



# CHALMERS

## Chalmers Publication Library

### **Numerical sensitivity of Linear Matrix Inequalities for shorter sampling periods**

This document has been downloaded from Chalmers Publication Library (CPL). It is the author's version of a work that was accepted for publication in:

**51st IEEE Conference on Decision and Control, CDC 2012, Maui, HI, 10 through 13 December 2012 (ISSN: 0191-2216)**

Citation for the published paper:

Lennartson, B. ; Middleton, R. (2012) "Numerical sensitivity of Linear Matrix Inequalities for shorter sampling periods". 51st IEEE Conference on Decision and Control, CDC 2012, Maui, HI, 10 through 13 December 2012 (Article number 6425801), pp. 4247-4252.

<http://dx.doi.org/10.1109/CDC.2012.6425801>

Downloaded from: <http://publications.lib.chalmers.se/publication/174498>

Notice: Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source. Please note that access to the published version might require a subscription.

Chalmers Publication Library (CPL) offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all types of publications: articles, dissertations, licentiate theses, masters theses, conference papers, reports etc. Since 2006 it is the official tool for Chalmers official publication statistics. To ensure that Chalmers research results are disseminated as widely as possible, an Open Access Policy has been adopted. The CPL service is administrated and maintained by Chalmers Library.

(article starts on next page)

# Numerical Sensitivity of Linear Matrix Inequalities for Shorter Sampling Periods

Bengt Lennartson *Member, IEEE*(\*) and Richard Middleton, *Fellow, IEEE*

**Abstract**—The numerical sensitivity of Linear Matrix Inequalities (LMIs) arising in the  $\mathcal{H}_\infty$  norm computation in discrete time is analyzed. Rapid sampling scenarios are examined comparing both shift and delta operator formulations of the equations. The shift operator formulation is shown in general to be arbitrarily poorly conditioned as the sampling rate increases. The delta operator formulation includes both recentering (to avoid cancellation problems) and rescaling, and avoids these difficulties. However, it is also shown that rescaling of the shift operator formulation gives substantial improvements in numerical conditioning, whilst recentering is of more limited benefit.

## I. INTRODUCTION

Linear Matrix Inequalities (LMIs) have been used extensively during the last decade for a range of control analysis and synthesis problems (see for example [1], [2]). One key reason for this is that LMIs can be solved very efficiently by applying interior-point methods to solve these problems.

LMIs are frequently formulated and solved as semidefinite programming (SDP) problems. One typical control problem is  $\mathcal{H}_\infty$  gain computation, with many other problems as extensions or variations on this method. However, it can be shown that when using the shift operator formulation of a discrete time system, with fast sampling, the LMI problem generically has very poor numerical conditioning. In this paper this problem is examined using the delta operator formulation.

One method, using sensitivity analysis, for studying the differential sensitivity of the solution to SDP problems under perturbation of input parameters has been investigated in e.g. [3], [4]. In this paper two more specific numerical sensitivity problems are analyzed for LMIs based on discrete-time shift operator models. These sensitivity issues are related to scaling and cancellation, two well known numerical problems for shift operator models that are conveniently solved using the delta operator [5]. More recently, delta operator based LMIs have been introduced, often related to  $\mathcal{H}_\infty$  robust control problems [6], [7].

A preliminary numerical analysis of both shift and delta operator LMIs was given in [8]. [9] gives a more detailed

B. Lennartson is with the Department of Signals and Systems, Chalmers University of Technology, Göteborg, Sweden; email: [bengt.lennartson@chalmers.se](mailto:bengt.lennartson@chalmers.se)

R. H. Middleton is with the School of Electrical Eng. & Computer Science, University of Newcastle, Australia; email: [Richard.Middleton@newcastle.edu.au](mailto:Richard.Middleton@newcastle.edu.au)

(\*) Three shorter sabbaticals 2003, 2006 and 2010 at the Centre for Complex Dynamic Systems & Control, University of Newcastle, Australia are very much appreciated.

analysis of the underlying mechanisms causing numerical sensitivity. A scaling of the input and output signals was then introduced in the same way as in [10], [11]. The shift operator based LMI was shown to become nearly singular for short sampling periods, not only for the requested optimal  $\gamma$ -value, but also for  $\gamma$ -values far away from the optimal solution. This ill-conditioned behavior generates significant errors in the computed norm for rapid sampling.

In this paper the basic discrete-time model is initially considered without any signal scaling, where the input signal is assumed to be piece-wise constant between the sampling instants. Then another ill-conditioned behavior appears, namely that the symmetric solution matrix  $P_q$  of the shift operator LMI increases with  $1/h$  for small sampling periods  $h$ . This fact makes it difficult to solve the shift operator LMI for small  $h$ . In fact, the increase of  $P_q$  for small  $h$  is shown to have more severe consequences than the more well known cancellation, which occurs both in LMIs and corresponding Riccati equations using shift operator models [8], [11], [5]. The problem with the increasing  $P_q$  can be avoided by solving the LMI for the scaled matrix  $P = hP_q$ . This means that the cancellation problem can be separated, and it is then shown to be of much less importance than the increasing  $P_q$  for small sampling periods  $h$ .

The paper starts with a brief presentation of the shift and delta operator models and their relationships. The corresponding LMIs are then given, and the numerical sensitivity of the different LMIs is analyzed. Finally, the different error contributions are illustrated by an example, where two well known LMI solvers are evaluated.

## II. SHIFT AND DELTA OPERATOR MODELS

The shift and the delta operator models are briefly presented in this section, including some useful transformations between them.

### A. Shift operator model

Consider the following discrete-time state space model on the shift operator form

$$\begin{bmatrix} qx(t_k) \\ y(t_k) \end{bmatrix} = \begin{bmatrix} A_q & B_q \\ C & D \end{bmatrix} \begin{bmatrix} x(t_k) \\ u(t_k) \end{bmatrix} \triangleq G_q \begin{bmatrix} x(t_k) \\ u(t_k) \end{bmatrix} \quad (1)$$

where the shift operator  $q$  is defined as  $qx(t_k) = x(t_{k+1})$ . The state vector  $x$ , the input signal  $u$  and the output signal  $y$  have dimensions  $n$ ,  $n_u$  and  $n_y$ , respectively. The discrete-time updates occur at times  $t_k$ ,  $k = 0, 1, 2, \dots$ , where the

time interval between two updates is the sampling period  $h = t_{k+1} - t_k$ .

### B. Delta operator model

By introducing the delta operator, cf. [5]

$$\begin{aligned}\delta x(t_k) &= (x(t_{k+1}) - x(t_k))h^{-1} \\ &= (A_q - I)h^{-1}x(t_k) + B_q h^{-1}u(t_k)\end{aligned}$$

the state space model (1) can be rewritten in delta operator form as

$$\begin{bmatrix} \delta x(t_k) \\ y(t_k) \end{bmatrix} = \begin{bmatrix} A_\delta & B_\delta \\ C & D \end{bmatrix} \begin{bmatrix} x(t_k) \\ u(t_k) \end{bmatrix} \triangleq G_\delta \begin{bmatrix} x(t_k) \\ u(t_k) \end{bmatrix} \quad (2)$$

Observe that the delta operator model is an exact representation of the discrete time system, assuming that the input signal is piece-wise constant as in the shift operator model. This can be compared with e.g. the discrete Euler approximation, where the system matrix  $I + hA_c$  is an approximation of  $A_q$ . The relation between the state space matrices on shift and delta operator form are

$$G_q = \begin{bmatrix} A_q & B_q \\ C & D \end{bmatrix} = \begin{bmatrix} I_n + hA_\delta & hB_\delta \\ C & D \end{bmatrix} \quad (3)$$

Introducing the matrices  $E = \text{diag}(I_n, 0)$  and  $T_h = \text{diag}(hI_n, I)$ , where the dimension of the lower identity matrix in  $T_h$  is given by the adjacent matrices, the relation between the the shift and delta operator form can be shortly expressed as  $G_q = E + T_h G_\delta$ .

There are two main reasons for introducing the delta operator for discrete-time models. Firstly, there is no natural transition and convergence from an ordinary discrete-time model in the shift operator  $q$  to the corresponding continuous-time model. Secondly, the shift operator exhibits bad numerical behavior for short sampling periods. Both problems are naturally solved by the delta operator.

### C. Convergence to continuous-time model

Assume a related continuous-time model

$$G_c = \begin{bmatrix} A_c & B_c \\ C & D \end{bmatrix} \quad (4)$$

and introduce the function  $\Gamma(h) = \sum_{k=0}^{\infty} (A_c h)^k / (k+1)!$ , which converges to  $\Gamma(h) = I + O(h)$  for short sampling periods. Furthermore, assume a piecewise constant input signal. Then the delta operator model  $G_\delta$  can be expressed as

$$G_\delta = \begin{bmatrix} \Gamma(h) & 0 \\ 0 & I \end{bmatrix} G_c \rightarrow G_c \quad \text{when } h \rightarrow 0 \quad (5)$$

On the other hand, the shift operator system matrix  $A_q = e^{A_c h} \approx I + A_c h \rightarrow I$  when  $h \rightarrow 0$  independently of  $A_c$ . This means that all eigenvalues of  $A_q$  converge to 1, and the information from the system behavior in  $A_c$  is completely lost when  $h \rightarrow 0$ .

These expressions explicitly show the convergence of the delta operator model  $G_\delta$  to its continuous-time counterpart  $G_c$ , and the bad numerical behavior of the shift operator model, see further comments in Middleton and Goodwin [5].

## III. LINEAR MATRIX INEQUALITIES

Computation of the  $\mathcal{H}_\infty$  norm using linear matrix inequalities (LMIs) is briefly presented in this section, both for systems on shift and delta operator form. It is well known that for a stable system  $\mathcal{G}$ , with input  $u$  and output  $y$ , the  $\mathcal{H}_\infty$  norm  $\|\mathcal{G}\|_\infty$  is given by the induced norm, cf. [12]  $\|\mathcal{G}\|_\infty = \sup_{\|u\| \neq 0} \frac{\|y\|}{\|u\|}$ . This norm can be calculated by solving linear matrix inequalities (LMIs) [13], [1].

### A. Shift operator LMI

For the shift operator case, the following lemma shows how to solve the  $\mathcal{H}_\infty$  norm.

*Lemma 1:* Consider a stable discrete-time system  $\mathcal{G}$  on shift operator form  $G_q$  (1). The  $\mathcal{H}_\infty$  norm  $\|\mathcal{G}\|_\infty < \gamma$ , if and only if there exists a symmetric matrix  $P_q$  such that

$$F_q(P_q, \gamma) = \text{diag}(-M_q(P_q, \gamma), P_q) > 0 \quad (6)$$

where

$$M_q(P_q, \gamma) = \begin{bmatrix} M_{q11}(P_q) & M_{q12}(P_q) \\ M'_{q12}(P_q) & M_{q22}(P_q, \gamma) \end{bmatrix} \quad (7)$$

and

$$\begin{aligned}M_{q11}(P_q) &= A'_q P_q A_q - P_q + C' C \\ M_{q12}(P_q) &= A'_q P_q B_q + C' D\end{aligned} \quad (8)$$

$$M_{q22}(P_q, \gamma) = B'_q P_q B_q + D' D - \gamma^2 I \quad \square$$

The minimal value of  $\gamma$  is obtained at the same time as the unknown  $P_q > 0$  is computed. This result is based on the bounded real lemma, see e.g. [13], where  $\|\mathcal{G}\|_\infty < \gamma$  iff the Riccati inequality  $M_{q11}(P_q) - M_{q12}(P_q)M_{q22}^{-1}(P_q, \gamma)M'_{q12}(P_q) < 0$  is satisfied for  $P_q > 0$  and  $M_{q22}(P_q, \gamma) < 0$ . A Schur complement then gives the LMI (6).

### B. Delta operator LMI

In the same way as for the shift operator case, an LMI can be formulated for delta operator models. The corresponding Riccati equation can be found in e.g. [5], [11]. Ones again, a Schur complement gives the following result.

*Lemma 2:* Consider a stable discrete-time system  $\mathcal{G}$  in delta operator form  $G_\delta$  (2). The  $\mathcal{H}_\infty$  norm  $\|\mathcal{G}\|_\infty < \gamma$ , if and only if there exists a symmetric matrix  $P$  such that

$$F_\delta(P, \gamma) = \text{diag}(-M_\delta(P, \gamma), P) > 0 \quad (9)$$

where

$$M_\delta(P, \gamma) = \begin{bmatrix} M_{\delta11}(P) & M_{\delta12}(P) \\ M'_{\delta12}(P) & M_{\delta22}(P, \gamma) \end{bmatrix} \quad (10)$$

and

$$\begin{aligned}M_{\delta11}(P) &= A'_\delta P + P A_\delta + h A'_\delta P A_\delta + C' C \\ M_{\delta12}(P) &= P B_\delta + h A'_\delta P B_\delta + C' D\end{aligned} \quad (11)$$

$$M_{\delta22}(P, \gamma) = h B'_\delta P B_\delta + D' D - \gamma^2 I \quad \square$$

Similar LMI's based on the delta operator can be found in e.g. [6], [14].

### C. Relation between delta operator and continuous LMIs

For a continuous-time system  $G_c$  (4), the LMI matrix corresponding to  $F_\delta(P, \gamma)$  is

$$F_c(P, \gamma) = \text{diag}(-M_c(P, \gamma), P) > 0 \quad (12)$$

where

$$M_c(P, \gamma) = \begin{bmatrix} A'_c P + P A_c + C' C & P B_c + C' D \\ B'_c P + D' C & D' D - \gamma^2 I \end{bmatrix} \quad (13)$$

Using (5) and (9)-(11), it means that

$$F_\delta(P, \gamma) = F_c(P, \gamma) + O(h)$$

Hence, the LMI formulation in the delta operator form illustrates very explicitly the convergence of the discrete-time solution to the corresponding continuous-time one. This is an expected but also essential convergence property in the following numerical sensitivity analysis.

### D. Relation between shift and delta operator LMIs

In the original shift operator model  $G_q$  in (3) recall that  $A_q = I + hA_\delta$ ,  $B_q = hB_\delta$ . This means that the block matrices (8) in  $M_q$  alternatively, using (11), can be expressed as

$$\begin{aligned} M_{q11}(P_q) &= A'_\delta P_q h + P_q h A_\delta + h A'_\delta P_q h A_\delta + C' C \\ &= M_{\delta11}(P_q h) \end{aligned}$$

$$M_{q12}(P_q) = (I + hA_\delta) P_q h B_\delta + C' D = M_{\delta12}(P_q h)$$

$$M_{q22}(P_q, \gamma) = h B'_\delta P_q h B_\delta + D' D - \gamma^2 I = M_{\delta22}(P_q h, \gamma)$$

Hence, we find that

$$M_q(P_q, \gamma) = M_\delta(P, \gamma) \quad (14)$$

where  $P = P_q h$ . Introducing this result in (6) and (9), we obtain the following result.

*Lemma 3:* The solution  $P_q > 0$  to the LMI  $F_q(P_q, \gamma) > 0$  for the system  $G_q$  (3) can alternatively be obtained as

$$P_q = \frac{P}{h} \quad (15)$$

where  $P > 0$  is the solution to the LMI  $F_\delta(P, \gamma) > 0$ . Furthermore, since  $F_\delta \rightarrow F_c$  when  $h \rightarrow 0$  the solution  $P > 0$  converges to the corresponding continuous-time solution. This implies that  $P_q = P/h$  increases without bound when  $h \rightarrow 0$ .  $\square$

Utilizing (15) in (6) and (14) means that  $F_q(P_q, \gamma)$  can alternatively be expressed as

$$\begin{aligned} F_q(P_q, \gamma) &= \text{diag}(-M_q(P/h, \gamma), P/h) \\ &= \text{diag}(-M_\delta(P, \gamma), P/h) \end{aligned} \quad (16)$$

The minimal  $\gamma$ -value of the LMI  $F_q(P_q, \gamma) > 0$  is normally obtained by an interior-point method [2], where the barrier function  $\phi(P_q, \gamma) = -\log \det F_q(P_q, \gamma)$  is introduced. Starting with a feasible solution such that  $F_q(P_q, \gamma) > 0$  means that  $\det F_q(P_q, \gamma) > 0$ . Decreasing  $\gamma$  means that finally  $\det F_q(P_q, \gamma)$  gets close to zero, and at the optimum  $F_q(P_q, \gamma)$  is approximately singular and  $\det F_q(P_q, \gamma) \approx 0$ .

On the other hand,  $\det F_q(P_q, \gamma) = \det(-M_\delta(P, \gamma)) \det(P/h) = \det(-M_\delta(P, \gamma)) O(1/h)$ . Hence, it is hard to determine if  $\det F_q(P_q, \gamma) \approx 0$  for small sampling periods, since this condition then includes a multiplication of a small value  $\det(-M_\delta(P, \gamma)) \approx 0$  with a large value  $\det(P/h) = O(1/h)$ . This analysis shows that the evaluation of the optimality condition is an ill-conditioned problem for short sampling periods, which is also confirmed by the numerical investigation in Section V.

### E. Scaled shift operator LMI

The ill-conditioned property of the LMI  $F_q(P_q, \gamma) > 0$  is easily avoided by observing that this condition according to (16) is based on the two conditions  $M_q(P_q, \gamma) = M_q(P/h, \gamma) < 0$  and  $P_q = P/h > 0$ . Since the latter one can be simplified to  $P > 0$ , the LMI  $F_q(P_q, \gamma) > 0$  can be replaced by the scaled shift operator LMI

$$F_S(P, \gamma) = \text{diag}(-M_q(P/h, \gamma), P) > 0 \quad (17)$$

where  $P = O(1)$  is the unknown matrix and we remind that also  $M_q(P/h, \gamma) = M_\delta(P, \gamma) = O(1)$ . Hence,  $F_S$  is numerically well behaved except for the cancellation in  $M_{q11}$ .

### F. Shift operator LMI without cancellation

The block matrix  $M_{q11} = A'_q P_q A_q - P_q + C' C$  in (8) includes for short sampling periods a cancellation between two large matrices. Since then  $A_q = I + O(h)$  and  $P_q = P/h$ , we find that

$$\begin{aligned} M_{q11} &= (I + O(h))' P/h (I + O(h)) - P/h + C' C \\ &= P_h/h - P/h + C' C \end{aligned}$$

where  $P_h \approx P = O(1)$ . This cancellation can be avoided either by using the delta operator formulation, or simply by introducing  $A_\Delta = hA_\delta$  and replacing the block matrix  $M_{q11}$  in  $M_q$  by

$$M_{\Delta11} = A'_\Delta P_q + P_q A_\Delta + A'_\Delta P_q A_\Delta + C' C \quad (18)$$

The LMI  $F_q(P_q, \gamma) > 0$  is then reformulated as

$$F_\Delta(P_q, \gamma) = \text{diag}(-M_\Delta(P_q, \gamma), P_q) \quad (19)$$

where

$$M_\Delta(P_q, \gamma) = \begin{bmatrix} M_{\Delta11}(P_q) & M_{q12}(P_q) \\ M'_{q12}(P_q) & M_{q22}(P_q, \gamma) \end{bmatrix} \quad (20)$$

Note that algebraically  $F_\Delta = F_q$ , but numerically the cancellation problem is avoided in  $F_\Delta$ .

In fact, we have now separated the two properties of the delta operator from an LMI perspective. The cancellation is avoided in  $F_\Delta$  (20), and the system scaling, including the  $1/h$  factor in both  $A_\delta$  and in the delta operator, is introduced in  $F_S$  by replacing  $P_q$  by  $P = hP_q$  as the unknown matrix.

## IV. ERROR ANALYSIS

We will now investigate the numerical error in the computation of the LMI matrices  $F_q$ ,  $F_\Delta$ ,  $F_S$ , and  $F_\delta$ . Especially the contribution from the cancellation in  $F_q$  and  $F_S$  will be analyzed as a function of the sampling period  $h$ .

### A. Errors in the LMI matrices

Motivated by floating point arithmetic implementations, we will use a relative error analysis [15]. It is well known that subtraction between two uncertain numbers being almost equal yields cancellation of digits. To be more precise, let  $a^\epsilon$  and  $b^\epsilon$  be stored representations of two numbers  $a$  and  $b$ . Then  $a^\epsilon = (1 + \epsilon_a)a$  and  $b^\epsilon = (1 + \epsilon_b)b$ , where  $|\epsilon_a| \leq \mu$  and  $|\epsilon_b| \leq \mu$ , and  $\mu$  is the machine precision ( $\mu \approx 2 \cdot 10^{-16}$  in MATLAB). Now assume that the subtraction is performed according to the IEEE-standard, [15]. Then the stored result is  $(a - b)^\epsilon = (1 + \epsilon_s)(a^\epsilon - b^\epsilon)$  with  $|\epsilon_s| \leq \mu$ .

For matrices similar expressions can be formulated introducing the Hadamard (entry-wise) matrix multiplication [16], the one matrix  $\mathbf{1}$ , and the relative error matrix  $\mathcal{E}$ , where  $[\mathbf{1}]_{ij} = 1$  and  $[\mathcal{E}]_{ij} = \epsilon_{ij}$ . Then  $A^\epsilon = (\mathbf{1} + \mathcal{E}_A) \circ A$  and  $(A - B)^\epsilon = (\mathbf{1} + \mathcal{E}_s) \circ ((\mathbf{1} + \mathcal{E}_A) \circ A - (\mathbf{1} + \mathcal{E}_B) \circ B)$ . Neglecting the quadratic error terms  $\mathcal{E}_s \circ \mathcal{E}_A \circ A$  and  $\mathcal{E}_s \circ \mathcal{E}_B \circ B$ , the error in the matrix subtraction can be expressed as

$$(A - B)^\epsilon = (\mathbf{1} + (\mathcal{E}_s + \mathcal{E}_A)) \circ (A - B) + (\mathcal{E}_A - \mathcal{E}_B) \circ B \quad (21)$$

Apply this formulation on  $M_{q11}$  in (8), by letting  $A = A'_q P_q A_q + C'_q C_q$  and  $B = P_q$ , which leads to  $A - B = M_{q11}$ . Since the analysis is focused on the cancellation, it is assumed for simplicity that  $A'_q P_q A_q + C'_q C_q$  is computed without error, but stored with a relative error as well as  $P_q$ . Based on (21), the error in the computation of  $M_{q11}$  can then be expressed as

$$M_{q11}^\epsilon(P_q) = (\mathbf{1} + \mathcal{E}_{q11}) \circ M_{q11}(P_q) + \mathcal{E}_P \circ P_q \quad (22)$$

where the relative errors  $\mathcal{E}_{q11}$  and  $\mathcal{E}_P$  are of size  $O(\mu)$ . The other block matrices in  $M_q$  are computed without cancellation, and can therefore be simplified to  $M_{q12}^\epsilon(P_q) = (\mathbf{1} + \mathcal{E}_{q12}) \circ M_{q12}(P)$  and  $M_{q22}^\epsilon(P, \gamma) = (\mathbf{1} + \mathcal{E}_{q22}) \circ M_{q22}(P, \gamma)$ , where  $\mathcal{E}_{q12} = O(\mu)$  and  $\mathcal{E}_{q22} = O(\mu)$ . Introduce the error matrix due to the cancellation in  $M_{q11}$

$$P^\epsilon = \text{diag}(\mathcal{E}_P \circ P, 0_{n_u \times n_u}) \quad (23)$$

Since  $P = O(1)$ , we observe that  $P^\epsilon = O(\mu)$ . Together with (22), reminding that  $P_q = P/h$  and  $M_q(P_q, \gamma) = M_\delta(P, \gamma)$ ,  $M_q^\epsilon(P_q, \gamma)$  can now be formulated as  $M_q^\epsilon(P/h, \gamma) = (\mathbf{1} + \mathcal{E}) \circ M_\delta(P, \gamma) + P^\epsilon/h$ , where  $\mathcal{E} = \begin{bmatrix} \mathcal{E}_{q11} & \mathcal{E}_{q12} \\ \mathcal{E}_{q12} & \mathcal{E}_{q22} \end{bmatrix} = O(\mu)$ . Based on this type of analysis the errors in the four LMI matrices  $F_q$  (6),  $F_\Delta$  (19),  $F_S$  (17), and  $F_\delta$  (9) are now be presented in the following theorem.

**Theorem 4:** Assume that  $M_\delta$  is computed and stored with relative error  $\mathcal{E}$  as

$$M_\delta^\epsilon(P, \gamma) = (\mathbf{1} + \mathcal{E}) \circ M_\delta(P, \gamma) \quad (24)$$

where  $\mathcal{E} = O(\mu)$ . The corresponding errors in  $F_q$ ,  $F_\Delta$ ,  $F_S$ , and  $F_\delta$  can then be expressed as

$$F_q^\epsilon(P/h, \gamma) = \text{diag}(-M_\delta^\epsilon(P, \gamma) - P^\epsilon/h, P/h) \quad (25)$$

$$F_\Delta^\epsilon(P/h, \gamma) = \text{diag}(-M_\delta^\epsilon(P, \gamma), P/h) \quad (26)$$

$$F_S^\epsilon(P, \gamma) = \text{diag}(-M_\delta^\epsilon(P, \gamma) - P^\epsilon/h, P) \quad (27)$$

$$F_\delta^\epsilon(P, \gamma) = \text{diag}(-M_\delta^\epsilon(P, \gamma), P) \quad (28)$$

where the error matrix due to the cancellation  $P^\epsilon = O(\mu)$ , and the storage error in the lower diagonal matrix in the LMI matrices  $F_x^\epsilon$  is neglected.  $\square$

This theorem highlights the fact that there are mainly two error sources in the shift operator based LMI calculations. The first has to do with the large value of  $P_q = P/h = O(1/h)$  for shorter sampling periods, resulting in ill-conditioned matrices  $F_q$  and  $F_\Delta$ . The other error source is the cancellation in  $M_q$ , resulting in the error term  $P^\epsilon/h$  in  $F_q$  and  $F_S$ . Both these error sources are avoided in the delta operator version  $F_\delta$ .

### B. Error sensitivity in the objective function

To investigate in more detail how the cancellation error influences the optimization, first assume that the LMI condition  $F > 0$  is handled by introducing a barrier function. The original minimization criterion  $\gamma$  is then replaced by the approximation

$$f(P, \gamma, \epsilon) = \theta\gamma - \log \det F(P, \gamma, \epsilon) \quad (29)$$

where the approximation error is reduced when the parameter  $\theta$  is increased [2]. Then the partial derivative

$$\begin{aligned} \frac{\partial f(P, \gamma, \epsilon)}{\partial \epsilon_i} &= -\frac{1}{\det F(P, \gamma, \epsilon)} \frac{\partial}{\partial \epsilon_i} \det F(P, \gamma, \epsilon) \\ &= -\text{tr}(F^{-1}(P, \gamma, \epsilon) \frac{\partial}{\partial \epsilon_i} F(P, \gamma, \epsilon)) \end{aligned} \quad (30)$$

This partial derivative is now analyzed for the scaled LMI problem  $F_S > 0$ , which only includes the cancellation error but not the ill-conditioning problem. The error term due to the cancellation  $P^\epsilon/h$ , defined in (23), depends on the error matrix  $\mathcal{E}_P$ . Hence, we investigate (30) with respect to the elements in this matrix  $\epsilon_{p_{ij}} = [\mathcal{E}_P]_{i,j}$ . First consider (27), where we only include the error matrix  $\mathcal{E}_P$  but not  $\mathcal{E}$  in  $M_\delta^\epsilon$ . This means that  $M_\delta^\epsilon$  is simplified to  $M_\delta$ , and we obtain

$$\begin{aligned} F_S(P, \gamma, \mathcal{E}_P) &= \text{diag}(-M_\delta(P, \gamma) - \frac{P^\epsilon}{h}, P) \\ &= \text{diag}(-M_\delta(P, \gamma) - \frac{1}{h} \text{diag}(\mathcal{E}_P \circ P, 0_{n_u \times n_u}), P) \end{aligned}$$

which gives

$$\left[ \frac{\partial F_S(P, \gamma, \mathcal{E}_P)}{\partial \epsilon_{p_{ij}}} \right]_{k,l} = \begin{cases} -\frac{p_{ij}}{h} & k = i, \ell = j \\ 0 & \text{otherwise} \end{cases}$$

where  $p_{ij} = [P]_{ij}$ . Since the inverse  $F(P, \gamma, \mathcal{E}_P)^{-1} = \text{diag}(- (M_\delta(P, \gamma) + P^\epsilon/h)^{-1}, P^{-1})$ , the partial derivative of the objective function (30) can now be expressed as

$$\frac{\partial f(P, \gamma, \mathcal{E}_P)}{\partial \epsilon_{p_{ij}}} = - \left[ \left( M_\delta(P, \gamma) + \frac{P^\epsilon}{h} \right)^{-1} \right]_{j,i} \frac{p_{ij}}{h}$$

This result is used in the following theorem.

**Theorem 5:** For the scaled LMI problem  $F_S > 0$  the sensitivity of the approximative objective function  $f(P, \gamma, \epsilon)$

in (29), with respect to the cancellation errors  $\epsilon_{p_{ij}} = [\mathcal{E}_P]_{i,j}$  in  $P^\epsilon$ , can approximately be determined as

$$\begin{aligned} f(P, \gamma, \epsilon_{p_{ij}}) &\approx f(P, \gamma, 0) + \epsilon_{p_{ij}} \frac{\partial f(P, \gamma, \mathcal{E}_P)}{\partial \epsilon_{p_{ij}}} \Big|_{\mathcal{E}_P=0} \\ &= f(P, \gamma, 0) + \epsilon_{p_{ij}} \frac{c_{ij}}{h} \end{aligned}$$

for  $i, j = 1, \dots, n$ , where

$$c_{ij} = -p_{ij} [M_\delta(P, \gamma)^{-1}]_{j,i}$$

□

Observe that  $f(P, \gamma, 0)$  becomes large at the optimum due to a large  $\theta$ , but also  $c_{ij}$ , since it includes the inverse of  $M_\delta$  that is near-singular at the optimum. This is valid for arbitrary sampling periods, which shows that the objective function will be sensitive to the error in  $P^\epsilon$  for sufficiently small sampling periods.

Now introduce a relative error in the determination of the optimal  $\gamma$  value for the actual LMI problems

$$e_\gamma = \frac{|\gamma - \gamma_0|}{\gamma_0} \quad (31)$$

where  $\gamma_0$  is the true optimal value and  $\gamma$  is the optimal value computed by the different LMIs. As already observed, for the scaled LMI problem  $F_S > 0$ , the objective function becomes more and more sensitive to the error in  $P^\epsilon$  due to cancellation for shorter sampling periods, according to Theorem 5. Since this error sensitivity is proportional to  $1/h$ , and the determination of the optimal  $\gamma$  value is based on this objective function, the relative error  $e_\gamma^S$  for the LMI  $F_S > 0$  is also expected to increase with the same factor  $1/h$ , that is

$$e_\gamma^S \approx \frac{\epsilon_S}{h} \quad \text{for small } h \quad (32)$$

where  $\epsilon_S$  is a constant factor. This error function is verified in the next section, where two different SDP solvers generate the same error behavior for shorter sampling periods  $h$ , and the value of  $\epsilon_S$  is shown to be of the same order as the machine precision.

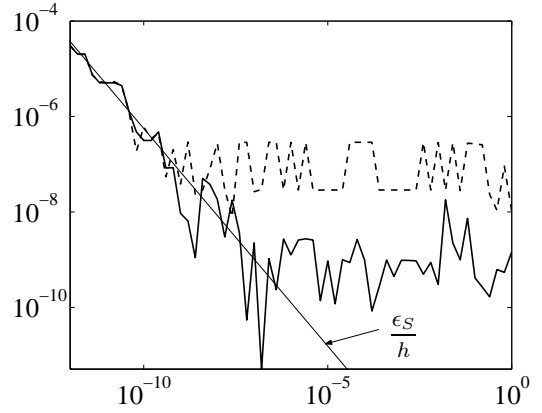
## V. NUMERICAL ILLUSTRATIONS

Consider the following resonant dynamic system with continuous-time transfer function

$$G_c(s) = \frac{12}{(s+1)(s^2+0.2s+1)(s^2+0.4s+4)}$$

where the accuracy in the computation of the corresponding discrete-time  $\mathcal{H}_\infty$  norm will be evaluated. This is done by solving the four different LMIs considered in this paper for various sampling periods  $h \leq 1$  using two different SDP solvers.

First a balanced state-space model is generated using the MATLAB function `ssbal`. Corresponding discrete-time state-space models with zero-order hold circuit at the control input are then computed for different sampling periods, both in shift and delta operator versions. Finally, the minimal  $\gamma$  value is computed for the different LMIs  $F_q$  (6),  $F_\Delta$  (19),  $F_S$  (17), and  $F_\delta$  (9), resulting in corresponding relative errors  $e_\gamma^q$ ,  $e_\gamma^\Delta$ ,  $e_\gamma^S$ , and  $e_\gamma^\delta$ .



**Fig. 1** Relative error as a function of the sampling period  $h$  for the SDP solvers SeDuMi (solid) and SDPA (dashed), when  $\gamma$  is minimized for the scaled LMI  $F_S > 0$ .

The LMIs are solved by two different semi definite programming (SDP) solvers SeDuMi [17] and SDPA [18]. They are all run on top of MATLAB via the user friendly interface YALMIP [19].

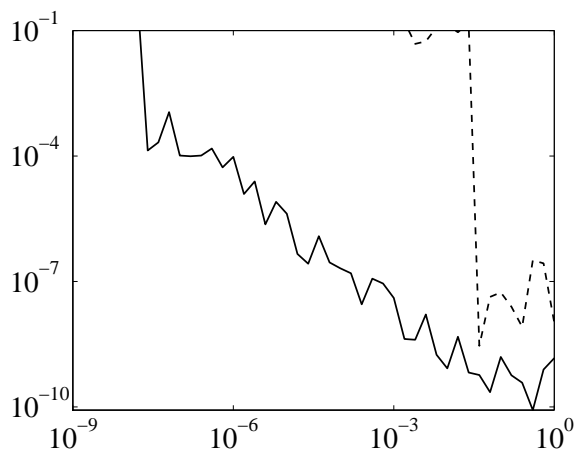
In order to obtain a correct relative error  $e_\gamma$  (31) the true gamma value  $\gamma_0$  is determined by the maximum frequency response based on the delta operator with a very dense grid (20000 points around the maximum at  $\approx 1$  rad/s). As an extra check, this result is compared to the continuous-time infinity norm function in MATLAB for very short sampling periods, and the corresponding discrete one for longer sampling periods, with a very small relative difference around  $10^{-11}$ .

The resulting relative errors  $e_\gamma^S$  for the scaled LMI  $F_S > 0$  are shown in Fig. 1 for the two different solvers. Since the cancellation remains in  $F_S$  the error  $e_\gamma^S$  increases for shorter sampling periods for both solvers approximately as  $\epsilon_S/h$ , the thin line in Fig. 1. The constant  $\epsilon_S \approx 5 \cdot 10^{-17}$ , which is of the same order as the machine precision in MATLAB ( $2 \cdot 10^{-16}$ ). This is reasonable since the error source is a cancellation. The result also coincide with the analysis in Section IV and especially Theorem 5 and (32).

Furthermore, note that for longer sampling periods the relative error depends more on the general accuracy of the individual LMI solver. No adjustments of the tuning parameters for the two solvers have been performed, and we observe that the solver SDPA is generally tuned to achieve less relative accuracy ( $\approx 10^{-7}$ ) than the SeDuMi solver.

Corresponding relative errors for the delta operator case agree with the more constant behavior of  $e_\gamma^S$  for longer sampling periods, with  $e_\gamma^\delta$  around  $10^{-7}$  for SDPA and  $10^{-8} - 10^{-9}$  for the SeDuMi solver. The difference is that the delta operator LMI behaves equally well also for very short sampling periods. For all practical choices of sampling periods we find on the other hand that the scaling mechanism in  $F_S$  is sufficient to achieve perfect numerical results. The cancellation error simply shows up only for unrealistically short sampling periods.

The ill-conditioned behavior in both  $F_q$  and  $F_\Delta$  is however



**Fig. 2** Relative error as a function of the sampling period  $h$  for the SDP solvers SeDuMi (solid) and SDPA (dashed), when  $\gamma$  is minimized for the shift operator LMI  $F_q > 0$ .

shown to be much more severe. The relative errors for the shift operator case  $e_\gamma^q$  are shown in Fig. 2. It is evident that especially the SDPA solver has great difficulties to generate accurate results also for moderate sampling periods. The removal of the cancellation in  $F_\Delta$  does not change the error behavior. The same severe errors as for the shift operator LMI are shown for shorter sampling periods. The reason is that the cancellation error according to Fig. 1 appears at much shorter sampling periods than the error caused by the ill-conditioning problem, still included in  $F_\Delta$ .

To summarize, a numerically robust solution for shorter sampling periods is to use the delta operator model, or to introduce the simple scaling of the unknown  $P$ -matrix. Both approaches work fine for all practical choices of sampling periods.

## VI. CONCLUSIONS

Numerical properties have been analyzed when the  $\mathcal{H}_\infty$  norm is calculated for discrete-time systems by Linear Matrix Inequalities (LMIs). In particular, the behavior for shorter sampling periods has been investigated. By analysis and numerical illustrations it has been shown that there are two main error sources when systems are modeled by the ordinary discrete-time shift operator. The quite well-known cancellation problem in the shift operator case is shown to be less important compared to the fact the LMI problem is fundamentally ill-conditioned for shorter sampling periods  $h$ . In a study of a numerical example, two different numerical solvers for LMIs exhibit problems with this ill-conditioned behavior.

All these numerical problems are solved by using a delta operator formulation of the LMI. Alternatively, it is possible to use the shift operator model but then apply a simple scaling of the matrix  $P$ . This transformation captures the system scaling mechanism in the delta operator model, but does not avoid the cancellation problem. The relative error then however becomes negligible for all practical choices of sampling periods.

## REFERENCES

- [1] S. Boyd, L. E. Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. SIAM Studies in Applied Mathematics, 1994.
- [2] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [3] S. A. Miller, "Sensitivity of solutions to semidefinite programs," Center for Control Engineering and Computation. University of California at Santa Barbara, Tech. Rep. CCEC-97-0519, <http://www-ccec.ece.ucsb.edu/techrpts/report-97-0519.ps>, 1997.
- [4] J. F. Sturm and S. Zhang, "On sensitivity of central solutions in semidefinite programming," *Mathematical Programming*, vol. 90, no. 2, pp. 205–227, 2001.
- [5] R. H. Middleton and G. C. Goodwin, *Digital Control and Estimation - A Unified Approach*. Prentice Hall, Englewood Cliffs, N. J., 1990.
- [6] E. Collins, W. Haddad, V. Chellaboina, and T. Song, "Robustness analysis in the delta-domain using fixed-structure multipliers," in *Proc. 36th IEEE Conference on Decision and Control*, San Diego, December 1997, pp. 3286–91.
- [7] R. Lin, F. Yang, and Q. Chen, "Design of robust non-fragile  $\mathcal{H}_\infty$  controller based on delta operator theory," *Journal of Control Theory and Applications*, vol. 5, no. 4, pp. 404–408, 2007.
- [8] B. Lennartson, R. Middleton, A. K. Christiansson, and T. McKelvey, "Low order sampled data  $\mathcal{H}_\infty$  control using the delta operator and LMIs," in *Proc. 43rd IEEE Conference on Decision and Control*, Bahamas, December 2004, pp. 4479–4484.
- [9] B. Lennartson, R. H. Middleton, and I. Gustafsson, "Numerical sensitivity of linear matrix inequalities using shift and delta operators," *IEEE Transactions on Automatic Control*, vol. 57, 2012.
- [10] M. Salgado, R. H. Middleton, and G. C. Goodwin, "Connection between continuous and discrete riccati equations with applications to Kalman filtering," *IEE Proceedings, Part D*, vol. 135, no. 1, pp. 28–34, 1988.
- [11] B. Lennartson, R. Middleton, and A. K. Christiansson, "Unified, periodic & sampled data  $\mathcal{H}_\infty$  control using the delta operator," in *Proc. 43rd IEEE Conference on Decision and Control*, Bahamas, December 2004, pp. 2364–2369.
- [12] M. Green and D. Limebeer, *Linear Robust Control*. Prentice Hall, Englewood Cliffs, N. J., 1995.
- [13] P. Gahinet and P. Apkarian, "A linear matrix inequality approach to  $\mathcal{H}_\infty$  control," *International Journal on Robust and Nonlinear Control*, vol. 4, pp. 421–448, 1994.
- [14] H. Lee, J. Park, and Y. Joo, "Further refinement on LMI-based digital redesign: delta-operator approach," *IEEE Transactions on Circuits and Systems Part II: Express Briefs*, vol. 53, no. 6, pp. 473–477, 2006.
- [15] B. N. Datta, *Numerical Linear Algebra and Applications*, 2nd ed. SIAM, 2010.
- [16] R. A. Horn and C. R. Johnson, *Topics in matrix analysis*. Cambridge University Press, 1994.
- [17] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization Methods and Software*, 1999. SeDuMi URL: <http://sedumi.ie.lehigh.edu>.
- [18] M. Kojima, K. Fujisawa, K. Nakata, and M. Yamashita, "SDPA (SemiDefinite Programming Algorithm) User's Manual - Version 6.20." Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tech. Rep., 2005. SDPA URL: <http://sdpa.indsys.chuo-u.ac.jp/sdpa>.
- [19] J. Löfberg, "YALMIP: a toolbox for modeling and optimization in MATLAB," in *Proc. 2004 IEEE International Symposium on Computer Aided Control Systems Design*, Taipei, September 2004, p. 284. YALMIP URL: <http://users.isy.liu.se/johanl/yalmip/>.