The PARMSR algorithm updated every L-customers period is composed of (1), (28), and (29); see, e.g., [18]–[20].

Theorem 4.1: The assertions of Theorem 2.1–2.3 hold in the present setting, if the conditions of the theorems are satisfied, respectively.

Proof: The key step is to verify that $\Sigma_{n=1}^{\infty} ab_n^{1-\delta}\varepsilon_n$ converges, a.s. The observation noise of the algorithm is (4), where f_{n+1} is defined by (28). Along the same lines as in [20], we first introduce the following notations:

$$\mu(n) = \left[\frac{n}{L}\right], \qquad \overline{\theta}_n = \theta_{\mu(n)}, \qquad \tilde{a}_n = a_{\mu(n)} \tag{30}$$

$$\beta_{(n-1)L+j} = \beta_{n,j}, \qquad j = 0, 1, \cdots, L-1 \tag{31}$$

$$\overline{\varepsilon}_n = J_t(T_n, \overline{\theta}_{n-1})\beta_n + J_\theta(T_n, \overline{\theta}_{n-1}) - f(\overline{\theta}_{n-1}) \quad \forall n \geq 1. \tag{32}$$

By (29) and (31) we have

$$\beta_{n+1} = \beta_n I_{[Q_n \geq 1]} + \frac{dx_{n+1}(\overline{\theta}_n)}{d\theta}, \qquad \forall n \geq 0.$$

From (28) and (29) it is derived that

$$\sum_{n=1}^{\infty} a_n^{1-\delta}\varepsilon_{n+1} = \sum_{n=1}^{\infty} a_n^{1-\delta}(f_{n+1} - f(\theta_n)) = \frac{1}{L}\sum_{n=L}^{\infty} \tilde{a}_n^{1-\delta}\overline{\varepsilon}_{n+1}$$

which implies that the almost sure convergence of $\Sigma_{n=1}^{\infty} a_n^{1-\delta}\varepsilon_{n+1}$ is equivalent to the almost sure convergence of $\Sigma_{n=1}^{\infty} \tilde{a}_n^{1-\delta}\overline{\varepsilon}_{n+1}$. The proof of the convergence of $\Sigma_{n=1}^{\infty} \tilde{a}_n^{1-\delta}\overline{\varepsilon}_{n+1}$ works the same way as in Lemma 3.1 and Lemmas 3.3–3.5 if, instead, we replace $\alpha_n, \theta_n, \varepsilon_n, a_n$ by $\beta_n, \overline{s}_n, \overline{\varepsilon}_n, \tilde{a}_n$, respectively. Details are omitted, for the brevity of the paper.

V. Concluding Remark

We have established the convergence rates of the PARMSR algorithm with fixed-length observation period for the $GI/G/1$ queueing systems. Along the same lines of the research, more precise convergence results for the PARMSR algorithms, such as a central limit theorem and a law of the iterated logarithm, could be derived.

References

[1] H. F. Chen, "Convergence rate of stochastic approximation for the case $f(x^0) = 0, f'(x^0) = 0$," Ins. Syst. Sci., The Chinese Academy Sci., Tech. Rep., 1995.

[2] ——, "Stochastic approximation and its new applications," in *Proc. 1994 Hong Kong Int. Wkshp. New Directions Contr. Manufacturing*, pp. 2–12.

[3] H. F. Chen and Y. M. Zhu, *Stochastic Approximation*. Shanghai Press Sci. and Technol., 1996.

[4] E. K. P. Chong and P. J. Ramadge, "Convergence of recursive optimization algorithms using infinitesimal perturbation analysis estimates," *Discrete Event Dynamic Syst.: Theory Appl.*, vol. 1, pp. 339–372, 1992.

[5] ——, "Optimization of queues using an infinitesimal perturbation analysis-based stochastic algorithm with general updates times," *SIAM J. Contr. Optim.*, vol. 31, pp. 698–732, 1993.

[6] ——, "Stochastic optimization of regenerative systems using infinitesimal perturbation analysis," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 1400–1410, July 1994.

[7] M. C. Fu, "Convergence of a stochastic approximation algorithm for the $GI/G/1$ queue using infinitesimal perturbation analysis," *J. Optim. Theory Appl.*, vol. 65, pp. 149–160.

[8] P. Glasserman, *Gradient Estimation via Perturbation Analysis*. Boston, MA: Kluwer, 1991.

[9] A. Gut, *Stopped Random Walks: Limit Theorems and Applications*. New York: Springer-Verlag, 1988.

[10] Y. C. Ho and X. R. Cao, *Perturbation Analysis of Discrete Event Dynamic Systems*. Boston, MA: Kluwer, 1991.

[11] Y. C. Ho, M. A. Eyler, and T. T. Chien, "A gradient technique for general buffer storage design in a serial production lines," *Int. J. Production Res.*, vol. 17, pp. 557–580, 1979.

[12] H. J. Kushner and F. J. Vázquez-Abad, "Stochastic approximation methods for systems of interest over an infinite horizon," *SIAM J. Contr. Optim.*, vol. 34, pp. 712–756, 1996.

[13] P. L'Ecuyer and P. W. Glynn, "Stochastic optimization by simulation: Convergence proofs for the $GI/G/1$ queue in steady-state," *Management Sci.*, vol. 40, pp. 1562–1578, 1994.

[14] P. L'Ecuyer, N. Giroux, and P. W. Glynn, "Stochastic optimization by simulation: Numerical experiments with the $M/M/1$ queue in steady-state," *Management Sci.*, vol. 40, pp. 1245–1261, 1994.

[15] H. Robbins and S. Monro, "A stochastic approximation method," *Ann. Math. Statist.*, vol. 22, pp. 400–407, 1951.

[16] W. F. Stout, *Almost Sure Convergence*. New York: Academic, 1974.

[17] R. Suri, "Perturbation analysis: The state of the art and research issues explained via the $GI/G/1$ queue," *Proc. IEEE*, vol. 77, pp. 114–137, 1989.

[18] R. Suri and Y. T. Leung, "Single run optimization of discrete event simulations—An empirical study using the $M/M/1$ queue," *IIE Trans.*, vol. 21, pp. 35–49, 1989.

[19] R. Suri and M. A. Zazanis, "Perturbation analysis gives strongly consistent sensitivity estimates for the $M/G/1$ queue," *Management Sci.*, vol. 34, pp. 39-64, 1988.

[20] Q. Y. Tang and H. F. Chen, "Convergence of perturbation analysis based optimization algorithm with fixed number of customers period," *Discrete Event Dynamic Systems: Theory and Appl.*, vol. 4, pp. 359–375, 1994.

Improved Upper Bounds for the Mixed Structured Singular Value

Minyue Fu and Nikita E. Barabanov

***Abstract*—In this paper, we take a new look at the mixed structured singular value problem, a problem of finding important applications in robust stability analysis. Several new upper bounds are proposed using a very simple approach which we call the multiplier approach. These new bounds are convex and computable by using linear matrix inequality (LMI) techniques. We show, most importantly, that these upper bounds are actually lower bounds of a well-known upper bound which involves the so-called D-scaling (for complex perturbations) and G-scaling (for real perturbations).**

***Index Terms*—Robust control, robust stability, robustness, structured singular value, uncertain systems.**

I. Introduction

This paper addresses the problem of the mixed structured singular value. The notion of structured singular value, or μ for short, was initially proposed by Doyle [4] for studying the robust stability

Manuscript received December 23, 1994; revised May 29, 1996. This work was supported by the Australian Research Council.

M. Fu is with the Department of Electrical and Computer Engineering, The University of Newcastle, Newcastle, NSW 2308, Australia.

N. E. Barabanov is with the Department of Software Engineering, Electrical Engineering University, St. Petersburg, Russia.

Publisher Item Identifier S 0018-9286(97)06600-2.

of linear systems which are subject to certain structured complex perturbations. It has also been extended to the case where there exist both real parameters and complex uncertainty, i.e., the so-called mixed perturbations; see Fan *et al.* [5].

The computational issue of μ has been studied in numerous papers; see Young *et al.* [16] for a summary. There is also a Matlab Toolbox available; see Balas *et al.* [1]. The best upper bound of μ known so far is given by Fan *et al.* [5] for a general μ problem where both real and complex perturbations are allowed. This bound is generalized from an earlier result of Doyle [4].

In this paper, we propose several new upper bounds for the structured singular value. The first new upper bound is derived by using the following trivial fact.

Lemma I.1: A matrix family $\mathcal{A}$ is nonsingular if there exists another matrix C (multiplier) such that the Hermitian part of the product AC is negative-definite for all $A \in \mathcal{A}$.

An interesting point regarding this upper bound is that it is convex. This upper bound holds an important conceptual value because it is extremely easy to derive and serves as a lower bound for other upper bounds. Because this upper bound is not easy to compute when there exist complex perturbations, we apply the well-known $\mathcal{S}$-procedure on it and derive a second new upper bound which is looser than the first one but convex and computable by using linear matrix inequality (LMI) techniques. The computation of this bound is numerically efficient when the number of real perturbation blocks is not large. When the number of real perturbation blocks is large, we derive another new upper bound which is even looser but more efficient to compute. This upper bound is still convex and computable by using the LMI method. *But most surprisingly, we find that all these new upper bounds are actually lower bounds of the upper bound of [5] mentioned above*.

The rest of this paper is organized as follows: Section II reviews the mixed structured singular value problem and the upper bound in [5]. Section III provides three new upper bounds. The relationships among these bounds and the upper bound of [5] are analyzed in Section IV. Section V compares the multiplier approach with an alternative approach for μ analysis. Some concluding remarks are given in Section VI.

II. Review of the Structured Singular Value

The notation needed for defining the mixed structured singular value is standard, and we simply duplicate it from [5].

Given $M \in \mathcal{C}^{n\times n}$ and nonnegative integers m_r, m_c, and m_C, with

$$m = m_r + m_c + m_C \le n \tag{1}$$

a block structure $\mathcal{K}$ is defined to be an m-tuple of positive integers

$$\mathcal{K} = [k_1, \cdots, k_{m_r}, k_{m_r+1}, \cdots, k_{m_r+m_c}, k_{m_r+m_c+1}, \cdots, k_m] \tag{2}$$

subject to $\Sigma_{i=1}^m k_i = n$. Also define the family of block-diagonal $n \times n$ matrices

$$\begin{aligned}\mathcal{X} = \{&\text{block diag}(\delta_1^r I_{k_1}, \cdots, \delta_{m_r}^r I_{k_{m_r}}, \delta_1^c I_{k_{m_r+1}}, \cdots,\\ &\delta_{m_c}^c I_{k_{m_r+m_c}}, \Delta_1^C, \cdots, \Delta_{m_C}^C)\colon \delta_i^r \in \mathcal{R}, \delta_i^c \in \mathcal{C},\\ &\Delta_i^C \in \mathcal{C}^{(k_{m_r+m_c+i})\times(k_{m_r+m_c+i})}\}.\end{aligned} \tag{3}$$

In the above, I_k denotes the $k \times k$ identity matrix.

Definition II.1 [4], [5]: The structured singular value $\mu_\mathcal{K}(M)$ of a complex $n \times n$ matrix M with respect to the block-structure $\mathcal{K}$ is defined to be zero if $I_n - \Delta M$ is nonsingular for all $\Delta \in \mathcal{X}$, and

$$\mu_\mathcal{K}(M) = (\min_{\Delta\in\mathcal{X}}\{\overline{\sigma}(\Delta)\colon \det(I_n - \Delta M) = 0\})^{-1} \tag{4}$$

otherwise, where $\overline{\sigma}(\ )$ denotes the largest singular value.

The exact computation of $\mu_\mathcal{K}(M)$ is very difficult. In fact, its computational complexity is known to be NP-hard; see, e.g., Poljak and Rohn [12]. In practice, an upper bound and a lower bound are computed to approximate the exact value. The upper bound is very important because it serves as a guaranteed margin for robust nonsingularity. In contrast, the lower bound is often used for checking the tightness of the upper bound. A popularly used upper bound for the structured singular value is given by Fan *et al.* [5]. To describe this upper bound, we introduce more notation:

$$\begin{aligned}\mathcal{D} = \{&\text{block diag}(D_1, \cdots, D_{m_r+m_c}, d_1 I_{k_{m_r+m_c+1}}, \cdots,\\ &d_{m_c+m_C} I_{k_m})\colon 0 < D_i = D_i^H \in \mathcal{C}^{k_i\times k_i}, d_i > 0\}\end{aligned} \tag{5}$$

$$\begin{aligned}\mathcal{G} = \{&\text{block diag}(G_1, \cdots, G_{m_r}, 0_{k_{m_r+1}}, \cdots, 0_{k_m})\colon\\ &G_i = G_i^H \in \mathcal{C}^{k_i\times k_i}\}\end{aligned} \tag{6}$$

where 0_k denotes a $k \times k$ zero matrix. For any $a \in \mathcal{R}$, define

$$\Phi_a(D, G) = M^H DM + j(GM - M^H G) - a^2 D. \tag{7}$$

Lemma II:2 [5]: Given M and $\mathcal{K}$, we have $\mu_\mathcal{K}(M) \le \nu_\mathcal{K}(M)$, where

$$\nu_\mathcal{K}(M) = \inf_{a>0}\{a\colon \exists D \in \mathcal{D}, G \in \mathcal{G}\colon \Phi_a(D, G) < 0\}. \tag{8}$$

Remark II.3: It is known that the gap between $\nu_\mathcal{K}(M)$ and $\mu_\mathcal{K}(M)$ can be very large when there are real perturbation blocks [16]. It is recently proved by Meinsma *et al.* [14] that $\nu_\mathcal{K}(M) = \mu_\mathcal{K}(M)$ if $2(m_r + m_c) + m_C \le 3$. For each case where $2(m_r + m_c) + m_C > 3$, examples exist to give a gap between $\mu_\mathcal{K}(M)$ and $\nu_\mathcal{K}(M)$.

III. New Upper Bounds

In this section, the three new upper bounds for $\mu_\mathcal{K}(M)$ as discussed in Section I are provided.

A. The First New Upper Bound

The first new upper bound is proposed by using the trivial fact in Lemma I.1. When applied to the μ problem, Lemma I.1 says that $\mu_\mathcal{K}(M) \le a$ $(a > 0)$ if there exists a multiplier $C \in \mathcal{C}^{n\times n}$ such that

$$C(I - \Delta M) + (I - \Delta M)^H C^H < 0 \tag{9}$$

for all $\Delta \in \mathcal{X}$ with $\overline{\sigma}(\Delta) \le a^{-1}$.

When M has the following special structure:

$$M = AB^H \tag{10}$$

where $A, B \in \mathcal{C}^{n\times q}$ and $q \le n$, the following equation becomes useful:

$$\det(I_n - \Delta M) = \det(I_n - M\Delta) = \det(I_q - B^H \Delta A). \tag{11}$$

An important special case of the above is $A = M$ and $B^H = I_n$.

With the rewriting of M in (10), the sufficient condition for μ above reduces to finding the multiplier $C \in \mathcal{C}^{q\times q}$ and checking

$$(I_q - B^H \Delta A)^H C + C^H (I_q - B^H \Delta A) < 0 \tag{12}$$

for all $\Delta \in \mathcal{X}$ with $\overline{\sigma}(\Delta) \le a^{-1}$. Redenoting Δ^H by Δ, the above is equivalent to

$$C^H (I_q - A^H \Delta B)^H + (I_q - A^H \Delta B) C < 0 \tag{13}$$

for all $\Delta \in \mathcal{X}$ with $\overline{\sigma}(\Delta) \le a^{-1}$.

Remark III.1: Condition (12) or (13) has an obvious advantage over (9) because the size of the multiplier in (12) or (13) is only $q \times q$ rather than $n \times n$.

Based on the analysis above, we propose a simple new upper bound as follows: given a matrix $M \in \mathcal{C}^{n \times n}$ of the form (10) and an associated block-structure $\mathcal{K}$, define

$$\hat{\mu}_{\mathcal{K}}(M) = \inf_{a>0} \{a: \exists\, C \in \mathcal{C}^{q \times q}: E(C, \Delta) > 0, \quad \forall \Delta \in \mathcal{X}, \overline{\sigma}(\Delta) \leq a^{-1}\} \tag{14}$$

where

$$E(C, \Delta) = C^H (I - A^H \Delta B)^H + (I - A^H \Delta B) C. \tag{15}$$

Obviously, we have

$$\mu_{\mathcal{K}}(M) \leq \hat{\mu}_{\mathcal{K}}(M). \tag{16}$$

Also note that the constraints on δ_i in (14) can be simply replaced with $\delta_i = \pm a^{-1}$.

B. The Second New Upper Bound

The computation of the first new upper bound is inconvenient in cases where there are complex perturbation blocks. To simplify the computation, we obtain a looser upper bound by applying the well-known $\mathcal{S}$-procedure [15] on the complex blocks (i.e., replacing them with D-scaling). To this end, we denote by Δ_R and Δ_C the real and complex sub-blocks of Δ, respectively

$$\Delta_R = \text{block diag}(\delta_1^r I_{k_1}, \cdots, \delta_{m_r}^r I_{k_{m_r}}) \tag{17}$$

$$\Delta_C = \text{block diag}(\delta_1^c I_{k_{m_r+1}}, \cdots, \delta_{m_c}^c I_{k_{m_r+m_c}}, \Delta_1^C, \cdots, \Delta_{m_C}^C) \tag{18}$$

$\mathcal{K}_R, \mathcal{K}_C, \mathcal{X}_R, \mathcal{X}_C, \mathcal{D}_R, \mathcal{D}_C, D_R, D_C, G_R$ are similarly defined (note that $G_C = 0$). We also rewrite (10) as follows:

$$M = AB^H = \begin{bmatrix} A_R \\ A_C \end{bmatrix} [B_R^H \quad B_C^H]. \tag{19}$$

Then, $E(C, \Delta)$ can be rewritten as

$$E(C, \Delta) = E(C, \Delta_R) - C^H B_C^H \Delta_C^H A_C - A_C^H \Delta_C B_C C \tag{20}$$

where

$$E(C, \Delta_R) = C^H (I - A_R^H \Delta_R B_R)^H + (I - A_R^H \Delta_R B_R) C. \tag{21}$$

Applying the $\mathcal{S}$-procedure on the complex blocks in (20), we obtain the following result.

Theorem III.2: Let $E(C, \Delta)$ be given in (20) and $a > 0$. Then, $E(C, \Delta) < 0$ for all $\Delta \in \mathcal{X}$ with $\overline{\sigma}(\Delta) \leq a^{-1}$ if there exists scalars $d_i, i = 1, \cdots, m_c + m_C$ to form $D_C \in \mathcal{D}_{\mathcal{K}_C}$ such that the following 2^{m_r} inequalities hold:

$$F(C, D_C, \Delta_R) = \begin{bmatrix} E(C, \Delta_R) + A_C^H D_C A_C & C^H B_C^H \\ B_C C & -a^2 D_C \end{bmatrix} < 0 \quad \forall \delta_i^r = \pm a^{-1}, i = 1, \cdots, m_r. \tag{22}$$

Proof: Suppose (22) holds for some D_C. Then, $F(C, D_C, \Delta_R) < 0$ for all $\Delta_R \in \mathcal{X}_R$ with $\overline{\sigma}(\Delta_R) \leq a^{-1}$ because $F(C, D_C, \Delta_R)$ is linear in Δ_R. Thus, for any $\Delta \in \mathcal{X}$ with $\overline{\sigma}(\Delta) \leq a^{-1}$

$$E(C, \Delta) \leq E(C, \Delta_R) + C^H B_C^H D_C^{-1/2} \Delta_C^H \Delta_C D_C^{-1/2} B_C C < 0.$$

The last step above follows from the Schur complement of (22). ■

Based on Theorem III.2, we define our second new upper bound as follows:

$$\overline{\mu}_{\mathcal{K}}(M) = \inf_{a>0} \{a: \exists\, C \in \mathcal{C}^{q \times q}, D_C \in \mathcal{D}_C: F(C, D_C, \Delta_R) < 0, \forall \delta_i^r = \pm a^{-1}, i = 1, \cdots, m_r\}. \tag{23}$$

This new upper bound can be computed using LMI methods. But the direct use of (22) will result in 2^{m_r} LMI's of size much larger than $q \times q$. Alternatively, we can significantly reduce the dimensions by introducing an additional variable matrix. This point is made precise in the following result.

Theorem III.3: The set of inequalities (22) holds if and only if there exists $K = K^H \in \mathcal{C}^{q \times q}$ such that the following $2^{m_r} + 1$ LMI's hold:

$$F_1(C, K, \Delta_R) = E(C, \Delta_R) + K < 0, \quad \forall \delta_i^r = \pm a^{-1}, i = 1, \cdots, m_r \tag{24}$$

$$F_2(C, K, D_C) = \begin{bmatrix} -K + A_C^H D_C A_C & C^H B_C^H \\ B_C C & -a^2 D_C \end{bmatrix} \leq 0. \tag{25}$$

Consequently, $\overline{\mu}_{\mathcal{K}}(M)$ can be alternatively expressed by

$$\overline{\mu}_{\mathcal{K}}(M) = \inf_{a>0} \{a: \exists C, K = K^H \in \mathcal{C}^{q \times q}, D_C \in \mathcal{D}_C: F_2(C, K, D_C) \leq 0, F_1(C, K, \Delta_R) < 0 \quad \forall \delta_i^r = \pm a^{-1}, i = 1, \cdots, m_r\}. \tag{26}$$

Proof: We first note that the inequality in (25) can be alternatively expressed by

$$K - A_C^H D_C A_C - a^{-2} C^H B_C^H D_C^{-1} B_C C \geq 0. \tag{27}$$

Suppose the inequalities in (22) hold for some D_C. Then, we simply take

$$K = A_C^H D_C A_C + a^{-2} C^H B_C^H D_C^{-1} B_C C. \tag{28}$$

Obviously, both (24) and (27) hold.

Conversely, suppose (24) and (27) hold for some K. Then, (22) follows trivially too.

Remark III.4: Obviously, $F_1(C, K, \Delta_R)$ and $F_2(C, K, D_C)$ are jointly linear in C, K, and D_C. The collective dimension of the LMI's in (24) and (25) is substantially less than that obtained by using (22). The tradeoff is the additional variable matrix. Fortunately, this tradeoff is not very costly for many robust stability problems because the matrix is $q \times q$ only and it is Hermitian.

C. The Third New Upper Bound

We now derive another new upper bound which is even looser than $\overline{\mu}_{\mathcal{K}}(M)$. However, this new upper bounds serves two purposes: 1) it bridges the gap between $\overline{\mu}_{\mathcal{K}}(M)$ and $\nu_{\mathcal{K}}(M)$, i.e., it is in between the two, and 2) it is numerically more efficient than $\overline{\mu}_{\mathcal{K}}(M)$ when the number of real perturbation blocks becomes large.

Our first step is to partition Δ_R into two parts:

$$\Delta_R = \text{diag}\{\Delta_{R_1}, \Delta_{R_2}\} \tag{29}$$

where Δ_{R_1} consists of the first m_1 repeated real blocks, while Δ_{R_2} are the remaining $m_r - m_1$ ones. Also, $A_R, B_R, D_R, \mathcal{K}_R$, and $\mathcal{X}_R$ are partitioned similarly. In particular

$$E(C, \Delta) = E(C, \Delta_{R_1}) - C^H B_{R_2}^H \Delta_{R_2}^H A_{R_2} - A_{R_2}^H \Delta_{R_2} B_{R_2} C - C^H B_C^H \Delta_C^H A_C - A_C^H \Delta_C B_C C \tag{30}$$

where

$$E(C, \Delta_{R_1}) = C^H (I - A_{R_1}^H \Delta_{R_1} B_{R_1})^H + (I - A_{R_1}^H \Delta_{R_1} B_{R_1}) C. \quad (31)$$

The motivation for the partition above stems from the fact that the number of LMI's involved in computing $\overline{\mu}_{\mathcal{K}}(M)$ is 2^{m_r}, which is an exponential number. With this partition, we will reduce this number to 2^{m_1} by converting the Δ_{R_2} block into additional scaling variables D_{R_2} and G_{R_2} [which are a part of D_R and G for $\nu_{\mathcal{K}}(M)$].

With the partition above, we obtain the following result.

Theorem III.5: Let $F(C, D_C, \Delta_R)$ be given in (22) and $a > 0$. Then, $F(C, D_C, \Delta_R) < 0$ for all $\Delta_R \in \mathcal{X}_R$ with $\overline{\sigma}(\Delta) \le a^{-1}$ if there exist $D_{R_2} \in \mathcal{D}_{R_2}$ and $G_{R_2} \in \mathcal{G}_{R_2}$ such that the following LMI's hold:

$$\begin{aligned} L(C, D_C, \Delta_{R_1}, D_{R_2}, G_{R_2}) &= \begin{bmatrix} L_{11} & C^H B_C^H & C^H B_{R_2}^H + jA_{R_2}^H G_{R_2} \\ B_C C & -a^2 D_C & 0 \\ B_{R_2} C - jG_{R_2} A_{R_2} & 0 & -a^2 D_{R_2} \end{bmatrix} \\ &< 0, \forall \delta_i^r = \pm a^{-1}, i = 1, \cdots, m_1 \end{aligned} \quad (32)$$

where

$$L_{11} = E(C, \Delta_{R_1}) + A_C^H D_C A_C + A_{R_2}^H D_{R_2} A_{R_2}.$$

Proof: Suppose $L(C, D_C, \Delta_{R_1}, D_{R_2}, G_{R_2}) < 0$ for all $\Delta_{R_1} \in \mathcal{X}_{R_1}$ with $\overline{\sigma}(\Delta_{R_1}) \le a^{-1}$. Define

$$Y = \begin{bmatrix} I_q & 0 \\ 0 & I_q \\ -\Delta_{R_2} A_{R_2} & 0 \end{bmatrix}. \quad (33)$$

It follows that

$$\begin{aligned} Y^H L(C, D_C, \Delta_{R_1}, D_{R_2}, G_{R_2}) Y &= F(C, D_C, \Delta_R) \\ &\quad + \mathrm{diag}\{A_{R_2}^H D_{R_2}^{1/2} (I - a^2 \Delta_{R_2}^H \Delta_{R_2}) D_{R_2}^{1/2} A_{R_2}, 0\}. \end{aligned} \quad (34)$$

Therefore, $F(C, D_C, \Delta_R) < 0$ for all $\Delta_R \in \mathcal{K}_R$ with $\overline{\sigma}(\Delta_R) \le a^{-1}$.

Using Theorem III.5, we can define the third new upper bound as follows:

$$\begin{aligned} \tilde{\mu}_{\mathcal{K}}(M) = \inf_{a>0} \{a\colon & \exists C \in \mathcal{C}^{q \times q}, D_{R_2} \in \mathcal{D}_{R_2}, G_{R_2} \in \mathcal{G}_{R_2}, \\ & D_C \in \mathcal{D}_C\colon L(C, D_C, \Delta_{R_1}, D_{R_2}, G_{R_2}) < 0, \\ & \forall \delta_i^r = \pm a^{-1}, i = 1, \cdots, m_1\} \end{aligned} \quad (35)$$

IV. Relationships Among the Upper Bounds

The sole purpose of this section is to prove the following result.

Theorem IV.1: Given M and $\mathcal{K}$, we have

$$\mu_{\mathcal{K}}(M) \le \hat{\mu}_{\mathcal{K}}(M) \le \overline{\mu}_{\mathcal{K}}(M) \le \tilde{\mu}_{\mathcal{K}}(M) \le \nu_{\mathcal{K}}(M). \quad (36)$$

The proof is based on the following result (see, e.g., Gahinet and Apkarian [8]).

Lemma IV.2: Given $U, V \in \mathcal{C}^{q \times n}$ and $Q = Q^H \in \mathcal{C}^{n \times n}$, there exists $C \in \mathcal{C}^{q \times q}$ such that

$$Q + U^H C V + V^H C^H U < 0 \quad (37)$$

if and only if both of the following conditions are satisfied:

$$U_\perp^H Q U_\perp < 0, \qquad V_\perp^H Q V_\perp < 0 \quad (38)$$

where $U_\perp$ and $V_\perp$ are any "null matrices" of U and V (i.e., they form the bases of the null spaces of U and V), respectively.

Proof of Theorem IV.1: The first inequality in (36) has been explained before [see (16)]. The second and third inequalities follow trivially from Theorems III.2 and III.5. To show the fourth inequality, we let M be in the form of (19) and need to prove the following: suppose there exists $D \in \mathcal{D}$ and $G \in \mathcal{G}$ such that $\Phi_a(D, G) < 0$ [recall the definition of $\Phi_a(D, G)$ in (7)]. Then, there exist $C \in \mathcal{C}^{q \times q}$ such that $L(C, D_C, \Delta_{R_1}, D_{R_2}, G_{R_2}) < 0$ for all Δ_{R_1} with $\overline{\sigma}(\Delta_{R_1}) \le a^{-1}$, or equivalently, with $\delta_i^r = \pm a^{-1}, i = 1, \cdots, m_1$. Two steps will be involved to achieve this purpose.

Step 1) We apply Lemma IV.2 to show that $\Phi_a(D, G) < 0$ is equivalent to the following condition: there exists some $C \in \mathcal{C}^{q \times q}$ such that

$$W(C, D, G) = \begin{bmatrix} C + C^H + A^H D A & C^H B^H + jA^H G \\ BC - jGA & -a^2 D \end{bmatrix} < 0. \quad (39)$$

To see this, we rewrite (39) as follows:

$$\begin{aligned} W(C, D, G) = \begin{bmatrix} A^H D A & -jA^H G \\ jGA & -a^2 D \end{bmatrix} + \begin{bmatrix} I \\ B \end{bmatrix} C [I \quad 0] \\ + \begin{bmatrix} I \\ 0 \end{bmatrix} C^H [I \quad B^H] < 0 \end{aligned} \quad (40)$$

and observe that the "null matrices" of $[I \quad 0]$ and $[I \quad B^H]$ are given by

$$\begin{bmatrix} 0 \\ I \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -B^H \\ I \end{bmatrix} \quad (41)$$

respectively. From Lemma IV.2, (26) holds if and only if $-a^2 D < 0$ and

$$BA^H DAB^H + j(BA^H G - GAB^H) - a^2 D < 0. \quad (42)$$

Note that the first condition is redundant and (42) is the same as $\Phi_a(D, G) < 0$.

Step 2) Partition $W(C, D, G)$ as follows:

$$W(C, D, G) = \begin{bmatrix} W_{11} & W_{12} & W_{13} & C^H B_C^H \\ W_{12}^H & -a^2 D_{R_1} & 0 & 0 \\ W_{13}^H & 0 & -a^2 D_{R_2} & 0 \\ B_C C & 0 & 0 & -a^2 D_C \end{bmatrix}$$

where

$$\begin{aligned} W_{11} &= C + C^H + A_{R_1}^H D_{R_1} A_{R_1} \\ &\quad + A_{R_2}^H D_{R_2} A_{R_2} + A_C^H D_C A_C \\ W_{12} &= C^H B_{R_1}^H + jA_{R_1}^H G_{R_1} \\ W_{13} &= C^H B_{R_2}^H + jA_{R_2}^H G_{R_2}. \end{aligned}$$

By rearranging its second, third, and fourth row and column blocks, $W(C, D, G)$ above becomes

$$\hat{W}(C, D, G) = \begin{bmatrix} W_{11} & C^H B_C^H & W_{12} & W_{13} \\ B_C C & -a^2 D_C & 0 & 0 \\ W_{12}^H & 0 & -a^2 D_{R_2} & 0 \\ W_{13}^H & 0 & 0 & -a^2 D_{R_1} \end{bmatrix}.$$

Obviously

$$W(C, D, G) < 0 \Leftrightarrow \hat{W}(C, D, G) < 0. \quad (43)$$

Define

$$Z = \begin{bmatrix} I_q & 0 & 0 \\ 0 & I_q & 0 \\ 0 & 0 & I_q \\ -\Delta_{R_1} A_{R_1} & 0 & 0 \end{bmatrix}. \quad (44)$$

It is a bit tedious but straightforward to verify the following equality:

$$\begin{aligned} Z^H\hat{W}(C,D,G)Z &= L(C,D_C,\Delta_{R_1},D_{R_2},G_{R_2}) \\ &\quad + \mathrm{diag}\{A_{R_1}^H D_{R_1}^{1/2}(I-a^2\Delta_{R_1}^H\Delta_{R_1})D_{R_1}^{1/2}A_{R_1},0,0\}. \end{aligned} \quad (45)$$

Hence

$$\Phi_a(D,G)<0 \Leftrightarrow \hat{W}(C,D,G)<0$$

implies that

$$L(C,D_C,\Delta_{R_1},D_{R_2},G_{R_2})<0$$

for all $\Delta_{R_1}\in\mathcal{X}_{R_1}$ with $\overline{\sigma}(\Delta_{R_1})\le a^{-1}$.

Remark IV.3: A simple example is given in [13] to show that the new upper bounds are indeed strictly less than $\nu_\mathcal{K}(M)$. For this example

$$\mathcal{X}=\begin{pmatrix}\mathcal{R} & \\ & \mathcal{R}\end{pmatrix};\qquad M=\begin{pmatrix}0 & 1\\ j & 0\end{pmatrix}.$$

It is shown [13] that $0=\mu_K(M)<\hat{\mu}_\mathcal{K}(M)=0.707<\nu_\mathcal{K}(M)=1$.

V. Comparison with an Alternative Approach

During the revision of this paper (see [7] for previous version), Iwasaki, Hara, and an anonymous reviewer brought to our attention an alternative approach for computing an upper bound for μ. This approach is based on the following result.

Lemma V.1 [11]: Given M and $\mathcal{K},\mu_\mathcal{K}(M)<1$ if and only if there exist matrices $R=R^H$, $Q=Q^H$, and F with appropriate dimensions such that

$$R+M^HF+F^HM+M^HQM<0 \quad (46)$$

$$\Delta^HR\Delta+F\Delta+\Delta^HF^H+Q>0 \quad (47)$$

for all $\Delta\in\mathcal{K}$ with $\overline{\sigma}(\Delta)\le 1$.

The conditions in Lemma V.1 are not suitable for computing μ in general because (47) is not convex in Δ. However, the convexity is guaranteed under the relaxation that $R\le 0$. In this case, the real parameters δ_i in (47) can be replaced with ±1. In particular, it is easily shown in [11] that the $D-G$ scaling technique of [5] gives a special choice of $R\le 0, F,$ and Q.

An interesting point by the aforementioned anonymous reviewer, which we appreciate, is that if a multiplier C exists for (9) (taking $a=1$), then $R\le 0, F,$ and Q also exist for (46) and (47). Hence, it seems that the multiplier approach can be improved further. However, the result below shows that this is not the case.

Theorem V.2: There exist $R=R^H\le 0$, $Q=Q^H$, and F such that (46) and (47) hold if and only if there exists C such that (9) holds.

Proof: To see the "if" part, we note a result in [13] which states that a multiplier C exists for (9) to hold (with $a=1$) if and only if there exists a different C such that

$$C(I-M\Delta)+(I-M\Delta)^HC^H<0,\forall\Delta\in\mathcal{X},\overline{\sigma}(\Delta)\le 1. \quad (48)$$

Now, take $Q=-C-C^H, F=CM,$ and sufficiently small $R<0$, and we obtain (46) and (47).

To see the "only if" part, we pre- and post-multiply (47) by M^H and M, respectively, to obtain

$$M^HQM+M^HF\Delta M+M^H\Delta^HF^HM+M^H\Delta^HR\Delta M\ge 0.$$

Canceling the Q term above using (46), we obtain

$$\begin{aligned} &M^HF(I-\Delta M)+(I-M^H\Delta^H)F^HM \\ &\quad +R-M^H\Delta^HR\Delta M<0. \end{aligned}$$

Reorganizing the terms above yields

$$\begin{aligned} &(M^HF+R)(I-\Delta M)+(I-M^H\Delta^H)(^HM+R) \\ &\quad -(I-\Delta M)^HR(I-\Delta M)<0. \end{aligned}$$

Since $R\le 0$, we obtain (9) with $C=(M^HF+R)$.

Remark V.3: The implication of the result above is that the multiplier approach is advantageous because it involves a much smaller number of variables (C versus Q,R,F).

VI. Concluding Remarks

In this paper, we have provided several new upper bounds for the mixed structured singular value. Despite of the fact that these upper bounds are derived based on a very simple multiplier approach, we have shown that they are indeed lower bounds of an upper bound given in [5]. The first upper bound serves a conceptual value as it is very easy to derive. The second upper bound is computable via LMI techniques and is suitable when the number of real perturbations m_r is not large. When m_r is large, the third upper bound needs to be used to avoid exponential growth in computation. The relaxation technique $\mathcal{S}$-procedure is the key in obtaining the second and third upper bounds. When the $\mathcal{S}$-procedure is applied to all real and complex perturbations, the upper bound in [5] follows.

We must stress that these new upper bounds still appear to be very coarse. In other words, the news of this paper is somewhat disappointing as it points out that our knowledge about μ is still very primitive despite of years of research. It is a challenging task to find better upper bounds which are also convex and efficiently computable. The computational issue of μ remains wide open. On the other hand, a recent negative result [6] shows that it is unlikely to have a polynomial algorithm which gives an upper bound for μ with relative error grows at a rate at most exponential (in fact it is impossible to have such a guarantee unless all NP problems are solvable in polynomial times).

The type of uncertainty analyzed by the μ framework is somewhat restrictive, i.e., it must be of the structure (3) and ℓ_∞ norm-bounded. It is not difficult to see that the multiplier approach can deal with a much larger class of uncertainty, namely, it allows any convex or even nonconvex set of uncertainty. This observation provides a simple connection between the μ theory and the absolute stability theory, where the sector-bounded uncertainty is typically used. The connection lies in the use of a multiplier. The idea of using multipliers for robustness analysis has been used by a number of authors; see, e.g., How and Hall [9] and Sparks and Beinstein [10] where some generalized Popov multipliers are used. We note, however, that a multiplier of this kind is a special type of $D-G$ scaling when the uncertainty is norm-bounded.

We also point out that when the multiplier approach is used to study the robust stability of a family of matrices, it has a simple link to quadratic stability. Namely, if the multiplier is restricted to a constant, positive-definite and Hermitian matrix, we face the problem of quadratic stability. More precisely, given a family of connected matrices $\mathcal{A}\subset\mathcal{C}^{q\times q}$, necessary and sufficient conditions for robust Hurwitz stability are: 1) there exists some $A_0\in\mathcal{A}$ which is Hurwitz stable, and 2) for every $\omega\in\mathcal{R}$, the matrix family $\mathcal{A}_\omega=\{A-j\omega I\colon A\in\mathcal{A}\}$ is robustly nonsingular.

Using the multiplier approach, a sufficient condition for $\mathcal{A}_\omega$ to be robustly nonsingular is the existence of a multiplier $C_\omega\in\mathcal{C}^{q\times q}$ such that

$$(A-j\omega I)^HC_\omega^H+C_\omega(A-j\omega I)<0,\qquad\forall A\in\mathcal{A}. \quad (49)$$

Taking $C_\omega=C=C^H>0$, the above reduces to quadratic stability

$$A^HC+CA<0,\qquad\forall A\in\mathcal{A}. \quad (50)$$

An example of using the multiplier approach to quadratic stability analysis can be found in Boyd and Yang [3] and Boyd *et al.* [2].

A follow-up paper [13] offers several other interesting properties of the multiplier approach. Namely, an equivalence among several multiplier schemes is established. The computation of the new upper bounds is formally formulated as an generalized eigenvalue problem which can be solved using LMI techniques. The continuity issue of the upper bounds is also studied.

References

[1] G. Balas, J. C. Doyle, K. Glover, A. Packard, and R. Smith, *The μ Analysis and Synthesis Toolbox*. Natick, MA: MathWorks and MUSYN, 1991.

[2] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Philadelphia, PA: SIAM, 1994.

[3] S. Boyd and Q. Yang, "Structured and simultaneous Lyapunov functions for system stability problems," *Int. J. Contr.*, vol. 49, pp. 2215–2240, 1989.

[4] J. C. Doyle, "Analysis of feedback systems with structured uncertainties," *Proc. Inst. Elec. Eng.*, vol. 129, pt. D, pp. 240–250, 1982.

[5] M. K. H. Fan, A. L. Tits, and J. C. Doyle, "Robustness in the presence of mixed parametric uncertainty and unmodeled dynamics," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 25–38, Jan. 1991.

[6] M. Fu, "The real structured singular value is hardly approximable," *IEEE Trans. Automat. Contr.*, to be published.

[7] M. Fu and N. E. Barabanov, "A simple upper bound of the structured singular value," in *Proc. European Contr. Conf.*, Rome, Italy, Sept. 1995, pp. 1027–1032.

[8] P. Gahinet and P. Apkarian, "A linear matrix inequality approach to H_∞ control," *Int. J. Robust Nonlinear Contr.*, Oct. 1992.

[9] J. P. How and S. R. Hall, "Connections between the Popov stability criterion and bounds for real parameter uncertainty," in *Proc. Amer. Contr. Conf.*, June 1995, pp. 1084–1089.

[10] A. G. Sparks and D. S. Bernstein, "The scaled Popov criterion and bounds for the real structured singular value," in *Proc. IEEE Conf. Decision Contr.*, Dec. 1994, pp. 2998–3002.

[11] T. Iwasaki, S. Hara, and T. Asai, "Well-posedness theorem: A Classification of LMI/BMI-reducible robust control problems," *Int. Symp. Intelligent Robotic Syst.*, Bangalore, India, Nov. 1995.

[12] S. Poljak and J. Rohn, "Checking robust nonsingularity is NP-hard," *Math. Contr., Signals, Syst.*, vol. 6, pp. 1–9, 1993.

[13] G. Meinsma, Y. Shrivastava, and M. Fu, "Some properties of an upper bound of μ," *IEEE Trans. Automat. Contr.*, vol. 41, pp. 1326–1330, Sept. 1996.

[14] ——, "A dual formulation of mixed μ and the losslessness of (D, G)-scaling," *IEEE Trans. Automat. Contr.*, to be published.

[15] V. A. Yakubovich, "$\mathcal{S}$-procedure in nonlinear control theory," *Vestnik Leninggradskogo Universiteta, Ser. Matematika*, pp. 62–77, 1971.

[16] P. M. Young, M. P. Newlin, and J. C. Doyle, "Practical computation of the mixed μ problem," in *Proc. Amer. Contr. Conf.*, 1992, pp. 2190–2194.

A Numerical Method in Optimal Production and Setup Scheduling of Stochastic Manufacturing Systems

H. Yan and Q. Zhang

***Abstract*— In this paper, we consider optimal production and setup scheduling in a failure-prone manufacturing system consisting of a single machine. The system can produce several types of products, but at any given time it can only produce one type of product. A setup is required if production is to be switched from one type of product to another. The decision variables are a sequence of setups and a production plan. The objective of the problem is to minimize the costs of setup, production, and surplus. An approximate optimality condition is given together with a computational algorithm for solving the optimal control problem.**

***Index Terms*—HJB equation, machine setup, numerical method, production planning, scheduling.**

I. Introduction

There is substantial literature on continuous-time Markov production planning. Such models are considered by Kimemia and Gershwin [10], Akella and Kumar [1], Boukas and Haurie [3], Haurie and Van Delft [8], and Sharifnia [16], among others. A crucial assumption that has been in these papers is that the machines are completely flexible and thus do not require any setup for switching the production from one product to another. Ideally, the assumption amounts to the possibility of simultaneous production of different products. An important class of manufacturing systems consists, however, of systems that have machines which involve setup costs and/or setup times, when switching from production of one product to that of another. Such systems have been considered by Gershwin [7], Sharifnia *et al.* [17], Connolly *et al.* [4], Hu and Caramanis [9], Bai and Elhafsi [2], and Srivatsan and Gershwin [19]. They have examined various possible heuristic policies and have carried out numerical computations and simulations. However, they have not studied the policies from the viewpoint of optimality or asymptotic optimality. Sethi and Zhang [15] formulated a continuous-time production and setup scheduling model. Using the theory of viscosity solutions of Hamilton–Jacobi–Bellman (HJB) equations, they were able to establish the optimality conditions. However, a closed form optimal solution in these cases is an impossible task to accomplish. In order to be able to use the optimality theory on real time production control, numerical methods for the model developed in [15] seem to be the only feasible approach.

It is the purpose of this paper to study numerical algorithms for solving the underlying problem. In this paper, we consider a manufacturing system consisting of a single failure-prone machine capable of producing a number of different products. A setup (with setup time or cost or both) is required if production is to be switched from one type of product to another. The problem is to find a sequence of setups and a production plan to minimize the total cost of setups, production, and surplus. Since a closed form solution to the problem is not available, one has to resort to a numerical approach

Manuscript received April 16, 1996. This work was supported in part by The Chinese University of Hong Kong under Direct Grant 220500850, ONR Grant N00014-96-1-0263, and a University of Georgia Faculty Research Grant.

H. Yan is with the Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Satin, Hong Kong.

Q. Zhang is with the Department of Mathematics, University of Georgia, Athens, GA 30602 USA.

Publisher Item Identifier S 0018-9286(97)06603-8.