligned} \qquad (6)$$

results in writing the closed-loop matrices as

$$\begin{aligned} \mathbf{A}_c &= \mathbf{A}_0 + \boldsymbol{\mathcal{B}}\boldsymbol{\Theta}\boldsymbol{\mathcal{C}}, \\ \mathbf{B}_c &= \mathbf{B}_0 + \boldsymbol{\mathcal{B}}\boldsymbol{\Theta}\boldsymbol{\mathcal{F}}, \\ \mathbf{C}_{c_i} &= \mathbf{C}_{0i} + \boldsymbol{\mathcal{E}}_i\boldsymbol{\Theta}\boldsymbol{\mathcal{C}}, \qquad i=1,2 \\ \mathbf{D}_{c_i} &= \mathbf{D}_{0i} + \boldsymbol{\mathcal{E}}_i\boldsymbol{\Theta}\boldsymbol{\mathcal{F}}. \qquad i=1,2 \end{aligned} \qquad (7)$$

At this point an introduction of the projection lemma (Scherer and Weiland, 2000) is useful.

Lemma 1. For arbitrary $\boldsymbol{\mathcal{P}},\boldsymbol{\mathcal{Q}}$ and a symmetric $\boldsymbol{\Psi}$, the MI

$$\boldsymbol{\mathcal{Q}}^T\boldsymbol{\Theta}\boldsymbol{\mathcal{P}} + \boldsymbol{\mathcal{P}}^T\boldsymbol{\Theta}^T\boldsymbol{\mathcal{Q}} + \boldsymbol{\Psi} < \mathbf{0} \qquad (8)$$

in the unstructured $\boldsymbol{\Theta}$ has a solution if and only if

$$\boldsymbol{\mathcal{P}}\boldsymbol{x} = \mathbf{0}\ \textit{or}\ \boldsymbol{\mathcal{Q}}\boldsymbol{x} = \mathbf{0} \ \Rightarrow\ \boldsymbol{x}^T\boldsymbol{\Psi}\boldsymbol{x} < \mathbf{0}\ \textit{or}\ \boldsymbol{x} = \mathbf{0}. \qquad (9)$$

If $\mathbf{W}_{\mathcal{P}}$ and $\mathbf{W}_{\mathcal{Q}}$ denote arbitrary matrices whose columns form a basis of the nullspaces of $\boldsymbol{\mathcal{P}}$ and $\boldsymbol{\mathcal{Q}}$ respectively, denoted Ker($\boldsymbol{\mathcal{P}}$) and Ker($\boldsymbol{\mathcal{Q}}$), (9) is equivalent to

$$\mathbf{W}_{\mathcal{P}}^T\boldsymbol{\Psi}\mathbf{W}_{\mathcal{P}}\ \textit{and}\ \mathbf{W}_{\mathcal{Q}}^T\boldsymbol{\Psi}\mathbf{W}_{\mathcal{Q}}. \qquad (10)$$

From (Gahinet and Apkarian, 1994) the MI for H_2 optimization is given by

$$\begin{bmatrix} -\mathbf{X}_2 & \mathbf{A}_c & \mathbf{0} \\ \mathbf{A}_c^T & -\mathbf{X}_2^{-1} & \mathbf{C}_{c_1}^T \\ \mathbf{0} & \mathbf{C}_{c_1} & -\mathbf{I} \end{bmatrix} < \mathbf{0} \qquad (11)$$

where

$$\mathbf{X}_2 = \mathbf{X}_2^T > \mathbf{0} \qquad (12)$$

and the optimal solution is given by minimizing

$$\mathrm{tr}\left(\mathbf{B}_0^T\mathbf{X}_2^{-1}\mathbf{B}_0\right) < \gamma_2^2 \qquad (13)$$

Using the shorthand of (7), the inequality (11) can be written with $\mathbf{X}_2$ and $\boldsymbol{\Theta}$ grouped terms as

$$\begin{aligned} &\begin{bmatrix} -\mathbf{X}_2 & \mathbf{A}_0 & \mathbf{0} \\ \mathbf{A}_0^T & \mathbf{X}_2^{-1} & \mathbf{C}_{01}^T \\ \mathbf{0} & \mathbf{C}_{01} & -\mathbf{I} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\mathcal{B}} \\ \mathbf{0} \\ \boldsymbol{\mathcal{E}}_1 \end{bmatrix} \boldsymbol{\Theta}\, [\mathbf{0}\ \boldsymbol{\mathcal{C}}\ \mathbf{0}] \\ &\qquad + \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\mathcal{C}}^T \\ \mathbf{0} \end{bmatrix} \boldsymbol{\Theta}^T\, [\boldsymbol{\mathcal{B}}^T\ \mathbf{0}\ \boldsymbol{\mathcal{E}}_1^T] < \mathbf{0} \end{aligned} \qquad (14)$$

Use of the projection lemma further states that (14) is solvable if and only if

$$\mathbf{W}_{\mathcal{P}}^T \begin{bmatrix} -\mathbf{X}_2 & \mathbf{A}_0 & \mathbf{0} \\ \mathbf{A}_0^T & \mathbf{X}_2^{-1} & \mathbf{C}_{01}^T \\ \mathbf{0} & \mathbf{C}_{01} & -\mathbf{I} \end{bmatrix} \mathbf{W}_{\mathcal{P}} < \mathbf{0}, \quad (15)$$

$$\mathbf{W}_{\mathcal{Q}}^T \begin{bmatrix} -\mathbf{X}_2 & \mathbf{A}_0 & \mathbf{0} \\ \mathbf{A}_0^T & \mathbf{X}_2^{-1} & \mathbf{C}_{01}^T \\ \mathbf{0} & \mathbf{C}_{01} & -\mathbf{I} \end{bmatrix} \mathbf{W}_{\mathcal{Q}} < \mathbf{0}, \quad (16)$$

where

$$\mathbf{W}_{\mathcal{P}} = \begin{bmatrix} \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{W}_{\mathcal{P}_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix}, \quad \mathbf{W}_{\mathcal{Q}} = \begin{bmatrix} \mathbf{W}_{\mathcal{Q}_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{W}_{\mathcal{Q}_2} & \mathbf{0} \end{bmatrix}, \quad (17)$$

and with

$$\mathrm{Im}\left[\mathbf{W}_{\mathcal{P}_1}\right] = \mathrm{Ker}\left[\boldsymbol{\mathcal{C}}\right], \quad \mathrm{Im}\begin{bmatrix} \mathbf{W}_{\mathcal{Q}_1} \\ \mathbf{W}_{\mathcal{Q}_2} \end{bmatrix} = \mathrm{Ker}\left[\boldsymbol{\mathcal{B}}^T \ \ \boldsymbol{\mathcal{E}}_1^T\right]. \quad (18)$$

Using the Schur complement (Scherer and Weiland, 2000) on (15) and (16) the conditions for a solution to (14) can be written as

$$-\mathbf{W}_{\mathcal{P}_1}^T \mathbf{X}_2^{-1} \mathbf{W}_{\mathcal{P}_1} + \mathbf{W}_{\mathcal{P}_1}^T \mathbf{C}_{01}^T \mathbf{C}_{01} \mathbf{W}_{\mathcal{P}_1} + (\mathbf{A}_0 \mathbf{W}_{\mathcal{P}_1})^T \mathbf{X}_2^{-1} (\mathbf{A}_0 \mathbf{W}_{\mathcal{P}_1}) < \mathbf{0} \quad (19)$$

and

$$(\mathbf{A}_0^T \mathbf{W}_{\mathcal{Q}_1} + \mathbf{C}_{01}^T \mathbf{W}_{\mathcal{Q}_2})^T \mathbf{X}_2 (\mathbf{A}_0^T \mathbf{W}_{\mathcal{Q}_1} + \mathbf{C}_{01}^T \mathbf{W}_{\mathcal{Q}_2}) - \mathbf{W}_{\mathcal{Q}_1}^T \mathbf{X}_2 \mathbf{W}_{\mathcal{Q}_1} - \mathbf{W}_{\mathcal{Q}_2}^T \mathbf{W}_{\mathcal{Q}_2} < \mathbf{0} \quad (20)$$

Thus, by finding an $\mathbf{X}_2$ and a $\boldsymbol{\Theta}$ that solves equations (14), (19) and (20) and minimizing (13), the H_2 norm for the transfer function, $\boldsymbol{\mathcal{T}}_{c_1}$, can be minimized.

Similarly for H_∞ a controller that fulfills the MI

$$\begin{bmatrix} -\mathbf{X}_\infty^{-1} & \mathbf{A}_c & \mathbf{B}_c & \mathbf{0} \\ \mathbf{A}_c^T & -\mathbf{X}_\infty & \mathbf{0} & \mathbf{C}_{c_2}^T \\ \mathbf{B}_c^T & \mathbf{0} & -\gamma_\infty \mathbf{I} & \mathbf{D}_{c_2}^T \\ \mathbf{0} & \mathbf{C}_{c_2} & \mathbf{D}_{c_2} & -\gamma_\infty \mathbf{I} \end{bmatrix} < \mathbf{0} \quad (21)$$

where

$$\mathbf{X}_\infty = \mathbf{X}_\infty^T > \mathbf{0} \quad (22)$$

is called γ-suboptimal (Gahinet and Apkarian, 1994), and by minimizing γ_∞ the optimal controller can be found.

Writing the H_∞ MI with $\mathbf{X}_\infty$ and $\boldsymbol{\Theta}$ grouped terms yields

$$\begin{bmatrix} -\mathbf{X}_\infty^{-1} & \mathbf{A}_0 & \mathbf{B}_0 & \mathbf{0} \\ \mathbf{A}_0^T & -\mathbf{X}_\infty & \mathbf{0} & \mathbf{C}_{02}^T \\ \mathbf{B}_0^T & \mathbf{0} & -\gamma_\infty \mathbf{I} & \mathbf{D}_{02}^T \\ \mathbf{0} & \mathbf{C}_{02} & \mathbf{D}_{02} & -\gamma_\infty \mathbf{I} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\mathcal{B}} \\ \mathbf{0} \\ \mathbf{0} \\ \boldsymbol{\mathcal{E}}_2 \end{bmatrix} \boldsymbol{\Theta} \begin{bmatrix} \mathbf{0} & \boldsymbol{\mathcal{C}} & \boldsymbol{\mathcal{F}} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\mathcal{C}}^T \\ \boldsymbol{\mathcal{F}}^T \\ \mathbf{0} \end{bmatrix} \boldsymbol{\Theta}^T \begin{bmatrix} \boldsymbol{\mathcal{B}}^T & \mathbf{0} & \mathbf{0} & \boldsymbol{\mathcal{E}}_2^T \end{bmatrix} < \mathbf{0} \quad (23)$$

Again use of the projection lemma states that (23) is solvable if and only if

$$\mathbf{V}_{\mathcal{P}}^T \begin{bmatrix} -\mathbf{X}_\infty^{-1} & \mathbf{A}_0 & \mathbf{B}_0 & \mathbf{0} \\ \mathbf{A}_0^T & -\mathbf{X}_\infty & \mathbf{0} & \mathbf{C}_{02}^T \\ \mathbf{B}_0^T & \mathbf{0} & -\gamma_\infty \mathbf{I} & \mathbf{D}_{02}^T \\ \mathbf{0} & \mathbf{C}_{02} & \mathbf{D}_{02} & -\gamma_\infty \mathbf{I} \end{bmatrix} \mathbf{V}_{\mathcal{P}} < \mathbf{0}, \quad (24)$$

$$\mathbf{V}_{\mathcal{Q}}^T \begin{bmatrix} -\mathbf{X}_\infty^{-1} & \mathbf{A}_0 & \mathbf{B}_0 & \mathbf{0} \\ \mathbf{A}_0^T & -\mathbf{X}_\infty & \mathbf{0} & \mathbf{C}_{02}^T \\ \mathbf{B}_0^T & \mathbf{0} & -\gamma_\infty \mathbf{I} & \mathbf{D}_{02}^T \\ \mathbf{0} & \mathbf{C}_{02} & \mathbf{D}_{02} & -\gamma_\infty \mathbf{I} \end{bmatrix} \mathbf{V}_{\mathcal{Q}} < \mathbf{0}, \quad (25)$$

where

$$\mathbf{V}_{\mathcal{P}} = \begin{bmatrix} \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{V}_{\mathcal{P}_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{V}_{\mathcal{P}_2} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix}, \quad \mathbf{V}_{\mathcal{Q}} = \begin{bmatrix} \mathbf{V}_{\mathcal{Q}_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \\ \mathbf{V}_{\mathcal{Q}_2} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad (26)$$

and with

$$\mathrm{Im}\begin{bmatrix} \mathbf{V}_{\mathcal{P}_1} \\ \mathbf{V}_{\mathcal{P}_2} \end{bmatrix} = \mathrm{Ker}\left[\boldsymbol{\mathcal{C}} \ \ \boldsymbol{\mathcal{F}}\right], \quad \mathrm{Im}\begin{bmatrix} \mathbf{V}_{\mathcal{Q}_1} \\ \mathbf{V}_{\mathcal{Q}_2} \end{bmatrix} = \mathrm{Ker}\left[\boldsymbol{\mathcal{B}}^T \ \ \boldsymbol{\mathcal{E}}_2^T\right]. \quad (27)$$

The use of Schur complement on (24) and (25) yields

$$\gamma_\infty^{-1}(\mathbf{C}_{02}\mathbf{V}_{\mathcal{P}_1} + \mathbf{D}_{02}\mathbf{V}_{\mathcal{P}_2})^T(\mathbf{C}_{02}\mathbf{V}_{\mathcal{P}_1} + \mathbf{D}_{02}\mathbf{V}_{\mathcal{P}_2}) + (\mathbf{A}_0\mathbf{V}_{\mathcal{P}_1} + \mathbf{B}_0\mathbf{V}_{\mathcal{P}_2})^T\mathbf{X}_\infty(\mathbf{A}_0\mathbf{V}_{\mathcal{P}_1} + \mathbf{B}_0\mathbf{V}_{\mathcal{P}_2}) - \mathbf{V}_{\mathcal{P}_1}^T\mathbf{X}_\infty\mathbf{V}_{\mathcal{P}_1} - \gamma_\infty\mathbf{V}_{\mathcal{P}_2}^T\mathbf{V}_{\mathcal{P}_2} < \mathbf{0} \quad (28)$$

and

$$\gamma_\infty^{-1}(\mathbf{B}_0^T\mathbf{V}_{\mathcal{Q}_1} + \mathbf{D}_{02}^T\mathbf{V}_{\mathcal{Q}_2})^T(\mathbf{B}_0^T\mathbf{V}_{\mathcal{Q}_1} + \mathbf{D}_{02}^T\mathbf{V}_{\mathcal{Q}_2}) + (\mathbf{A}_0^T\mathbf{V}_{\mathcal{Q}_1} + \mathbf{C}_{02}^T\mathbf{V}_{\mathcal{Q}_2})^T\mathbf{X}_\infty^{-1}(\mathbf{A}_0^T\mathbf{V}_{\mathcal{Q}_1} + \mathbf{C}_{02}^T\mathbf{V}_{\mathcal{Q}_2}) - \mathbf{V}_{\mathcal{Q}_1}^T\mathbf{X}_\infty^{-1}\mathbf{V}_{\mathcal{Q}_1} - \gamma_\infty\mathbf{V}_{\mathcal{Q}_2}^T\mathbf{V}_{\mathcal{Q}_2} < \mathbf{0} \quad (29)$$

In short it will be possible to synthesize a dynamic controller for the mixed H_2/H_∞ problem if the matrices $\mathbf{X}_2, \mathbf{X}_\infty$ and $\boldsymbol{\Theta}$ that fulfills the MIs in equations (14), (19), (20), (23), (28) and (29) can be found. The mixed H_2/H_∞ controller would then be given by $\boldsymbol{\Theta}$. Furthermore, by minimizing γ_2 in (13) and γ_∞ in (23), (28) and (29) the optimal controller can be found.

3. EVOLUTIONARY ALGORITHM

The purpose of the EA is to search for the matrices $\mathbf{X}_2$, $\mathbf{X}_\infty$ and $\mathbf{\Theta}$ that solve the MIs. By also minimizing γ_2 and γ_∞, the optimal controller $\mathbf{\Theta}$ can be found. Previous attempts to synthesize a mixed H_2/H_∞ controller have used the constraint of setting $\mathbf{X}_2 = \mathbf{X}_\infty$ (Scherer and Weiland, 2000), which is the condition for changing MIs into the analytically solvable LMIs. Another advantage of using EA to solve the problem compared to analytical methods is that the inverse of $\mathbf{X}_2$ and $\mathbf{X}_\infty$ need no special considerations. Analytical methods would have required a reformulation of the problem before a solution could be found.

Both $\mathbf{X}_2$ and $\mathbf{X}_\infty$ are subject to the constraints of (12) and (22) respectively, which means that both matrices must be symmetric and positive definite. To obtain symmetric positive definite matrices the expression

$$\mathbf{X} = \mathbf{M}^T\mathbf{M} \tag{30}$$

is used. By ensuring that $\mathbf{M}$ is real and nonsingular, the resulting $\mathbf{X}$ will be real, symmetric and positive definite. So, by letting the EA search for the real nonsingular matrices $\mathbf{M}_2$ and $\mathbf{M}_\infty$ and using equation (30), the implementation of the EA can be less restrictive with regard to the search domain when finding the matrices $\mathbf{X}_2$ and $\mathbf{X}_\infty$. However, formula (30) is ambiguous and will for matrices $\mathbf{M}$ and $-\mathbf{M}$ produce the same $\mathbf{X}$. Thus, in order to avoid ambiguousity the constraint

$$\det(\mathbf{M}) > 0 \tag{31}$$

should be implemented. Matrices $\mathbf{M}$, that do not meet the constraint in (31) can, however, easily be conformed to meet the constraint by multiplication with -1. Since there are no constraints on $\mathbf{\Theta}$, no special considerations have to met for this matrix when designing the EA.

3.1 *Individuals*

Before designing the fitness function it is necessary to determine how the matrices $\mathbf{M}_2$, $\mathbf{M}_\infty$ and $\mathbf{\Theta}$ should be combined. Having a separate population of matrices for each of the matrices $\mathbf{M}_2$, $\mathbf{M}_\infty$, and $\mathbf{\Theta}$ and determining a fitness value for each of the possible combinations would be infeasible. The number of fitness evaluations in each generation would be exponential in relation to the number of populations and the population sizes. In order to avoid the high number of calculations it would have been necessary to limit the population sizes, thus, also limiting the search space of the EA.

By choosing an individual to consist of a combination of $\mathbf{M}_2$, $\mathbf{M}_\infty$ and $\mathbf{\Theta}$, such that one matrix $\mathbf{M}_2$ is combined with only one matrix $\mathbf{M}_\infty$ and one matrix $\mathbf{\Theta}$, the number of fitness evaluations in each generation will not be exponential, but will be equal to the population size. This will ensure that the EA will have a wide search space, since the population size can be chosen higher than for the combination of all matrices $\mathbf{M}_2$, $\mathbf{M}_\infty$ and $\mathbf{\Theta}$.

The drawback of using the above mentioned combination of $\mathbf{M}_2$, $\mathbf{M}_\infty$ and $\mathbf{\Theta}$ into a single individual is that the possibility of losing matrices that, combined with other matrices, would fulfill the MIs, is high. A matrix, $\mathbf{M}_2$, in a specific individual is dependent on the other matrices, $\mathbf{M}_\infty$ and $\mathbf{\Theta}$, in that individual in order to receive a good fitness. This means that one ill fit matrix and two very fit matrices in an individual will result in a poor fitness value for that individual. However, the drawback can be reduced, which will be described later in section 3.3.

3.2 *Fitness Function*

The fitness function can be defined in many different ways. In this paper the fitness function is chosen to be expressed as adjusted fitness (Koza, 1994). Adjusted fitness is written in the form

$$\boldsymbol{\mathcal{F}}_a = \frac{1}{1+f} \tag{32}$$

where $f \geq 0$ is sought minimized. This results in a maximum value of 1 for the adjusted fitness function, $\boldsymbol{\mathcal{F}}_a$. The reason for choosing adjusted fitness is that, when f approaches 0, the importance of small changes is exaggerated. So, as the population improves, greater emphasis is placed on small differences, thus, making the difference between a good individual and a great one.

The MIs in formulas (14), (19), (20), (23), (28), and (29) all involve negative definiteness. It is then necessary to define a function that maps the fulfillment of an MI, involving negative definiteness, into $\mathbb{R}$. One such function can be defined as

$$f(\cdot) = \begin{cases} \rho \cdot \lambda_{max}(\cdot) + \beta & \text{for} \lambda_{max}(\cdot) \geq 0 \\ 0 & \text{for} \lambda_{max}(\cdot) < 0 \end{cases} \tag{33}$$

where λ_{max} is the largest eigenvalue of the matrix in the MI and ρ and β are penalty factors. The offset β is included since the MIs are strict, and thus a $\lambda_{max} = 0$ cannot be allowed to yield $f(\cdot) = 0$. The slope of $f(\cdot)$ ensures that large positive values of λ_{max} results in high values for $f(\cdot)$ while decreasing values for λ_{max} results in a decreasing value for $f(\cdot)$. Assigning functions $f_1, \ldots, f_6$, of the form $f(\cdot)$, to the MIs in (14), (19), (20), (23), (28), and (29) respectively and inserting into (32) results in the fitness function, $\boldsymbol{\mathcal{F}}(\mathbf{M}_2, \mathbf{M}_\infty, \mathbf{\Theta})$.

$$\mathcal{F}(\mathbf{M}_2, \mathbf{M}_\infty, \mathbf{\Theta}) = \frac{1}{1 + \sum_{i=1}^{6} f_i} \tag{34}$$

By looking at the MIs in formulae (23), (28), and (29) it is seen that γ_∞, which have not yet been defined, is included. γ_∞ could be set as a constant value, however, this would be very restrictive and would limit the possibility of finding a feasible mixed H_2/H_∞ controller using the EA. γ_∞ could also be found iterative, though this might result in having to include γ_∞ as a variable in the individuals in the EA. However, another possibility is to define an expression for γ_∞ based on the existing variables, $\mathbf{M}_2$, $\mathbf{M}_\infty$, and $\mathbf{\Theta}$, thus, indirectly implementing the iteration as part of the EA. By introducing a weighting

$$\zeta = \frac{\gamma_2}{\gamma_\infty} \tag{35}$$

an expression for γ_∞ will be given as

$$\frac{\sqrt{\mathrm{tr}\big(\mathbf{B}_0^T(\mathbf{M}_2^T\mathbf{M}_2)^{-1}\mathbf{B}_0\big)}}{\zeta} < \gamma_\infty. \tag{36}$$

When attempting to find the optimal controller, which will be described later, the weighting, ζ, can be viewed as the factor that determines how much the controller should be optimized for γ_2 compared to γ_∞. It is easily seen that even though (36) is strict, it will be possible to insert the expression for γ_∞ in formula (37) into the MIs containing γ_∞, when $\zeta_\epsilon = \zeta + \epsilon$ and $\epsilon > 0$ and ϵ arbitrarily small.

$$\gamma_\infty = \frac{\sqrt{\mathrm{tr}\big(\mathbf{B}_0^T(\mathbf{M}_2^T\mathbf{M}_2)^{-1}\mathbf{B}_0\big)}}{\zeta_\epsilon} \tag{37}$$

With γ_∞ defined, $\mathcal{F}(\mathbf{M}_2, \mathbf{M}_\infty, \mathbf{\Theta})$ is now fully defined with respect to $\mathbf{M}_2$, $\mathbf{M}_\infty$, and $\mathbf{\Theta}$ and can be calculated. The fitness value of $\mathcal{F}(\mathbf{M}_2, \mathbf{M}_\infty, \mathbf{\Theta})$ indicates how close the individual $(\mathbf{M}_2, \mathbf{M}_\infty, \mathbf{\Theta})$ is to a feasible mixed H_2/H_∞ controller, $\mathbf{\Theta}$. Thus, if $\mathcal{F}(\mathbf{M}_2, \mathbf{M}_\infty, \mathbf{\Theta}) = 1$ then $\mathbf{\Theta}$ is a feasible mixed H_2/H_∞ controller for the system. However, even though $\mathbf{\Theta}$ is a feasible mixed controller it will most likely not be the optimal mixed controller.

As mentioned in section 2, the optimal controller can be found by minimizing γ_2 and γ_∞. From formulae (13) and (37) it is seen that both γ_2 and γ_∞ is expressed by $\mathrm{tr}\big(\mathbf{B}_0^T(\mathbf{M}_2^T\mathbf{M}_2)^{-1}\mathbf{B}_0\big)$, and the degree of optimization of γ_2 compared to γ_∞ is given by ζ. So for $\mathcal{F}(\mathbf{M}_2, \mathbf{M}_\infty, \mathbf{\Theta}) = 1$ and by minimizing $\mathrm{tr}\big(\mathbf{B}_0^T(\mathbf{M}_2^T\mathbf{M}_2)^{-1}\mathbf{B}_0\big)$, the desired optimal mixed controller can be found. The conditions can be combined into a joint fitness function

$$\mathcal{F}_{opt} = \frac{1}{1 + \mathrm{tr}\big(\mathbf{B}_0^T(\mathbf{M}_2^T\mathbf{M}_2)^{-1}\mathbf{B}_0\big) + \sum_{i=1}^{6} f_i} \tag{38}$$

This joint fitness function is, however, not without flaws, and these flaws will be described in detail in section 4.

3.3 *Crossover*

For simplification the crossover operation will be performed so that two parent individuals creates two offspring. Furthermore, when crossover is performed only one matrix type from the parent individuals will be used in the operation, whereas the two remaining matrix types will be transferred directly to the offspring. An example would be that two parents

$$p_1 : ({}^1\mathbf{M}_2, {}^1\mathbf{M}_\infty, {}^1\mathbf{\Theta}),$$
$$p_2 : ({}^2\mathbf{M}_2, {}^2\mathbf{M}_\infty, {}^2\mathbf{\Theta})$$

would result in two offspring

$$o_1 : ({}^1\mathbf{M}_2, {}^2\mathbf{M}_\infty, {}^1\mathbf{\Theta}),$$
$$o_2 : ({}^2\mathbf{M}_2, {}^1\mathbf{M}_\infty, {}^2\mathbf{\Theta}).$$

The probability for which of the three matrix types that is transferred should be equal in order to gain maximum effect of the operation. In this case the probability should thus, be $1/3$.

The interchanging of matrices in the above example reduces the drawbacks mentioned in section 3.1, since recombination of the matrices in the different individuals now will be performed in a limited way. However, in order to add diversity to the population, convex combination of the interchanged matrices will also be performed. Thus, the offspring of the above example would, using convex combination, be

$$o_1 : ({}^1\mathbf{M}_2, \alpha \cdot {}^1\mathbf{M}_\infty + (1-\alpha) \cdot {}^2\mathbf{M}_\infty, {}^1\mathbf{\Theta})$$
$$o_2 : ({}^2\mathbf{M}_2, (1-\alpha) \cdot {}^1\mathbf{M}_\infty + \alpha \cdot {}^2\mathbf{M}_\infty, {}^2\mathbf{\Theta})$$

where $0<\alpha<1$. It should be noted that offspring of the matrices ${}^p\mathbf{M}_2$ and ${}^p\mathbf{M}_\infty$, will result in offspring matrices, ${}^o\mathbf{X}_2$ and ${}^o\mathbf{X}_\infty$, which, after application of this convex crossover operation, will not be convex combinations of the parent matrices, ${}^p\mathbf{X}_2$ and ${}^p\mathbf{X}_\infty$, found from using ${}^p\mathbf{M}_2$ and ${}^p\mathbf{M}_\infty$ in (30). Since $\mathbf{\Theta}$ is used directly in the individuals it can easily be seen that for $\mathbf{\Theta}$ the offspring will actually be a convex combination of the parents. The probability for whether direct transfer or convex combination will be performed on the transferred matrix type could be set to any value, however, a probability of $1/2$ would be reasonable.

3.4 *Mutation*

Two ways of performing mutation on the individuals will be presented in this paper. The first way is to perform the mutation on a single element of one of the matrices, $\mathbf{M}_2$, $\mathbf{M}_\infty$ and $\mathbf{\Theta}$. A Gaussian distributed random number with zero mean and deviation σ_i, $N(0, \sigma_i)$, is added to the element that is selected to be mutated. Since the random number is Gaussian distributed, the probability that the mutation will result in minor changes is high, though it also depends on the size of the deviation σ_i. The deviation σ_i will be based on the very successful Rechenberg's '1/5 success rule' (Eiben *et al.*, 1999), which states that 1/5 of all mutations performed should be successful. If the success rate is lower than 1/5, the deviation is decreased according to

$$\sigma_{i+1} = c \cdot \sigma_i \qquad 0.817 \leq c \leq 1 \tag{39}$$

and for success rates higher than 1/5 the deviation is increased according to

$$\sigma_{i+1} = \sigma_i / c \qquad 0.817 \leq c \leq 1 \tag{40}$$

By noticing that a success rate higher than 1/5 would result from the parents being distributed unevenly around the optimum, the deviation is then increased to compensate for that. Similarly, a success rate lower than 1/5 results from the parents being evenly distributed around the optimum, and the deviation is then decreased in order to heighten precision and increase convergence around the optimum.

In order to obtain further diversity in the population there exist a possibility that an entire matrix in an individual will be multiplied with a scalar value. The scalar value is a Gaussian distributed random number with mean 1 and deviation σ_i, $N(0, \sigma_i)$. In the EA, the probability for an entire matrix to be mutated can be set equal to the probability for mutation of a single element. Thus, the impact on the population when mutating an entire matrix will be limited and will not cause the population to diverge.

4. VALIDATION

After having developed the theory for using EAs to synthesize a mixed H_2/H_∞ controller, an EA was developed to examine the feasibility of this approach. The EA was developed in Java™ and tested on several simple plants. These tests resulted in a variety of experiences.

First, it should be mentioned that it was possible to find a feasible mixed H_2/H_∞ controller for small simple plants. It was also possible optimize the mixed controller, even though the resulting controller might not have been the optimal mixed controller.

The EA requires a vast amount of computations. Since the search area for the EA is very large, the population size had to be above 20 in order to obtain a usable controller for a system with two plant states and one controller state. Larger population sizes resulted in increasingly better results. Expanding the system with either one plant or controller state, resulted in the matrices $\mathbf{M}$ going from 3×3 to 4×4. This resulted in a higher population size needed for finding feasible controllers, due to an increased search area. Thus, using the EA to synthesize a controller for increasingly larger plants, resulted in an exponential reduction in the performance of the EA.

Using only the fitness function given in (38) caused the EA to fail. This was caused by a contradiction between optimizing the term $\mathrm{tr}\big(\mathbf{B}_0^T(\mathbf{M}_2^T\mathbf{M}_2)^{-1}\mathbf{B}_0\big)$ and fulfilling the MIs. The problem was solved by using the fitness function of (34) to search for a feasible controller. When a feasible controller had been found, the fitness function was changed to the one given in (38). By reevaluating the entire population, the new fitness function could then be used to search for the optimal mixed controller. However, the term β used in (33) had to be set to the value of $\mathrm{tr}\big(\mathbf{B}_0^T(\mathbf{M}_2^T\mathbf{M}_2)^{-1}\mathbf{B}_0\big)$ for the first feasible controller found. If this was not done, the MIs would become unfulfilled when the term $\mathrm{tr}\big(\mathbf{B}_0^T(\mathbf{M}_2^T\mathbf{M}_2)^{-1}\mathbf{B}_0\big)$ was being minimized, and the controllers found would be unfeasible.

It is unknown whether feasible controllers can be gathered in several separate areas of the search space. If it is possible, then the developed EA would surely be stuck in the first area of feasible controllers encountered, regardless of whether the optimal controller is contained in that area or not. Thus, the controller found by the developed EA might not be the optimal controller.

5. REFERENCES

Eiben, Agoston E., Robert Hinterding and Zbigniew Michalewicz (1999). Parameter control in evolutionary algorithms. *IEEE Transactions on Evolutionary Computation* **3**(2), 124–141.

Gahinet, Pascal and Pierre Apkarian (1994). Linear matrix inequality approach to h_∞ control. *International Journal of Robust and Nonlinear Control* **4**(4), 421–448.

Koza, John R. (1994). *Genetic Programming.* 2nd. 4 ed.. The MIT Press.

Scherer, Carsten and Siep Weiland (2000). *Linear Matrix Inequalities in Control.* 3rd ed.. Lecture Notes DISC Course.