

where

$$v_\gamma(t, e) = \frac{1}{2} e^T P(x) e = \frac{1}{2} e^T \begin{bmatrix} \gamma K_p & M \\ M & \gamma M \end{bmatrix} e \quad (\text{A.3})$$

and

$$v_{ci}(t, \hat{\beta}_i, \hat{d}_i) = (\hat{\beta}_i - \beta)^T \Gamma_i^{-1} (\hat{\beta}_i - \beta) + (\hat{d}_i - d_i)^T \bar{\Gamma}_i^{-1} (\hat{d}_i - d_i) + g_i^{-1} \epsilon_i. \quad (\text{A.4})$$

Then take the time derivative of  $v_c$  along the solution trajectory of (A.1) using (3.10) to show that

$$\frac{d}{dt} v_c \leq -q |e|^2 \quad (\text{A.5})$$

for some  $q > 0$  for sufficiently large  $\lambda_{\min}(K_p)$  and  $\gamma$ , which further implies that  $e \in L_{2 \cap L_\infty}$  and  $\hat{\beta}_i, \hat{d}_i \in L_\infty$  for  $i = 1, \dots, n$ . As a result,  $u \in L_\infty$  and, hence,  $\dot{e} \in L_\infty$  so that from Barbalet's Lemma [17] it follows that  $e \rightarrow 0$  as  $t \rightarrow \infty$ . Q.E.D.

#### REFERENCES

- [1] J. T. Wen and D. S. Bayard, "New class of control laws for robotic manipulators — Part I: Nonadaptive case," *Int. J. Contr.*, vol. 47, no. 5, pp. 1361–1385, 1988.
- [2] D. S. Bayard and J. T. Wen, "New class of control law for robotic manipulators — Part II: Adaptive case," *Int. J. Contr.*, vol. 47, no. 5, pp. 1387–1406, 1988.
- [3] Y. H. Chen, "Adaptive robust model-following control and application to robot manipulators," *ASME Trans. Dynam. Syst., Measurement, Contr.*, vol. 109, pp. 209–215, 1987.
- [4] —, "Adaptive robust observers of nonlinear uncertain systems," Personal correspondence; *Int. J. Contr.*, 1989, submitted for publication.
- [5] M. S. Chen, "Adaptive estimation of bounds on parameter uncertainty with an application in robust control," Personal correspondence.
- [6] M. Corless and G. Leitmann, "Adaptive control of systems containing uncertain functions and unknown functions with uncertain bounds," *J. Optimiz. Theory Appl.*, vol. 41, pp. 155–168, 1983.
- [7] —, "Adaptive control for uncertain dynamic systems," in *Dynamical Systems and Microphysics, Control Theory and Mechanics*, A. Blaquiere and G. Leitmann, Eds. New York: Academic, 1984.
- [8] J. J. Craig, P. Hsu, and S. S. Sastry, "Adaptive control of mechanical manipulators," *Int. J. Robotics Res.*, vol. 6, no. 2, pp. 16–28, 1987.
- [9] L. C. Fu and T. Y. Fu, "How effective can a local PID feedback control really be for trajectory tracking of robot motion?" in *Proc. 12th Symp. Automat. Contr.*, Shing-Chu, Taiwan, 1988, pp. 135–144.
- [10] —, "Robust adaptive independent joint control of robot manipulators," in *Proc. 12th Symp. Automat. Contr.*, Shing-Chu, Taiwan, 1988, pp. 363–371.
- [11] P. Hsu, M. Bodson, S. S. Sastry, and B. Paden, "Adaptive identification and control for manipulators without using joint accelerations," in *Proc. IEEE Int. Conf. Robotics Automation*, Raleigh, NC, 1987.
- [12] D. E. Koditschek, "Quadratic Lyapunov functions for mechanical systems," in *Proc. IEEE Int. Conf. Robotics Automation*, Raleigh, NC, 1987.
- [13] —, "Strict global Lyapunov functions for mechanical systems," in *Proc. IEEE Int. Conf. Robotics Automation*, Raleigh, NC, 1987.
- [14] R. Middleton, and G. Goodwin, "Adaptive computed torque control for rigid link manipulators," in *Proc. IEEE Conf. Decision Contr.*, Athens, Greece, 1986.
- [15] B. J. Oh, M. Jamshidi, and H. Seraji, "Decentralized adaptive control," in *Proc. IEEE Int. Conf. Robotics Automation*, Philadelphia, PA, 1988, pp. 1016–1021.
- [16] B. Paden and B. Riedle, "A positive-real modification of a class of nonlinear controller for robot manipulators," in *Proc. Amer. Contr. Conf.*, Atlanta, GA, 1988, pp. 1782–1785.
- [17] V. M. Popov, *Hyperstability of Control Systems*. New York: Springer-Verlag, 1973.

- [18] H. Seraji, "Adaptive independent joint control of manipulators: Theory and experiment," in *Proc. IEEE Int. Conf. Robotics Automation*, Philadelphia, PA, 1988, pp. 854–861.
- [19] S. N. Singh, "Adaptive model following control of nonlinear robotic system," *IEEE Trans. Automat. Contr.*, vol. AC-30, no. 11, pp. 1099–1100, 1985.
- [20] J. E. Slotine and W. Li, "On the adaptive control of robot manipulators," *Int. J. Robotics Res.*, vol. 6, no. 3, pp. 49–59, 1987.
- [21] B. S. Yuan and W. J. Book, "Decentralized adaptive control of robot manipulators with robust stabilization design," in *Proc. Amer. Contr. Conf.*, Atlanta, GA, 1988, pp. 102–107.

#### Stability of Learning Control with Disturbances and Uncertain Initial Conditions

Greg Heinzinger, Dan Fenwick, Brad Paden, and Fumio Miyazaki

**Abstract**—In this note, we investigate the effects of state disturbances, output noise, and errors in initial conditions on a class of learning control algorithms. We present a simple learning algorithm and exhibit, via a concise proof, bounds on the asymptotic trajectory errors for the learned input and the corresponding state and output trajectories. Furthermore, these bounds are continuous functions of the bounds on the initial condition errors, state disturbances, and output noise, and the bounds are zero in the absence of these disturbances.

#### INTRODUCTION

Learning control is a name attributed to a class of self-tuning processes whereby the system performance of a specified task improves based on the previous performances of identical tasks. This is an advantage when controlling systems that cannot be modeled accurately. But the idea of a self-learning system is in itself aesthetically appealing in that it represents a significant step in the development of an intelligent fully autonomous control system.

Consider the "learning" system depicted in Fig. 1.  $u_i(\cdot)$  denotes an input trajectory. The plant  $P$  has desired output trajectory  $y_d(\cdot)$  and actual output  $y_i(\cdot)$  due to  $u_i(\cdot)$ . The learning operator  $l$  compares  $y_d(\cdot)$  and  $y_i(\cdot)$  and adds an update term to  $u_i(\cdot)$  to produce  $u_{i+1}(\cdot)$ . These trajectories are functions of  $t \in [0, T]$  and the updates occur sequentially in time. The trajectories are supported on finite intervals of the time axis and the iteration from  $i$  to  $i+1$  occurs from one interval to the next. In a sense,  $u_i(\cdot)$  is a parameter that is adaptively tuned. In contrast to typical adaptive control schemes, the parameter  $u_i(\cdot)$  belongs to an infinite-dimensional space.

Manuscript received September 23, 1989; revised July 18, 1990. Paper recommended by Past Associate Editor, B. H. Krogh. The work of G. Heinzinger was supported by the Berkeley Fellowship and by Project MICRO of the University of California, funded by Semiconductor Research Corporation, Harris Corporation, Intel, Phillips Research Labs, Fairchild, Hewlett Packard, Texas Instruments, AMD, Inference Corporation, and Siemens Corporation. The work of D. Fenwick and B. Paden was supported by the National Science Foundation under Grant 8421415.

G. Heinzinger is with the Department of Engineering and Computer Science and the Electronics Research Laboratory, University of California at Berkeley, Berkeley, CA 94720.

D. Fenwick and B. Paden are with the Department of Mechanical Engineering, University of California at Santa Barbara, Santa Barbara, CA 93106.

F. Miyazaki is with the Department of Engineering Science, Osaka University, Osaka, Japan.

IEEE Log Number 9103081.

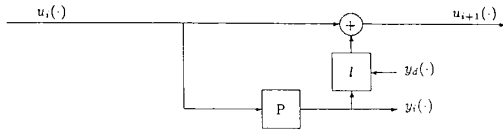


Fig. 1. Basic learning system.

As an example, a learning control scheme for a robot manipulator would record measurements as it moved an object from point A to point B; it would then use this data to improve its performance the next time it moved the same object from point A to point B. In some applications the need to repeat a trajectory multiple times is a distinct disadvantage of learning control. In many applications, however, repetitive tasks are commonly performed making learning control a very natural solution. Another advantage of learning control is that it is easy to implement and allows simple models and control schemes to be used, while compensating for unmodeled dynamics and complex phenomenon such as stiction. It is also appealing because it is similar to some of our own learning processes; we may practice a task (say throwing a ball) many times before we are able to find inputs to a complex system (our body) to accomplish the task.

Recently, there have been a number of efforts toward defining and analyzing learning control schemes [2]–[7]. In model-based learning schemes [5], the inputs corresponding to the desired and actual trajectories are computed from estimated system parameters and the resulting input errors fed to the learning operator. In this scheme, the performance of the algorithm depends on the quality of the parameter estimates. A more common approach is to operate on the output errors directly, and the model-based learning scheme in [5] is shown in [6] to be a special case of this more general approach. The basic strategy of these techniques is to use an iteration of the form  $u_{i+1}(\cdot) = l(u_i(\cdot), y_d(\cdot) - y_i(\cdot))$ , where the operator  $l(\cdot, \cdot)$  remains to be specified.

For time-invariant mechanical systems Arimoto *et al.* [2] and Craig [3] present conditions on the learning operator that guarantee system convergence upon repeated application of the learning algorithm. One shortcoming of these analyses is that they are small-signal analyses that require the assumption that the initial trajectory (and thus all subsequent ones) lies in a neighborhood of the desired trajectory. In addition, no investigation is presented as to the size or existence of these neighborhoods. Both Hauser [6] and Bondi *et al.* [4] remove this assumption by developing global analyses, proving convergence of the input sequence  $u_i(\cdot)$  with any initial trajectory. Another extension of Hauser [6] allows time-varying systems. This is important because we wish to improve the performance of the plant as much as possible using conventional feedback control methods. The learned input  $u_i(\cdot)$  is a feed-forward term that further improves the performance for a specific task. Therefore, for most applications we have the situation shown in Fig. 2, and the learning algorithm operates on the system between  $u_i(\cdot)$  and  $y_i(\cdot)$ , which is time-varying.

For examples illustrating the use of learning to various systems, and for more detailed expositions, the reader is referred to [1]–[7]. In addition, Sugie and Ono [7] demonstrate the need for differentiation in the learning operator for systems without direct feed-through terms.

In the remainder of this note, we consider the effects of uncertainties and disturbances on the learning algorithm. Specifically, does the performance of the algorithm continuously degrade as errors and disturbances are introduced? For a practical implementation we would like to know that the learning algorithm causes the input,

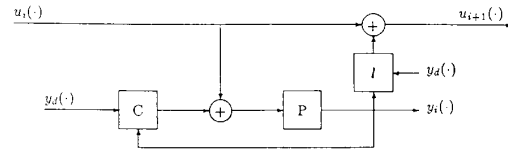


Fig. 2. Learning control application for a plant with a feedback controller attached.

state, and output trajectory errors to be asymptotically bounded when there are 1) errors in the initial state, 2) bounded state disturbances, and 3) bounded output disturbances. In addition, we would like to understand how these bounds depend on the disturbances, and they should decrease to zero as the disturbances do.

In the last few years researchers have begun to answer these questions. Arimoto *et al.* [8] deal with time-invariant mechanical systems and use a small-signal analysis to demonstrate the effects of initial state errors and differentiable state disturbances; the proof once again assumes the initial trajectory is in some small neighborhood. In [4], Bondi *et al.* present a global analysis for time-invariant mechanical systems. While both papers deal with time-invariant mechanical systems, neither one answers all the questions posed above.

In this note, we consider the following class of nonlinear time-varying systems described by the following state-space equations:

$$\begin{aligned} \dot{x}(t) &= f(x(t), t) + B(x(t), t)u(t) \\ y(t) &= g(x(t), t). \end{aligned} \quad (1.1)$$

As mentioned before, this is significant because we can apply our results to a plant and feedback configuration as shown in Fig. 2. The learning operator studied operates on the derivative of the previous output error in a memoryless linear fashion

$$u_{i+1}(t) = u_i(t) + L(\dot{y}_d(t) - \dot{y}_i(t)) \quad (1.2)$$

where  $L$  is a memoryless linear map.

The stability results for this more general class of systems represents the main contribution of our work. The global proof we present is simple, concise, and complete, and it makes explicit the dependence of the bounds on the disturbances and errors.

## II. MAIN RESULT

In this section we present a theorem making explicit the asymptotic errors in performance due to disturbances and errors in the initial conditions. In addition, these errors are given by bounds, which are continuous functions of the bounds on the initial condition errors and the disturbances.

The description of the system, assumptions, notation, and update law are similar to those in [6]; the proof technique is similar to many in that it proceeds in a straightforward manner showing that we have a “contraction” on the input sequence, implying the convergence results.

As specified in (1.1), the class of nonlinear time-varying systems considered is described by the following state-space equations:

$$\begin{aligned} \dot{x}_i(t) &= f(x_i(t), t) + B(x_i(t), t)u_i(t) + w_i(t) \\ y_i(t) &= g(x_i(t), t) \end{aligned} \quad (2.1)$$

where, for all  $t \in [0, T]$ ,  $x_i(t) \in \mathbb{R}^n$ ,  $u_i(t) \in \mathbb{R}^r$ ,  $y_i(t) \in \mathbb{R}^m$ , and  $w_i(t) \in \mathbb{R}^n$ . The functions  $f: \mathbb{R}^n \times [0, T] \rightarrow \mathbb{R}^n$  and  $B: \mathbb{R}^n \times [0, T] \rightarrow \mathbb{R}^{n \times r}$  are piecewise continuous in  $t$ ; and  $g: \mathbb{R}^n \times [0, T] \rightarrow \mathbb{R}^m$  is differentiable in  $x$  and  $t$ , with partial derivatives  $g_x(\cdot, \cdot)$  and  $g_t(\cdot, \cdot)$ . We consider inputs  $u_i: [0, T] \rightarrow \mathbb{R}^r$ , not necessarily continuous. In addition, we assume the following properties.

A1: For each fixed  $x_i(0)$  with  $w_i(\cdot) \equiv 0$  the output map  $\theta: C([0, T], \mathbb{R}^r) \times \mathbb{R}^n \rightarrow C([0, T], \mathbb{R}^m)$  and the state map  $\mathcal{S}: C([0, T], \mathbb{R}^r) \times \mathbb{R}^n \rightarrow C([0, T], \mathbb{R}^n)$  are one-to-one. In this notation  $y_i(\cdot) = \theta(u_i(\cdot), x_i(0))$  and  $x_i(\cdot) = \mathcal{S}(u_i(\cdot), x_i(0))$ .

A2: The distance  $w_i(\cdot)$  is bounded by  $b_w$  on  $[0, T]$  (i.e.,  $\|w_i(t)\| \leq b_w$  on the interval  $[0, T]$ ).

A3: The functions  $f(\cdot, \cdot)$ ,  $B(\cdot, \cdot)$ ,  $g_x(\cdot, \cdot)$  and  $g_t(\cdot, \cdot)$  are uniformly globally Lipschitz in  $x$  on the interval  $[0, T]$ . That is,  $\|h(x_1, t) - h(x_2, t)\| \leq k_h \|x_1(t) - x_2(t)\| \forall t \in [0, T]$  and some  $k_h < \infty \in \mathbb{R}$  ( $h \in \{f, B, g_x, g_t\}$ ).

A4: The operators  $B(\cdot, \cdot)$  and  $g_x(\cdot, \cdot)$  are bounded on  $\mathbb{R}^n \times [0, T]$ .

A5: All functions are assumed to be measurable and integrable.

Assumption A1 implies that given an achievable, desired output trajectory ( $y_d$ ) and initial state ( $x_d(0)$ ), there exist unique input ( $u_d$ ) and state ( $x_d$ ) trajectories corresponding to this output trajectory. Assumption A4 on  $g_x(\cdot, \cdot)$  implies that  $g$  is uniformly globally Lipschitz in  $x$  on  $[0, T]$ .

The function  $w_i(t)$  represents both deterministic and random disturbances of the system; it may be stiction, nonreproducible friction, modeling errors, etc. This is important to include since these are present in physical systems. Assumption A2 restricts these disturbances to be bounded, but they may be discontinuous (e.g., stiction in mechanical systems).

We consider the update law given by

$$u_{i+1}(t) = (1 - \gamma)u_i(t) + \gamma u_0(t) + L(y_i(t), t)[\dot{y}_d(t) - \dot{y}_i(t)] \quad 0 \leq \gamma < 1 \quad (2.2)$$

where  $L: \mathbb{R}^m \times [0, T] \rightarrow \mathbb{R}^{r \times m}$  is bounded. This learning operator is similar to those used in [2], [6], [8]–[10] in that it updates the system input in an *affine* fashion.

The use of  $\gamma$  will be discussed in Section II. For an initial reading  $\gamma$  is best ignored (i.e.,  $\gamma = 0$ ) since its inclusion does not alter the proof in any essential way.

The following norm is used to simplify the expression of our results.

**Definition 2.3:** We define the  $\lambda$  norm for a function  $h: [0, T] \rightarrow \mathbb{R}^k$  by

$$\|h(\cdot)\|_\lambda \triangleq \sup_{t \in [0, T]} e^{-\lambda t} \|h(t)\|. \quad (2.3)$$

**Remarks:** From this definition we can see that  $\|h\|_\lambda \leq \|h\|_\infty \leq e^{\lambda T} \|h\|_\lambda$  for  $\lambda > 0$  (where  $\|h\|_\infty \triangleq \sup_{t \in [0, T]} \|h(t)\|$ ), implying that these two norms are equivalent.

For clarification of the remaining discussion, function parameters will be shown in subscript notation with the dependence on time implied unless otherwise stated. In particular

$$\begin{aligned} g_{xi} &\triangleq \frac{\partial}{\partial x} g(x, t) \Big|_{x=x_i(t)}, & g_{xd} &\triangleq \frac{\partial}{\partial x} g(x, t) \Big|_{x=x_d(t)}, \\ g_{ti} &\triangleq \frac{\partial}{\partial t} g(x, t) \Big|_{x=x_i(t)}, & g_{td} &= \frac{\partial}{\partial t} g(x, t) \Big|_{x=x_d(t)}, \\ f_i &\triangleq f(x_i(t), t), & f_d &\triangleq f(x_d(t), t), \\ u_i &\triangleq u_i(t), & u_d &\triangleq u_d(t), \\ w_i &\triangleq w_i(t), & B_i &\triangleq B(x_i(t), t), \\ B_d &\triangleq B(x_d(t), t), & L_i &\triangleq L(y_i(t), t) \end{aligned}$$

and  $k_{g_x}$ ,  $k_{g_t}$ ,  $k_f$ ,  $k_B$ ,  $k_g$  are the Lipschitz constants for  $g_x(\cdot, \cdot)$ ,  $g_t(\cdot, \cdot)$ ,  $f(\cdot, \cdot)$ ,  $B(\cdot, \cdot)$ , and  $g(\cdot, \cdot)$ , respectively. We now state the main result.

**Theorem 2.4:** Let the system described by (2.1) satisfy assumptions A1–A5 and use the update law (2.2). Given a desired output trajectory  $y_d(\cdot)$  and an initial state  $x_d(0)$ , which are achievable,<sup>1</sup> if

$$\|(1 - \gamma)I - L(g(x, t), t)g_x(x, t)B(x, t)\| \leq \rho < 1 \quad \forall (x, t) \in \mathbb{R}^n \times [0, T]$$

and the initial state error is bounded ( $\|x_d(0) - x_i(0)\| \leq b_{x_0}$ ), then as  $i \rightarrow \infty$  the error between  $u_i$  and  $u_d$  is bounded. In addition, the state and output asymptotic errors are bounded. These bounds depend continuously on the bound on the initial state error, bound on the state disturbance, and  $\gamma$ ; as  $b_{x_0}$ ,  $b_w$ , and  $\gamma$  tend to zero, these bounds also tend to zero.

The proof's main idea is to show that  $\|\delta u_{i+1}\|_\lambda \leq \bar{\rho} \|\delta u_i\|_\lambda + \epsilon$  where  $0 \leq \bar{\rho} < 1$  ( $\delta u_i \triangleq u_d - u_i$ ). This implies that  $\limsup_{i \rightarrow \infty} \|\delta u_i\|_\lambda \leq \frac{1}{1 - \bar{\rho}} \epsilon$ . Hence, the majority of the proof is a calculation to show that this relationship holds. The results follow quickly once we have established this. Intuitively, the condition on  $L$  says that if we push on the system (through  $u$ ), we can observe a change in the output, and we take an appropriate action to reduce the error.

**Proof:** From (2.1) and the update law (2.2) the error for the iterate  $i + 1$  can be written as

$$u_d - u_{i+1} = u_d - (1 - \gamma)u_i - \gamma u_0 - L_i[\dot{y}_d - \dot{y}_i] \quad (2.5)$$

$$\begin{aligned} &= (1 - \gamma)(u_d - u_i) + \gamma(u_d - u_0) \\ &\quad - L_i \begin{bmatrix} g_{xd}(f_d + B_d u_d) + g_{td} \\ -g_{xi}(f_i + B_i u_i + w_i) - g_{ti} \end{bmatrix} \end{aligned} \quad (2.6)$$

$$\begin{aligned} &= [(1 - \gamma)I - L_i g_{xi} B_i](u_d - u_i) + \gamma(u_d - u_0) \\ &\quad - L_i \begin{bmatrix} g_{xd}(f_d + B_d u_d) + g_{td} \\ -g_{xi}(f_i + B_i u_d + w_i) - g_{ti} \end{bmatrix}. \end{aligned} \quad (2.7)$$

Taking norms, using the bounds, and using the Lipschitz conditions yields

$$\begin{aligned} \|u_d - u_{i+1}\| &\leq \|(1 - \gamma)I - L_i g_{xi} B_i\| \|u_d - u_i\| \\ &\quad + \gamma \|u_d - u_0\| \\ &\quad + \|L_i\| \left\{ \begin{aligned} &\|g_{xd} - g_{xi}\| \|f_d B_d u_d\| \\ &+ \|g_{xi}\| [\|f_d - f_i\| + \|B_d - B_i\| \|u_d\| + \|w_i\|] \\ &+ \|g_{td} - g_{ti}\| \end{aligned} \right\} \end{aligned} \quad (2.8)$$

$$\begin{aligned} &\leq \rho \|u_d - u_i\| + \gamma \|u_d - u_0\| \\ &\quad + b_L \left\{ \begin{aligned} &k_{g_x} \|x_d - x_i\| b_d \\ &+ b_{g_x} [k_f \|x_d - x_i\| + k_B \|x_d - x_i\| b_{ud} + b_w] \\ &+ k_{g_t} \|x_d - x_i\| \end{aligned} \right\}, \end{aligned} \quad (2.9)$$

where  $b_x$ ,  $b_{g_x}$  are the norm bounds for  $L(\cdot, \cdot)$ ,  $g_x(\cdot, \cdot)$ , respectively, and

$$b_d \triangleq \sup_{t \in [0, T]} \|f_d + B_d u_d\| \text{ and } b_{ud} \triangleq \sup_{t \in [0, T]} \|u_d\|.$$

Defining  $k_1 \triangleq b_L [k_{g_x} b_d + b_{g_x} (k_f + k_B b_{ud}) + k_{g_t}]$ , (2.9) sim-

<sup>1</sup> By achievable we mean that the desired input trajectory  $u_d(\cdot)$  exists and is bounded.

plifies to

$$\|u_d - u_{i+1}\| \leq \rho \|u_d - u_i\| + k_1 \|x_d - x_i\| + b_L b_{g_x} b_w + \gamma \|u_d - u_0\|. \quad (2.10)$$

Now writing the integral expression for  $x(t)$  and taking norms we obtain

$$\|x_d - x_i\| = \|x_d(0) - x_i(0)\| \quad (2.11)$$

$$\begin{aligned} & + \int_0^t ((f_d + B_d u_d) - (f_i + B_i u_i + w_i)) d\tau \\ & \leq \|x_d(0) - x_i(0)\| \\ & + \int_0^t \left( \|f_d - f_i\| + \|B_d - B_i\| \|u_d\| \right. \\ & \quad \left. + \|B_i\| \|u_d - u_i\| + \|w_i\| \right) d\tau \quad (2.12) \end{aligned}$$

$$\begin{aligned} & \leq \|x_d(0) - x_i(0)\| \\ & + \int_0^t \left( (k_f + k_B b_{ud}) \|x_d - x_i\| \right. \\ & \quad \left. + b_B \|u_d - u_i\| + b_w \right) d\tau \quad (2.13) \end{aligned}$$

where  $b_B$  is the norm bound on  $B(\cdot, \cdot)$ . Defining  $k_3 \triangleq (k_f + k_B b_{ud})$ ,  $\|\delta x_i\| \triangleq \|x_d - x_i\|$ , and using a basic integral inequality (see [11, p. 96]) we have that

$$\begin{aligned} \|\delta x_i\| & \leq \|x_d(0) - x_i(0)\| e^{k_3 t} \\ & + \int_0^t e^{k_3(t-\tau)} (b_B \|u_d(\tau) - u_i(\tau)\| + b_w) d\tau. \quad (2.14) \end{aligned}$$

Combining (2.10) and (2.14) yields

$$\begin{aligned} \|\delta u_{i+1}\| & \leq \rho \|\delta u_i\| \\ & + k_1 \left\{ \int_0^t e^{k_3(t-\tau)} (b_B \|\delta u_i(\tau)\| + b_w) d\tau \right\} \\ & + b_L b_{g_x} b_w + \gamma \|\delta u_0\| \quad (2.15) \end{aligned}$$

$$\begin{aligned} & \leq \rho \|\delta u_i\| + k_1 b_B \int_0^t e^{k_3(t-\tau)} \|\delta u_i(\tau)\| d\tau \\ & + k_1 \|\delta x_i(0)\| e^{k_3 t} + k_1 b_w \int_0^t e^{k_3(t-\tau)} d\tau \\ & + b_L b_{g_x} b_w + \gamma \|\delta u_0\|. \quad (2.16) \end{aligned}$$

Multiplying (2.16) by  $e^{-\lambda t}$ , defining  $k \triangleq \max \{k_1 b_B, k_3\}$ , and assuming  $\lambda > k$  we have that

$$\begin{aligned} e^{-\lambda t} \|\delta u_{i+1}\| & \leq \rho e^{-\lambda t} \|\delta u_i\| + k \int_0^t e^{-\lambda \tau} \|\delta u_i(\tau)\| e^{(k-\lambda)(t-\tau)} d\tau \\ & + k_1 \|\delta x_i(0)\| e^{(k_3-\lambda)t} \\ & + k_1 b_w \int_0^t e^{-\lambda \tau} e^{(k_3-\lambda)(t-\tau)} d\tau \\ & + b_L b_{g_x} b_w e^{-\lambda t} + \gamma e^{-\lambda t} \|\delta u_0\|. \quad (2.17) \end{aligned}$$

Noticing that the integrals are strictly increasing and that for a constant  $\|k\|_\lambda = k$  we obtain

$$\begin{aligned} \|\delta u_{i+1}\|_\lambda & \leq \left[ \rho + \frac{k}{\lambda - k} (1 - e^{(k-\lambda)T}) \right] \|\delta u_i\|_\lambda \\ & + k_1 \|\delta x_i(0)\| + \frac{k_1 b_w}{\lambda - k_3} (1 - e^{(k_3-\lambda)T}) \\ & + b_L b_{g_x} b_w + \gamma \|\delta u_0\|_\lambda. \quad (2.18) \end{aligned}$$

Defining  $\bar{\rho} = \rho + \frac{k}{\lambda - k} (1 - e^{(k-\lambda)T})$  and  $k_4 \triangleq b_L b_{g_x} + \frac{k_1}{\lambda - k_3} (1 - e^{(k_3-\lambda)T})$ , we have that

$$\begin{aligned} \|\delta u_{i+1}\|_\lambda & \leq \bar{\rho} \|\delta u_i\|_\lambda + k_1 \|\delta x_i(0)\| \\ & + k_4 b_w + \gamma \|\delta u_0\|_\lambda \quad (2.19) \end{aligned}$$

$$\|\delta u_{i+1}\|_\lambda \leq \bar{\rho} \|\delta u_i\|_\lambda + \epsilon. \quad (2.20)$$

Where  $\epsilon$  combines the norm bounds of the initial state errors, state disturbances, and bias contribution. Since  $\rho < 1$ , we can find a  $\lambda > k$  which makes  $\bar{\rho} < 1$ . Thus,  $u_i$  converges to the neighborhood of  $u_d$  of radius  $(\frac{1}{1-\bar{\rho}})\epsilon$  with respect to the  $\lambda$  norm.

Implying that

$$\limsup_{i \rightarrow \infty} \|\delta u_i\|_\lambda \leq \left( \frac{1}{1-\bar{\rho}} \right) \epsilon. \quad (2.21)$$

Using (2.14), and similar manipulations we obtain

$$\|\delta x_i\|_\lambda \leq \|\delta x_i(0)\| + \int_0^t e^{(k_3-\lambda)(t-\tau)} \|\delta u_i\|_\lambda d\tau \quad (2.22)$$

$$\leq \|\delta x_i(0)\| + \frac{1}{\lambda - k_3} (1 - e^{(k_3-\lambda)T}) \|\delta u_i\|_\lambda. \quad (2.23)$$

$$\text{So } \limsup_{i \rightarrow \infty} \|\delta x_i\|_\lambda \leq \|\delta x_i(0)\| + \frac{1}{\lambda - k_3} (1 - e^{(k_3-\lambda)T}) \left( \frac{1}{1-\bar{\rho}} \right) \epsilon.$$

To obtain the result for  $y_i$  we use the fact that  $g$  is Lipschitz in  $x$ . ■

Equation (2.19) clearly illustrates the influence of the initial state error, state disturbance, and bias term in degrading our bound on the asymptotic errors. We see that this bound on the degradation is continuous in these factors. Furthermore, in the absence of these terms  $\epsilon = 0$  implying convergence of the algorithm to the desired trajectories.

It is also possible to include terms in the update law depending on the position error and its integral, with the same results as Theorem 2.4; the details of this are shown in [1].

### III. DISCUSSION

The learning update law presented in Section II guarantees learning, provided certain conditions are met. In practice, however, there are modifications that improve the performance of the algorithm.

If at any time the update law produces an input that we know to be out of range, and the allowable set of inputs is a convex set, then projecting back into this set will improve performance. This is evident from our analysis since this will further decrease  $\|u_d - u_{i+j}\|$ . For many robot manipulators the allowable set of joint torques is a hypercube and projection is easily implemented.

Theorem 2.4 implies that as the iteration number approaches  $\infty$  the trajectory errors are less than certain bounds. In practice, however, we desire to stop the process in a finite time, and we desire the error to be as small as possible at this time. In this situation the bias term may be helpful, and varying the update operator as the iterations progress may further improve performance. The bias term is initially useful to keep the input from wandering excessively, but with time we want to decrease its influence by decreasing  $\gamma$ . Once the input has converged fairly well we may want to begin decreasing the learning gain (the size of  $L$ ) to cause the input to average out random disturbances—improving the accuracy of the input that we choose after a finite time. It is easily

seen that these modifications do not change the results of Theorem 2.4, provided that the condition on the update law is satisfied for all  $L_i$  and  $\gamma_i$ .

The class of systems considered is fairly general, and a closer examination of the results reveals that what is essential is that when we apply an input to the system, we can observe a corresponding output, and we act upon this output with the learning operator. The stability of the system may affect the convergence rate, but not the actual convergence of the learning algorithm.

It is important to remember that learning control is not a form of dynamic feedback. It cannot be used to stabilize a system nor to change its performance for a general trajectory. Therefore, in applications it is desirable to use a robust feedback controller to improve the system performance, and as explained earlier, this is the motivation for considering time-varying systems. Learning control iteratively updates a feed-forward term to provide a finer and finer "open loop" performance along a specific trajectory—it is not intended to make up for a poor feedback controller design.

In conclusion, we believe that the learning algorithm presented is applicable to a wide variety of problems. The stability of learning in the presence of disturbances and initial condition errors allows us to use the learning algorithm with confidence in applications. We further conjecture that these results can be extended to other update laws, allowing the differentiation to be replaced by say a lead filter with better noise response; this constitutes an interesting area for future research.

#### ACKNOWLEDGMENT

The authors would like to express their thanks to R. Minnichelli, R. Murray, and P. Jacobs for their help in the form of discussions and suggestions.

#### REFERENCES

- [1] G. Heinzinger, D. Fenwick, B. Paden, and F. Miyazaki, "Robust learning control," in *Proc. 28th IEEE Conf. Decision Contr.*, Tampa, FL, Dec. 1989.
- [2] S. Arimoto, S. Kawamura, and F. Miyazaki, "Bettering operation of robots by learning," *J. Robot. Syst.*, vol. 1, no 2, pp. 440–447, 1984.
- [3] J. J. Craig, "Adaptive control of manipulators through repeated trials," in *Proc. Amer. Contr. Conf.*, San Diego, CA, June 1984, pp. 1566–1573.
- [4] P. Bondi, G. Casalino, and L. Gambardella, "On the iterative learning control theory for robotic manipulators," *IEEE J. Robot. Automation*, vol. 4, pp. 14–22, Feb. 1988.
- [5] C. G. Atkeson and J. McIntyre, "Robot trajectory learning through practice," in *Proc. IEEE Conf. Robot. Automation*, San Francisco, CA, Apr. 1986, pp. 1737–1742.
- [6] J. Hauser, "Learning control for a class of nonlinear systems," in *Proc. 26th IEEE Conf. Decision Contr.*, Los Angeles, CA, Dec. 1987, pp. 859–860.
- [7] T. Sugie and T. Ono, "On an iterative learning control law for dynamical systems," in *Proc. IFAC 10th World Conf.*, Munich, Germany, June 1987, pp. 339–344.
- [8] S. Arimoto, S. Kawamura, and F. Miyazaki, "Convergence, stability, and robustness of learning control schemes for robot manipulators," in *Proc. Int. Symp. Robot Manipulators; Modeling, Contr., Education*, Albuquerque, NM, Nov. 1986.
- [9] —, "Mathematical theory of learning with applications to robot control," in *Adaptive and Learning Systems*, K. S. Narendra, Ed., New York: Plenum 1986.
- [10] M. Togai and O. Yamano, "Learning control and its optimality: Analysis and its application to controlling industrial robots," in *Proc. 24th IEEE Conf. Decision Cont.* Fort Lauderdale, FL, Dec. 1985, pp. 248–253.
- [11] T. M. Flett, *Differential Analysis*. Cambridge: Cambridge University Press, 1980.

## Discrete-Time Filtering for Linear Systems with Non-Gaussian Initial Conditions: Asymptotic Behavior of the Difference Between the MMSE and LMSE Estimates

Richard B. Sowers and Armand M. Makowski

**Abstract**—We consider the one-step prediction problem for discrete-time linear systems in correlated plant and observation Gaussian white noises, with non-Gaussian initial conditions. We investigate the large time asymptotics of  $\epsilon_t$ , the expected squared difference between the MMSE and LMSE (or Kalman) estimates of the state at time  $t$  given past observations. We characterize the limit of the error sequence  $\{\epsilon_t, t = 0, 1, \dots\}$  and obtain some related rates of convergence; a complete analysis is provided for the scalar case. The discussion is based on explicit representations for the MMSE and LMSE estimates, recently obtained by the authors, which display the dependence of these quantities on the initial distribution.

#### I. INTRODUCTION

Consider the time-invariant linear discrete-time stochastic system

$$\begin{aligned} X_0^o &= \xi, \quad X_{t+1}^o = AX_t^o + W_{t+1}^o \\ Y_t &= HX_t^o + V_{t+1}^o \quad t = 0, 1, \dots, \end{aligned} \quad (1.1)$$

where the matrices  $A$  and  $H$  are of dimension  $n \times n$  and  $n \times k$ , respectively. This system is defined on some underlying probability triple  $(\Omega, \mathcal{F}, P)$  which carries all the random elements considered in this note, namely the  $\mathbb{R}^n$ -valued plant process  $\{X_t^o, t = 0, 1, \dots\}$ , the  $\mathbb{R}^k$ -valued observation process  $\{Y_t, t = 0, 1, \dots\}$  and the  $\mathbb{R}^{n+k}$ -valued noise process  $\{(W_{t+1}^o, V_{t+1}^o), t = 0, 1, \dots\}$ . Throughout this note we make assumptions A.1–A.3, where:

A.1: the process  $\{(W_{t+1}^o, V_{t+1}^o), t = 0, 1, \dots\}$  is a stationary zero-mean  $\mathbb{R}^{n+k}$ -valued Gaussian white noise sequence [2, p. 22] with covariance structure  $\Gamma$  given by

$$\Gamma := \text{Cov} \begin{pmatrix} W_{t+1}^o \\ V_{t+1}^o \end{pmatrix} = \begin{pmatrix} \Gamma_w & \Gamma_{wv} \\ \Gamma_{vw} & \Gamma_v \end{pmatrix} \quad t = 0, 1, \dots; \quad (1.2)$$

A.2: the initial state  $\xi$  has distribution  $F$  with finite first and second moments  $\mu$  and  $\Delta$ , respectively, and is independent of the noise process  $\{(W_{t+1}^o, V_{t+1}^o), t = 0, 1, \dots\}$ ; and

A.3: the covariance matrices  $\Gamma_w$  and  $\Delta$  are positive definite, thus invertible.

For each  $t = 0, 1, \dots$ , we form the conditional mean  $\hat{X}_{t+1}^o := E[X_{t+1}^o | Y_0, Y_1, \dots, Y_t]$  or MMSE estimate of  $X_{t+1}^o$  on the basis of  $\{Y_0, Y_1, \dots, Y_t\}$ . In general,  $\hat{X}_{t+1}^o$  is a nonlinear function of  $\{Y_0, Y_1, \dots, Y_t\}$ , in contrast to the corresponding LMSE or Kalman estimate of  $X_{t+1}^o$  which is by definition linear, and which we denote by  $\hat{X}_{t+1}^K$ . We then calculate  $\epsilon_{t+1} := E[\|\hat{X}_{t+1}^o - \hat{X}_{t+1}^K\|^2]$ <sup>1</sup>

Manuscript received March 30, 1989; revised July 9, 1990 and August 3, 1990. Paper recommended by Past Associate Editor, D. H. Owens. The work of R. B. Sowers was supported by an ONR Graduate Fellowship. The work of A. M. Makowski was supported in part by the Engineering Research Centers Programs NSF CDR 88-03012 and in part by the National Science Foundation under Grant ECS 83-51836.

R. B. Sowers was with the Department of Electrical Engineering and the Systems Research Center, University of Maryland, College Park, MD 20742. He is now with the Center for Applied Mathematical Sciences, Department of Mathematics, University of Southern California, Los Angeles, CA 90089-1113.

A. M. Makowski is with the Department of Electrical Engineering and the Systems Research Center, University of Maryland, College Park, MD 20742.

IEEE Log Number 9103082.

<sup>1</sup>Here  $\|v\|$  denotes the Euclidean norm of the vector  $v$  in  $\mathbb{R}^n$ .