discretization" of S^1 or of a "linear interpolation" becomes clear. Let $\delta_{\mathrm{I}} \le \delta/3$. Note that the slow feedback $z_0 \mapsto \omega^{(1)}_{\delta_{\mathrm{I}}}(z_0)$ is Lipschitz on S^1. Thus, we can assume that an equidistant discretization $\{z_0^1, \cdots, z_0^k\} \subset (0, 2\pi]$ exists such that the slow feedback is a linear interpolation. Note, furthermore, that the finite set of fast feedbacks $y \mapsto \omega_{\delta_{\mathrm{II}}}(z_0^i, y)$, $i = 1, \cdots, k$, has a common Lipschitz constant. Let $\delta_{\mathrm{I}} \le \delta/3$. Then, we can find an equidistant discretization $\{y^1, \cdots, y^l\} \subset (0, 2\pi]$ such that, for any z_0^i, the fast feedback is a linear interpolation. For slow states $z \in [z_0^i, z_0^{i+1}]$ and fast states $y \in [y^j, y^{j+1}]$, we interpolate setting

$$\omega_\delta(z, y) := \omega_{\delta_{\mathrm{II}}}\left(z_0^i, y\right) + \frac{z - z_0^i}{z_0^{i+1} - z_0^i}\left(\omega_{\delta_{\mathrm{II}}}\left(z_0^{i+1}, y\right) - \omega_{\delta_{\mathrm{II}}}\left(z_0^i, y\right)\right).$$

Then, the average $(1/T)\int_0^T A(y(t), \omega_\delta(z, y(t)))\,dt$ coincides with $\omega^{(1)}_{\delta_{\mathrm{I}}}(z)$ by the (affine) linearity of $\omega \mapsto A(y, \omega)$. We chose $\epsilon_\delta \in (0, 1]$ small enough such that for all $\epsilon \in (0, \epsilon_\delta]$ the right-hand side of (15) becomes smaller than $\delta/3$ and the proof is finished. $\square$

IV. An Example

We illustrate the result of the previous section by the following example of a singularly perturbed oscillator:

$$A(y, u) = \begin{pmatrix} 0 & 1 \\ -1 + y_1 u & 0 \end{pmatrix}, \quad g(y) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

with control range $\Omega = [-1, 1]$. Again, we write $(z_1, z_2) := (s_1, s_2)/\sqrt{s_1^2 + s_2^2}$. According to our open-loop analysis, we know that a constant control function $u(t) = \omega \in [-1, 1]$ can only asymptotically provide a zero Lyapunov exponent for the slow motion s, which follows from the fact that $y_1(t)\omega = \alpha \sin(t + \beta)\omega$ with positive constants $\alpha, \beta > 0$ and that accordingly the averaged slow subsystem is just an undamped linear oscillator. Thus, the orbits are circles. When the control function is allowed to change with respect to time the picture becomes different: the averages $(1/2\pi)\int_0^{2\pi} y_1(t)u(t)\,dt$ can take any value in the interval $[-2/\pi, 2/\pi]$, and the control range of the averaged slow subsystem becomes

$$\Omega^{(1)} = \left\{\begin{pmatrix} 0 & 1 \\ -1 + v & 0 \end{pmatrix} : v \in [-2/\pi, 2/\pi]\right\}.$$

For constant averaged control functions, the orbits of the averaged slow subsystem are ellipses, "horizontally stretched" for $v > 0$, "vertically stretched" for $v < 0$. Intuitively, to stabilize the averaged slow subsystem, we have to choose $v \in [-2/\pi, 2/\pi]$ nonnegative for (z_1, z_2) in the first or third quadrant and nonpositive for (z_1, z_2) in the second or fourth quadrant. This motivates the following composite state feedback for stabilization:

$$u1(t) = -z_1(t)z_2(t)y_1(t).$$

In fact, this feedback produces a singularly perturbed ordinary differential equation with corresponding averaged slow subsystem

$$\frac{d}{dt}\begin{pmatrix} s_1(t) \\ s_2(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 - 2z_1(t)z_2(t)/\sqrt{2} & 0 \end{pmatrix}\begin{pmatrix} s_1(t) \\ s_2(t) \end{pmatrix}.$$

Straightforward calculations show that this averaged differential equation has a negative Lyapunov exponent. Similarly, the system can be destabilized by the composite state feedback

$$u2(t) = z_1(t)z_2(t)y_1(t).$$

References

[1] L. Arnold, H. Crauel, and J.-P. Eckmann, Eds., "Lyapunov exponents," in *Lecture Notes in Mathematics*. Berlin, Germany: Springer, 1991, p. 1486.

[2] Z. Artstein, "Stabilization with relaxed controls," *Nonlinear Anal., TMA*, vol. 7, no. 11, pp. 1163–1173, 1983.

[3] Z. Artstein and V. G. Gaitsgory, "Tracking fast trajectories along a slow dynamics: A singular perturbation approach," *SIAM J. Contr. Optim.*, vol. 35, no. 5, pp. 1487–1507, 1997.

[4] Z. Artstein and A. Vigodner, "Singularly perturbed ordinary differential equations with dynamic limits," *Proc. Royal Soc. Edinburgh A*, vol. 126, no. 3, pp. 541–569, 1996.

[5] F. Colonius and W. Kliemann, "Minimal and maximal Lyapunov exponents of bilinear control systems," *JDE*, vol. 105, no. 5, pp. 232–275, 1993.

[6] M. Corless and L. Glielmo, "On the exponential stability of singularly perturbed systems," *SIAM J. Contr. Optim.*, vol. 30, no. 6, pp. 1338–1360, 1992.

[7] V. G. Gaitsgory, "A new reduction technique for a class of singularly perturbed optimal control problems," *IEEE Trans. Automat. Contr.*, vol. 40, pp. 721–724, 1995.

[8] G. Grammel, "Averaging of singularly perturbed systems," *Nonlinear Anal., TMA*, vol. 28, no. 11, pp. 1851–1865, 1997.

[9] ——, "Singularly perturbed control systems: Recent progress," in *Proc. 35th IEEE Conf. Decision Contr.*, Kobe, Japan, 1996, pp. 505–510.

[10] L. Gruene, "Discrete feedback stabilization of semilinear control systems," *ESAIM: Contr., Optim. Calculus Variations*, vol. 1, pp. 207–227, 1996.

[11] P. V. Kokotovic, H. K. Khalil, and J. O'Reilly, *Singular Perturbation Methods in Control: Analysis and Design*. San Diego, CA: Academic, 1986.

[12] E. D. Sontag, *Mathematical Control Theory*. Berlin, Germany: Springer-Verlag, 1990.

Convergence Results of the Analytic Center Estimator

Er-Wei Bai, Minyue Fu, Roberto Tempo, and Yinyu Ye

***Abstract*—The analytic center approach for bounded error parameter estimation was recently proposed as an alternative to the well-known least squares and Chebyshev estimates. In this paper, we show the asymptotic performance of this approach and prove that the analytic center converges to the true parameter under mild conditions.**

***Index Terms*—Analytic center, bounded error parameter estimation, convergence analysis, membership set identification.**

I. Introduction

Consider a single input–single output discrete-time system

$$y_i = \phi_i^T\theta + v_i, \qquad i = 1, 2, \cdots, n, \tag{1.1}$$

Manuscript received December 18, 1998; revised October 18, 1999. Recommended by Associate Editor, J. Chen. This work was supported in part by the National Science Foundation, the ARC (Australia), and the CENS-CNR.

E.-W. Bai is with the Department of Electrical and Computer Engineering, University of Iowa, Iowa City, IA 52242 USA (e-mail: erwei@icaen.uiowa.edu).

M. Fu is with the Department of Electrical and Computer Engineering, University of Newcastle, N.S.W., 2308 Australia (e-mail: eemf@hartley.newcastle.edu.au).

R. Tempo is with the CENS-CNR, Politecnico di Torino, Torino, Italy (e-mail: tempo@polito.it).

Y. Ye is with the Department of Management Science, University of Iowa, Iowa City, IA 52242 USA (e-mail: yyye@dollar.biz.uiowa.edu).

Publisher Item Identifier S 0018-9286(00)02167-X.

where $y_i \in \mathbf{R}$ is the system output, $\phi_i \in \mathbf{R}^m$ is the measurable regressor, $\theta \in \mathbf{R}^m$ is the unknown parameter vector to be identified, and $v_i \in \mathbf{R}$ is the noise. In the bounded error parameter estimation setting, it is assumed that the noise v_i is bounded by $\epsilon > 0$, i.e.,

$$|v_i| \leq \epsilon \tag{1.2}$$

for $i = 1, 2, \cdots, n$. Then, the membership set is defined by

$$\Omega^n = \bigcap_{i=1}^{n} \left\{ \hat{\theta} \in \mathbf{R}^m \colon -\epsilon \leq y_i - \phi_i^T \hat{\theta} \leq \epsilon \right\}. \tag{1.3}$$

The membership set includes the parameter estimates consistent with (1.1), the input–output data, and the noise bound (1.2). The goal is to compute a specific estimate in the membership set enjoying some optimality properties. A large body of research exists (see, for example, [7] and [9]) in this area. The most popular estimate along this direction is the Chebyshev center θ_c [10] of the set Ω^n

$$\theta_c = \arg \min_{\hat{\theta} \in \Omega^n} \max_{\eta \in \Omega^n} \|\hat{\theta} - \eta\|$$

where $\|\cdot\|$ is any l_p norm. If $p = \infty$, this is the best worst case estimate of the true but unknown system parameter vector in the sense that it minimizes the maximum "distance" between θ_c and the unknown parameter vector that generated the data. However, it is well known that the Chebyshev center is sensitive to outliers, and moreover, online sequential implementation of the Chebyshev center does not seem feasible. To overcome these difficulties, an analytic center estimate θ_a was proposed in [2]–[4]. In particular, it was shown in [2] that the analytic center θ_a minimizes the logarithmic average output error and can be implemented in a sequential form. The complexity of this sequential algorithm for computing a sequence of analytic centers up to observation time n is linear in terms of the number of Newton iterations.

To be more specific, let us rewrite the membership set

$$\begin{aligned} \Omega^n &= \bigcap_{i=1}^{n} \left\{ \hat{\theta} \in \mathbf{R}^m \colon -\epsilon \leq y_i - \phi_i^T \hat{\theta} \leq \epsilon \right\} \\ &= \{ \hat{\theta} \in \mathbf{R}^m \colon (A^n)^T \hat{\theta} \leq c^n \} \end{aligned} \tag{1.4}$$

where

$$\begin{aligned} A^n &= (a_1, a_2, \cdots, a_{2n}) \\ &\doteq (\phi_1, -\phi_1, \phi_2, -\phi_2, \cdots, \phi_n, -\phi_n) \in \mathbf{R}^{m \times 2n} \end{aligned}$$

and

$$c^n = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{2n-1} \\ c_{2n} \end{pmatrix} \doteq \begin{pmatrix} \epsilon + y_1 \\ \epsilon - y_1 \\ \vdots \\ \epsilon + y_n \\ \epsilon - y_n \end{pmatrix} \in \mathbf{R}^{2n}.$$

Suppose the set Ω^n is bounded and has a nonempty interior. Then, the analytic center θ_a of Ω^n is an interior point of Ω^n maximizing the (dual) potential function

$$\theta_a = \arg \max_{\hat{\theta} \in \Omega^n} \psi(\hat{\theta}) = \arg \max_{\hat{\theta} \in \Omega^n} \frac{1}{2n} \sum_{i=1}^{2n} \ln\left(c_i - a_i^T \hat{\theta}\right). \tag{1.5}$$

The following theorem was given in [2] concerning the analytic center.

Property 1.1: Consider the system (1.1), the noise bound (1.2), and the membership set (1.3). Then, the analytic center is the solution to the following optimization problem:

$$\begin{aligned} \theta_a &= \arg \max_{\hat{\theta} \in \Omega^n} \prod_{i=1}^{n} \left(\epsilon^2 - \left(y_i - \phi_i^T \hat{\theta} \right)^2 \right) \\ &= \arg \max_{\hat{\theta} \in \Omega^n} \sum_{i=1}^{n} \ln\left(\epsilon^2 - \left(y_i - \phi_i^T \hat{\theta} \right)^2 \right). \end{aligned} \tag{1.6}$$

The main result of the present paper is to show the convergence of the analytic center to the true parameter under various conditions. A preliminary version was published in the conference [5].

II. Convergence Analysis

To avoid some unnecessary complications, we modify the noise bound and assume that

$$|v_i| \leq \epsilon - \xi < \epsilon \tag{2.1}$$

for some arbitrarily small constant $\xi > 0$. Further, we define two conditions, as follows.

Condition 1: $n_0 > 0$ and $\beta_1 \geq \beta_2 > 0$ exist such that for all $n \geq n_0$

$$\beta_1 I \geq \frac{1}{n} \sum_{i=1}^{n} \phi_i \phi_i^T \geq \beta_2 I. \tag{2.2}$$

In the literature, this condition is referred to as the weak persistent excitation condition [6]. Condition 1 and (2.1) guarantee that the set Ω^n is bounded and has a nonempty interior.

Condition 2:

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \frac{v_i}{\epsilon^2 - v_i^2} \phi_i = 0. \tag{2.3}$$

Condition 2 is satisfied if v_i and ϕ_i are independent random variables with $\mathbf{E}(\phi_i) = 0$ or v_i is symmetric; i.e., the probability density function $q(v_i) = q(-v_i)$, which implies $\mathbf{E}(v_i/(\epsilon^2 - v_i^2)) = 0$. The above condition also holds when ϕ_i is a deterministic time function and v_i is a symmetric independent random variable or when v_i is a deterministic time function, but ϕ_i is an independent random variable with $\mathbf{E}(\phi_i) = 0$.

Three results are presented below. The first one answers the question when the analytic center gives a correct estimate; i.e., $\theta_a = \theta$. The second theorem, which is the main result in this section, studies the convergence properties of the analytic center. The third result is a consequence of the second theorem, which shows the convergence of θ_a to the unknown θ with probability one if v_i is a random variable with zero mean.

Lemma 2.1: Under Condition 1, $\theta_a = \theta$ if and only if

$$\frac{1}{n} \sum_{i=1}^{n} \frac{v_i}{\epsilon^2 - v_i^2} \phi_i = 0. \tag{2.4}$$

Proof: Condition 1 guarantees the existence and uniqueness of the analytic center θ_a for $n \geq n_0$. Further, θ_a is the solution of

$$\frac{d\psi(\hat{\theta})}{d\hat{\theta}} = 0$$

where the potential function $\psi(\hat{\theta})$ is defined in (1.5), which is equivalent to

$$\frac{1}{n} \sum_{i=1}^{n} \frac{\phi_i \left(y_i - \phi_i^T \theta_a \right)}{\epsilon^2 - \left(y_i - \phi_i^T \theta_a \right)^2} = 0. \tag{2.5}$$

If $\theta_a = \theta$, then this equation reduces to (2.4). Conversely, if (2.4) holds, then $\theta_a = \theta$ is a solution to the above. By the uniqueness of θ_a under Condition 1, this is the only solution. Hence, $\theta_a = \theta$ if and only if (2.4) holds.

Theorem 2.1: The parameter estimation error given by the analytic center has the following bound for $n \geq n_0$:

$$\|\theta_a - \theta\| \leq \left\| \left(\frac{1}{n} \sum_{i=1}^{n} \frac{\phi_i \phi_i^T}{\epsilon^2 - v_i^2} \right)^{-1} \right\| \cdot \left\| \frac{1}{n} \sum_{i=1}^{n} \frac{2v_i}{\epsilon^2 - v_i^2} \phi_i \right\|. \quad (2.6)$$

In particular, $\theta_a \to \theta$ as $n \to \infty$ if Conditions 1 and 2 hold.

Proof: Define the parameter estimation error $e = \theta - \theta_a$ and prediction error $\rho_i = y_i - \phi_i^T \theta_a$. Then

$$\rho_i = v_i + \phi_i^T e.$$

Subsequently, (2.5) gives the unique solution to θ_a and is equivalent to

$$-\sum_{i=1}^{n} \frac{v_i \phi_i}{\epsilon^2 - \rho_i^2} = \sum_{i=1}^{n} \frac{\phi_i \phi_i^T}{\epsilon^2 - \rho_i^2} e. \quad (2.7)$$

Define

$$\eta_i = \frac{v_i \phi_i}{\epsilon^2 - \rho_i^2} - \frac{v_i \phi_i}{\epsilon^2 - v_i^2}.$$

Then, it is straightforward to check that

$$\eta_i = \frac{\phi_i \phi_i^T}{\epsilon^2 - \rho_i^2} \frac{v_i^2 + v_i \rho_i}{\epsilon^2 - v_i^2} e.$$

Adding $\sum_{i=1}^{n} \eta_i$ to both sides of (2.7), we obtain

$$\begin{aligned} -\sum_{i=1}^{n} \frac{v_i \phi_i}{\epsilon^2 - v_i^2} &= \sum_{i=1}^{n} \frac{\phi_i \phi_i^T}{\epsilon^2 - \rho_i^2} \left(1 + \frac{v_i^2 + v_i \rho_i}{\epsilon^2 - v_i^2} \right) e \\ &= \left(\sum_{i=1}^{n} \frac{\epsilon^2 + v_i \rho_i}{(\epsilon^2 - v_i^2)(\epsilon^2 - \rho_i^2)} \phi_i \phi_i^T \right) e \\ &= \left(\sum_{i=1}^{n} \frac{d_i}{\epsilon^2 - v_i^2} \phi_i \phi_i^T \right) e \end{aligned}$$

where

$$d_i = \frac{\epsilon^2 + v_i \rho_i}{\epsilon^2 - \rho_i^2}.$$

Noting that $|v_i| < \epsilon$ and $|\rho_i| < \epsilon$, we get

$$d_i \geq \tfrac{1}{2}.$$

Thus

$$e = -\left(\sum_{i=1}^{n} \frac{2d_i}{\epsilon^2 - v_i^2} \phi_i \phi_i^T \right)^{-1} \left(\sum_{i=1}^{n} \frac{2v_i}{\epsilon^2 - v_i^2} \phi_i \right) \quad (2.8)$$

and

$$\|\theta - \theta_a\| \leq \left\| \left(\frac{1}{n} \sum_{i=1}^{n} \frac{\phi_i \phi_i^T}{\epsilon^2 - v_i^2} \right)^{-1} \right\| \cdot \left\| \frac{1}{n} \sum_{i=1}^{n} \frac{2v_i}{\epsilon^2 - v_i^2} \phi_i \right\|.$$

Moreover, from Condition 1 and the fact that

$$0 < \xi(2\epsilon - \xi) \leq \epsilon^2 - v_i^2 \leq \epsilon^2$$

the first term is always bounded. Then, the convergence of θ_a to θ follows from Condition 2. This completes the proof.

The geometric interpretation of the result is that the Hessian of the logarithmic barrier at the analytic center of a polyhedron defines two ellipsoids inscribing and outscribing the polyhedron [1], [2].

Corollary 2.1: Consider the system (1.1). Suppose that the noise v_i is a sequence of independently (not necessarily identically) distributed random variables of zero mean with the bound given by (2.1). Further, assume that the regressor ϕ_i satisfies Condition 1 and is independent of v_i. Then, the analytic center θ_a converges to the true but unknown θ with probability one as $n \to \infty$.

Proof: The hypothesis implies that the random variables $v_i/(\epsilon^2 - v_i^2)$ are independent with zero mean and finite convariance. Then, by the law of large numbers [8], Condition 2 is satisfied with probability one as $n \to \infty$. Therefore, the conclusion follows.

III. Some Remarks

It is interesting to compare the convergence conditions for the analytic center approach with the recursive least squares (RLS) given by

$$\begin{aligned} \theta_{ls} &= \left(\frac{1}{n} \sum_{i=1}^{n} \phi_i \phi_i^T \right)^{-1} \frac{1}{n} \sum_{i=1}^{n} \phi_i y_i \\ &= \theta + \left(\frac{1}{n} \sum_{i=1}^{n} \phi_i \phi_i^T \right)^{-1} \frac{1}{n} \sum_{i=1}^{n} \phi_i v_i. \end{aligned} \quad (3.1)$$

For the convergence of the RLS estimate, Condition 1 remains the same while Condition 2 is replaced with a slightly simpler condition

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} v_i \phi_i^T = 0. \quad (3.2)$$

The main tradeoff is that RLS is much simpler, but does not guarantee that the solution lies in the membership set all of the time. On the other hand, the analytic center requires more computation [2], in particular, when n is large, but at the same time, it is guaranteed to lie in the membership set. Therefore, it would be nice to start with the analytic center and then switch to RLS when n becomes large. This mixed approach is to start with an analytic center estimator and then switch to an RLS estimator when the estimate begins to converge. The key point, however, is to determine the switching time. To further elaborate on this, we assume zero mean and independence of the noise (not necessarily identical distributions). Then, the covariance of the RLS estimate error becomes

$$\mathrm{Cov}(\theta_{ls} - \theta) = R_n^{-1} \frac{1}{n^2} \sum_{i=1}^{n} \sum_{k=1}^{n} \phi_i \phi_k^T \mathbf{E}(v_i v_k) R_n^{-1}$$

where

$$R_n = \frac{1}{n} \sum_{i=1}^{n} \phi_i \phi_i^T.$$

Denote

$$\sigma^2 = \max_i \mathbf{E}\left(v_i^2\right) = \max_i \int_{-\epsilon}^{\epsilon} x^2 q_i(x)\, dx$$

where $q_i(x)$ is the probability density function of v_i. Then

$$\mathrm{Cov}(\theta_{ls} - \theta) \leq \frac{\sigma^2}{n} \left(\frac{1}{n} \sum_{i=1}^{n} \phi_i \phi_i^T \right)^{-1} \leq \frac{\sigma^2}{\beta_2 n} I \quad (3.3)$$

where β_2 is defined in (2.2). Note that $\sigma^2 \leq \epsilon^2$ because $|v_i| \leq \epsilon$ and

$$\int_{-\epsilon}^{\epsilon} x^2 q_i(x)\, dx \leq \epsilon^2 \int_{-\epsilon}^{\epsilon} q_i(x)\, dx \leq \epsilon^2.$$

Accordingly, $\mathrm{Cov}(\theta_{ls} - \theta) \leq (\epsilon^2/\beta_2 n)I$ and

$$\mathbf{E}(\|\theta_{ls} - \theta\|^2) = \mathrm{Trace}\{\mathrm{Cov}(\theta_{ls} - \theta)\} \leq \frac{m\epsilon^2}{\beta_2 n}.$$

Now, to determine an appropriate switching time, we use the well-known Chebyshev inequality in the following lemma [8].

Lemma 3.1: Consider any random variable x. Then, for any constant $\rho > 0$

$$\mathrm{Prob}\{\|x - \mathbf{E}(x)\| \geq \rho\} \leq \frac{\mathbf{E}(\|x - \mathbf{E}(x)\|^2)}{\rho^2}.$$

Using the above lemma, we consider two arbitrary tolerance parameters $\rho > 0$ and $\delta > 0$. The former is used to characterize the "distance" between θ and θ_{ls}, whereas the latter is for bounding the probability $\mathrm{Prob}\{\|\theta_{ls} - \theta\| \geq \rho\}$. We may choose the switching time n determined by

$$n \geq \frac{m\epsilon^2}{\beta_2 \delta \rho^2}. \tag{3.4}$$

By the Chebyshev inequality, we obtain

$$\mathrm{Prob}\{\|\theta_{ls} - \theta\| \geq \rho\} \leq \frac{\mathbf{E}(\|\theta_{ls} - \theta\|^2)}{\rho^2} \leq \frac{m\epsilon^2}{\rho^2 \beta_2 n} \leq \delta. \tag{3.5}$$

Clearly, no guarantee exists that, if $n \geq m\epsilon^2/\beta_2\delta\rho^2$, the RSL estimate lies in the membership set. However, with a probability of at least $1 - \delta$, the RLS estimate θ_{ls} is ρ-close to the true θ, provided that (3.4) is satisfied, and thus, for small enough $\rho > 0, \theta_{ls}$ is likely to be in the membership set. Therefore, we can "safely" switch the estimate from the analytic center to least squares. The above idea is easily implementable.

REFERENCES

[1] K. Afkhamie, Z. Q. Luo, and K. M. Wong, "Interior point column generation algorithms for adaptive filtering," Dept. Elect. Comput. Eng., McMaster Univ., 1997. Tech. Rep..

[2] E. W. Bai, Y. Ye, and R. Tempo, "Bounded error parameter estimation: A sequential analytic center approach," *IEEE Trans. Automat. Contr.*, vol. 44, pp. 1107–1117, 1999.

[3] ——, "Recent results on the analytic center approach for bounded error parameter estimation," in *Learning, Control and Hybrid Systems*, Y. Yamamoto and S. Hara, Eds. Berlin, Germany: Springer-Verlag, 1998, pp. 245–254.

[4] ——, "Open problems in sequential parametric estimation," in *Open Problems in Mathematical Systems Theory and Control*, V. Blondel, E. Sontag, M. Vidyasagar, and J. Willems, Eds. Berlin, Germany: Springer-Verlag, 1998, pp. 19–22.

[5] E. W. Bai, M. Fu, R. Tempo, and Y. Ye, "Analytic center approach to parameter estimation: Convergence analysis," in *Proc. 1998 ICASSP*, Seattle, WA.

[6] G. Goodwin and K. S. Sin, *Adaptive Filtering, Prediction and Control*. New York: Prentice-Hall, 1984.

[7] M. Milanese and A. Vicino, "Optimal estimation theory for dynamic systems with set membership uncertainty: An overview," *Automatica*, vol. 27, pp. 997–1009, 1991.

[8] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991.

[9] H. Piet-Lahanier and E. Walter, "Exact description of feasible parameter sets and minimax estimator," *Int. J. Adap. Contr. Signal Proc.*, vol. 8, pp. 5–14, 1994.

[10] J. Traub, G. Wasilkowski, and H. Wozniakowski, *Information Based Complexity*. New York: Academic, 1988.

Additional Dynamics in Transformed Time-Delay Systems

Keqin Gu and Silviu-Iulian Niculescu

***Abstract*—In studying the stability of time delay systems, many published results use a transformation to transform a system with single time delay to a system with distributed delay. In this article, the inherent limitations of such approaches are studied. Specifically, it is shown that such a transformation incurs additional dynamics that can be characterized by appropriate additional eigenvalues. The critical delay values when such additional eigenvalues cross the imaginary axis can be explicitly calculated. If the smallest of such delays is less than the stability delay limit of the original system, then any stability criteria obtained using such transformation will be conservative. Some examples are also included.**

***Index Terms*—Razumikhin theorem, stability, time-delay systems.**

I. INTRODUCTION

The study of stability of time-delay systems has been active in recent years. See, for example, [25], [14], [19], [21], [6], [7], [24], and [23] and the references therein. See also [16] for an overview.

For the stability of the system

$$\dot{x}(t) = Ax(t) + A_d x(t - r) \tag{1}$$

where $x(t) \in \mathcal{R}^n$; $A, A_d \in \mathcal{R}^{n \times n}$, many published results use the fact

$$\begin{aligned} x(t-r) &= x(t) - \int_{-r}^{0} \dot{x}(t+\tau)\, d\tau \\ &= x(t) - \int_{-r}^{0} [Ax(t+\tau) + A_d x(t+\tau-r)]\, d\tau \end{aligned}$$

to transform the above system to a distributed delay

$$\begin{aligned} \dot{x}(t) = &(A + A_d)x(t) \\ &- A_d \int_{-r}^{0} [Ax(t+\tau) + A_d x(t+\tau-r)]\, d\tau. \end{aligned} \tag{2}$$

This transforms a system with *discrete delay* (or pointwise delay) (1) to one with *distributed delay* (2), according to the terminology in [10]. In this paper, we will focus on the special case that A and A_d are constant matrices.

It should be realized that systems (1) and (2) are indeed different. The complete prediction of future trajectory of (1) requires the knowledge of $x(t)$ over a time interval of length r, for example, $t \in [-r, 0]$. However, the knowledge of $x(t)$ over a time interval of length $2r$, for example, $t \in [-r, r]$, is needed to predict the future trajectory of (2).

It was proven in [15] (in a more general case of systems with structured uncertainty) that the stability of (2) implies the stability of (1),

Manuscript received January 26, 1999; revised March 8, 1999. Recommended by Associate Editor, R. Datko. This work was supported in part by the National Science Foundation under Grant INT-9818312 and the CNRS under Grant 7404, 1999–2002.

K. Gu is with the Department of Mechanical and Industrial Engineering, Southern Illinois University at Edwardsville, Edwardsville, IL 62026 USA (e-mail: kgu@siue.edu).

S.-I. Niculescu is with Heudiasyc (UMR CNRS 6599), Université de Technologie de Compiègne, Centre de Recherche de Royallieu, 60205, Compiegne, France (e-mail: silviu@hds.utc.fr).

Publisher Item Identifier S 0018-9286(00)02168-1.