

# Control of Smart Exercise Machines—Part II: Self-Optimizing Control

Perry Y. Li, *Member, IEEE*, and Roberto Horowitz, *Member, IEEE*

**Abstract**—This is the second part of a two-part paper on the design of an intelligent controller for a class of exercise machines. The control objective is to cause the user to exercise in a manner that optimizes a criterion related to the user's mechanical power. The optimal exercise strategy is determined by an *a priori* unknown biomechanical behavior, called the Hill surface, of the individual user. Consequently, the control scheme must simultaneously: 1) identify the user's biomechanical behavior; 2) optimize the controller; and 3) stabilize the system to the estimated optimal states. In Part I of this paper, a dynamic damping controller was proposed which satisfies the safety requirement and is capable of causing the user to execute an arbitrary exercise strategy if the user's biomechanical behavior is known. In this second part of the paper, we address the self-optimization problem in which both the determination and the eventual execution of the optimal exercise strategy are accomplished, when the user's biomechanical behavior is unknown. This is achieved by a combination of an adaptive controller and a reference generator. The latter switches the desired exercise strategy between a training strategy and the estimated optimal strategy. Depending on the switching scheme chosen, it is shown that, asymptotically, the user will either execute the optimal exercise with probability one or operate close to it. Experimental results of the overall system verify the efficacy of the design.

**Index Terms**—Adaptive control, biomechanics, hybrid systems, intelligent control, passivity, robotics, self optimization, velocity field control.

## I. INTRODUCTION

THIS PAPER and the companion paper [1] are concerned with the formulation, analysis, and implementation of intelligent exercise machine control systems to enable individual users to exercise optimally. In this control problem, which is formulated in [1], the biomechanical behavior of the user with respect to the exercise machine is modeled by a *Hill surface* relating the force that the user can exert to the position and velocity of the exercising motion. Each exercise strategy can, therefore, be encoded by a velocity field on the configuration space of the machine and corresponds to a trajectory on the Hill surface of the user. The control objective is to manipulate the assistive/resistive force on the exercise machine, so that the user exercises according to the optimal velocity field specific to the user's Hill surface. A typical optimal exercise would maximize the user's mechanical power

output throughout the exercise motion. To ensure that the exercise machine is safe to operate, the closed-loop control system is also constrained to interact passively with the user. To this end, a dynamic damping controller was proposed in [1] which satisfies this closed-loop passivity requirement and, if the user's Hill surface is known, is capable of causing the user to execute an arbitrarily *specified* velocity field. In particular, if the optimal velocity field is specified, then the user would perform the optimal exercise.

Unfortunately, the Hill surface is generally unknown *a priori*, since it is both user specific and varies with the user's motivation and fatigue state during the exercise session. Thus, the dynamic damping controller in [1] is inadequate in causing the user to execute the optimal velocity field, nor can the optimal velocity field be specified in the first place. The present paper addresses these inadequacies by developing an adaptive version of the dynamic damping controller and by designing a reference velocity field generator, so that the optimal velocity field can be ultimately specified.

The problem at hand belongs to the class of "self-optimizing control" problems [2] in which a plant with some *a priori* unknown parameters is required to perform an optimal task with respect to a performance index. In our present problem, the user must execute the optimal velocity field specific to his/her Hill surface, despite the fact that the Hill surface is unknown *a priori*. The main difference between a self-optimizing control problem and a typical adaptive or learning control problem lies in the fact that, since the optimal task generally depends on the unknown plant parameters, it cannot be explicitly specified beforehand and has to be determined online. Indeed, without knowledge of the plant, it may not even be possible to decide if the plant is, in fact, operating optimally. In contrast, in a typical adaptive or learning control problem (e.g., [3]–[6] and others), the task (such as following a given trajectory) is explicitly specified *a priori*, and the objective is to perform this task with good precision in the presence of parametric and nonparametric uncertainties in the plant and environment dynamics. If the optimal task can be specified explicitly, then one may develop an adaptive controller to enable that task to be performed. Typically, the estimates of the plant parameters do not even have to converge to the true ones. In fact, vanishing of parameter estimation error seldom occurs in direct adaptive control schemes.

The basic difficulty in the present problem is that, in order to determine the optimal velocity field, one has to have an adequate knowledge of the user's Hill surface. This requires the user to exercise in a manner that explores the different

Manuscript received February 22, 1997; revised September 2, 1997. Recommended by Guest Editor H. Kobayashi.

P. Y. Li is with the Department of Mechanical Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: pli@me.umn.edu).

R. Horowitz is with the Department of Mechanical Engineering, University of California, Berkeley, CA 94720-1740 USA (e-mail: horowitz@me.berkeley.edu).

Publisher Item Identifier S 1083-4435(97)09031-5.

regions of the Hill surface. In other words, the exercise must be sufficiently exciting. On the other hand, the optimal exercise which we would like the user to perform is not sufficiently exciting. Hence, a conflict exists between the objectives of exercising optimally and of identification. This conflict is the essence of the *dual control* problem posed in [7]. This difficulty is illustrated by a result in adaptive linear quadratic (LQ) control which states that the set of controller parameters which induce optimality is a thin set in the set of all possible convergence points [8]. Hence, it is not likely that the controller will converge to the optimal controller. The (static) extremum control problem, where the optimal output is to be achieved for an unknown plant, was considered in [9]. In [9], the estimated optimal input is updated on-line and is superimposed with a dither signal before it is applied. Because of the dither, the input signal never really optimizes the objective criterion. Intuitively, if changes of the plant characteristics or the badness of the parameter estimates can be detected, the dithering or the excitation signal can be applied only at those instances. A similar idea in adaptive control can be found in [10] and [11], where the strength of an excitation signal is controlled by an internal signal which vanishes as the control objective is achieved.

In our approach, instead of modifying the strength of the excitation, we require the user to perform two types of velocity fields, the “training” velocity field and the estimated optimal velocity field. These are time multiplexed together, and the relative frequency of their application is modulated in such a manner that the “training” velocity field will be infrequently applied as the Hill surface becomes well known. To implement this idea, a control structure which is both hierarchical and hybrid (i.e., discrete and continuous dynamics interact) is proposed. At the lower level, a *continuous-state* adaptive version of the dynamic damping controller in [1] is constructed which is capable of causing the user to execute any specified velocity field, despite uncertainty in the Hill surface. At the higher level, a *discrete-state* supervisor generates a reference velocity field by switching between the “training” velocity field and the estimated optimal velocity field. The decision to switch between them is made discretely, based on an *optimality error* signal which monitors the convergence of the parameter estimates, as well as the instances when the user’s Hill surface may have varied. The discrete supervisor can either be *deterministic* or *probabilistic*. For the deterministic case, the exercise velocity converges after, at most, a finite number of training phases to a velocity field close to the optimal one, with assignable closeness. In the probabilistic case, it is shown that the velocity converges asymptotically to the optimal velocity field with almost certainty.

The remainder of this paper is organized as follows. In Section II, we briefly review the exercise machine control problem and present additional preliminary background necessary to develop the self-optimizing control system. In Section III, an adaptive version of the dynamic damping controller is presented. In Section IV, we describe the reference generator and the combined result of the adaptive controller and the reference generator. Experimental results are given in Section V. Section VI contains concluding remarks.

## II. BACKGROUND

### A. Control Problem Review

The single-degree-of-freedom experimental exercise machine used in our research was described in [1]. It consists of a rigid link connected to a dc motor at one end and is attached to a handle at the other. The user exercises by rotating the link while holding the handle. The motor provides for the resistive or assistive force to cause the user to execute the desired exercise strategy. The configuration space of this system is the circle  $\mathcal{G} = [0, 2\pi)$ , so that  $x \in \mathcal{G}$  is the angle of the rigid link with respect to a reference angle. The dynamics of this setup are given by

$$M(x(t))\ddot{x}(t) + C(x(t))\dot{x}^2(t) = F(t) + T(t) \quad (1)$$

where  $F(t)$  is the generalized force (torque) generated by the user, and  $T(t)$  is the torque generated by the motor, which is the manipulated control variable.  $M(x) \in \mathbb{R}^+$  is the generalized inertia and  $C(x)\dot{x}^2$  is the Coriolis and centripetal force, and they are related by  $(d/dx)M(x) = 2C(x)$ .

We assume that the Hill surface which relates the user’s force  $F(t)$  to the position and the velocity of the exercise motion, when the effort and the fatigue state of the user are constant, is given by

$$F_h(x, \dot{x}) := a(x) - b(x)\dot{x}; \quad a(x), b(x) > 0. \quad (2)$$

Notice that the Hill surface is affine in, and decreases monotonically with, the velocity of motion.  $a(x)$  and  $b(x)$  are arbitrary continuous functions of  $x$  and are assumed to be unknown *a priori*.

The overall objective is to enable the user to exercise in a manner that maximizes, at all times, the weighted human power:

$$J_\rho(F, \dot{x}) = F\dot{x}^\rho \quad \text{with } \rho > 0 \quad (3)$$

subject to the constraint that the force, position and velocity are related by the Hill surface, i.e.,

$$F(t) = F_h(x(t), \dot{x}(t)).$$

Therefore, the user must exercise according to the *optimal* desired velocity field given by

$$V^*(x) = \frac{\rho}{\rho + 1} \frac{a(x)}{b(x)}. \quad (4)$$

In other words, the motor torque  $T(t)$  in (1) must be manipulated so that  $\dot{x}(t) \rightarrow V^*(x(t))$ .

### B. Linear Parameterization of the Unknown Functions

In order to estimate the unknown functions  $a(x)$  and  $b(x)$  of the Hill surface in (2), they are linearly parameterized via an integral representation of the first kind.

A sufficiently smooth function  $k: \mathcal{G} \times \mathcal{G} \rightarrow \mathbb{R}$  is a symmetric kernel with a finite eigen function expansion if  $k(x, \sigma) =$

$k(\sigma, x) \forall x, \sigma \in \mathcal{G}$  and if there exist  $\psi_1(\cdot), \dots, \psi_N(\cdot): \mathcal{G} \rightarrow \mathfrak{R}$ , and  $\zeta_1, \dots, \zeta_N \in \mathfrak{R}$  such that

$$\begin{aligned} \int_{\mathcal{G}} k(x, \sigma) \psi_i(\sigma) d\sigma &= \zeta_i \psi_i(x) \\ \int_{\mathcal{G}} \Psi^T(\sigma) \Psi(\sigma) d\sigma &= I_N \\ k(x, \sigma) &= \Psi(x) \mathbf{Z} \Psi^T(\sigma) \\ \mathbf{Z} &= \text{diag}(\zeta_1, \dots, \zeta_N) \end{aligned} \quad (5)$$

where  $\Psi(\cdot) = [\psi_1(\cdot), \dots, \psi_N(\cdot)]$ .

*Assumption 1:* For a given effort level and fatigue state of the user, the human force is given by the Hill surface in (2), i.e.,  $F(t) = F_h(x(t), \dot{x}(t))$ , with  $a(x) \in [\underline{a}(x), \infty)$  and  $b(x) \in [\underline{b}(x), \bar{b}(x)]$ , where  $\underline{a}(x), \underline{b}(x), \bar{b}(x) > 0$ .

Moreover,  $a, b: \mathcal{G} \rightarrow \mathfrak{R}^+$  can be represented by an integral equation of the first kind [12]:

$$\begin{bmatrix} a(x) \\ b(x) \end{bmatrix} = \int_{\mathcal{G}} k(x, \sigma) \begin{bmatrix} c_a(\sigma) \\ c_b(\sigma) \end{bmatrix} d\sigma \quad (6)$$

where  $c_a(\cdot)$  and  $c_b(\cdot)$  are squared integrable, and  $k: \mathcal{G} \times \mathcal{G} \rightarrow \mathfrak{R}$  is a symmetric kernel with a finite eigen function expansion. Furthermore, the vector function  $\mathbf{c}: \mathcal{G} \rightarrow \mathfrak{R}^2$ ,  $\mathbf{c}(x) := [c_a(x), c_b(x)]^T$  lies in a convex set  $\mathcal{P}_c$ , so that  $a(x) \in [\underline{a}(x), \infty)$  and  $b(x) \in [\underline{b}(x), \bar{b}(x)]$ .

The functions  $k(\cdot, \cdot)$  and  $c_a(\cdot)$  and  $c_b(\cdot)$  are referred to as the kernel and the influence functions, respectively. Typically,  $k(\cdot, \cdot)$  is assumed to be known and  $c_a(\cdot), c_b(\cdot)$  are assumed to be unknown and have to be identified. Integral representations have been found to be useful in functional adaptive and learning control [6], since they enable a wide class of nonlinear functions to be parameterized linearly, while ensuring that the estimates of these functions remain smooth.

The unknown Hill surface  $F_h(x, \dot{x})$  can now be linearly parameterized in terms of a known regressor and an unknown parameter. Define the functional regressors and the corresponding vector regressor as follows:

$$\phi^c(t, \sigma) := k(x(t), \sigma) [1, -\dot{x}(t)]^T \quad \forall \sigma \in \mathcal{G} \quad (7)$$

$$\boldsymbol{\phi}(t) := \int_{\mathcal{G}} \phi^c(t, \sigma) \otimes \Psi(\sigma) d\sigma \in \mathfrak{R}^{2N} \quad (8)$$

where  $\Psi(\sigma)$  is the collection of eigen functions in (5) and  $\otimes$  denotes the Kronecker product. Define the parameter vectors by

$$\Theta := \begin{pmatrix} \boldsymbol{\theta}_a \\ \boldsymbol{\theta}_b \end{pmatrix} = \begin{bmatrix} \int_{\mathcal{G}} \Psi^T(\sigma) c_a(\sigma) d\sigma \\ \int_{\mathcal{G}} \Psi^T(\sigma) c_b(\sigma) d\sigma \end{bmatrix}. \quad (9)$$

Since  $\mathbf{c}(\sigma)$  is constrained to lie in the convex set  $\mathcal{P}_c$ , there is a related convex set, denoted by  $\mathcal{P}_\theta \subset \mathfrak{R}^{2N}$ , such that  $\Theta \in \mathcal{P}_\theta$ .

Therefore, if  $F(t) = F_h(x(t), \dot{x}(t))$ , then

$$F(t) = \int_{\mathcal{G}} \phi^c(t, \sigma)^T \mathbf{c}(\sigma) d\sigma = \boldsymbol{\phi}(t)^T \Theta. \quad (10)$$

In (10), the functional and vector regressors  $\phi^c(t, \sigma)$  and  $\boldsymbol{\phi}(t)$  are assumed to be known, whereas the influence function

vector  $\mathbf{c}(\sigma)$  and the parameter vector  $\Theta$  are unknown *a priori*, since they depend on the effort and fatigue state of the user and must be identified. In actual implementation of the estimation algorithm to be presented, it is more convenient to use the integral representation directly (i.e., the influence function vector  $\mathbf{c}(\sigma)$  is directly estimated). However, the equivalent finite-dimensional vector parameterization (i.e., in terms of  $\Theta \in \mathfrak{R}^{2N}$ ) is more convenient in presenting the analysis. We use only the more familiar finite dimensional vector parameterization in the remainder of the paper.

### C. Force Observer

The human force  $F(t)$  in (1) cannot generally be measured, even with a force sensor, unless invasive techniques are used. This is because  $F(t)$  is the generalized force that actuates all the inertia components in the system, including the limbs of the user. For the estimation of the unknown parameter vector  $\Theta$  in (9), we shall utilize a force observer to obtain  $\mathcal{F}(t)$ , which is the stable filtered output of  $F(t)$ , using only position and velocity measurements

$$\mathcal{F}(t) = \frac{\lambda}{\lambda + s} F(t) \quad \lambda > 0. \quad (11)$$

Since  $T(t)$ ,  $M(x)$ , and  $C(x)$  are known quantities, it can be verified from (1) that  $\mathcal{F}(t)$  can be computed by

$$\begin{aligned} \mathcal{F}(t) &= \lambda M(x) \dot{x} - \frac{\lambda}{\lambda + s} r(t) \\ r(t) &:= T(t) + C(x(t)) \dot{x}(t) + \lambda M(x(t)) \dot{x}(t). \end{aligned} \quad (12)$$

The filtered force  $\mathcal{F}(t)$  is related to the unknown parameter vector  $\Theta$  in (9) by

$$\mathcal{F}(t) = \boldsymbol{\rho}(t)^T \Theta \quad (13)$$

where  $\boldsymbol{\rho}(t)$  is the filtered regressor vector given by

$$\boldsymbol{\rho}(t) := \frac{\lambda}{s + \lambda} \boldsymbol{\phi}(t). \quad (14)$$

### D. Dynamic Damping Controller

A dynamic damping control law was designed in [1] to enable the user to perform an exercise specified by a desired velocity field  $V_d(x, t) > 0$ . It has the additional property that the closed-loop exercise system interacts passively with the user to ensure that the machine is safe to operate. The controller takes the position and velocity of the exercise motion  $[x(t), \dot{x}(t)]$  as input and generates the motor generalized force  $T(t)$  as the output. It takes the form

$$\begin{pmatrix} T \\ M_2 \dot{v}_2 \end{pmatrix} = -\mathbf{D}(x, \mathbf{v}_a, t) \begin{pmatrix} \dot{x} \\ v_2 \end{pmatrix} \quad (15)$$

where  $\mathbf{v}_a := (\dot{x} \ v_2)^T \in \mathfrak{R}^2$ . In (15),  $v_2$  is an internal state, and  $M_2$  is a controller parameter. They can be interpreted as the velocity and the inertia of a fictitious flywheel (see [1] for details).  $\mathbf{D}(x, \mathbf{v}_a, t) \in \mathfrak{R}^{2 \times 2}$  is a positive definite matrix and has the following structure (see [1, eqs. (31)–(35)]):

$$\mathbf{D}(x, \dot{x}, v_2, t) = \mathbf{D}_1(x, \dot{x}, v_2, t) + F_d(x, t) \mathbf{B}_1(x, t) \quad (16)$$

where  $\mathbf{D}_1(x, \dot{x}, v_2, t)$  is skew symmetric,  $\mathbf{B}_1(x, t)$  is positive definite, and

$$F_d(x, t) := F_h(x, V_d(x, t)) = a(x) - b(x)V_d(x, t) \quad (17)$$

is the force that the user would exert if  $\dot{x} = V_d(x, t)$ . In this paper, we assume that the matrix  $\mathbf{B}_1(x, t)$  is defined using [1, eq. (32)], so that its off-diagonal elements are skew symmetric. Let  $\phi_d(t) \in \mathbb{R}^{2N}$  be a regressor vector defined by

$$\phi_d(t) := \begin{pmatrix} 1 \\ -V_d(x(t), t) \end{pmatrix} \otimes \int_G k(x(t), \sigma) \Psi(\sigma) d\sigma \quad (18)$$

then

$$F_d(x(t), t) = \phi_d(t)^T \Theta. \quad (19)$$

Given a desired velocity field  $V_d(x, t) > 0$ , the dynamic damping controller is fully determined without the knowledge of the user's Hill surface, except for the function  $F_d(x, t)$ .

### E. Self-Optimizing Control Strategy

The dynamic damping controller is, by itself, insufficient to enable the user to exercise optimally. For this to happen, the dependence on the knowledge of the user's Hill surface in (15) has to be removed, and a mechanism is necessary to specify the desired velocity field so that  $V_d(x, t) \rightarrow V^*(x)$ , where  $V^*(x)$  is the optimal velocity field in (4) which is *a priori* unknown.

To accomplish this objective, a control scheme which consists of an *adaptive controller* and a *reference generator* is proposed (Fig. 1). The reference generator commands a reference velocity field, and the adaptive controller makes certain that it is faithfully executed. In turn, the adaptive controller passes to the reference generator the estimate  $\hat{\Theta}(t)$  of the Hill surface parameter  $\Theta$ , which the reference generator uses to determine the reference velocity field. Intuitively, if the parameter estimate generated by the adaptive controller is accurate, then the reference generator will be able to compute  $V^*(x)$ , which will then be tracked by the adaptive controller. For this reason, the adaptive controller should, in addition to making sure that the reference velocity field is executed, estimate the system parameters  $\Theta$  accurately. On the other hand, to correctly identify  $\Theta$ , the reference velocity field must contain sufficient information. Thus, the reference generator described in Section IV time multiplexes a "training" velocity field, which provides the necessary excitation, with an estimate of the optimal velocity field based on an estimate of the Hill surface. As the estimate of  $\Theta$  improves, the "training" velocity field will be infrequently applied so that  $V_d(x, t) \rightarrow V^*(x)$  in some sense.

### III. ADAPTIVE DYNAMIC DAMPING CONTROL

Given the desired velocity field  $V_d(x, t)$ ,  $F_d(x, t)$  is the only term in the dynamic damping controller (15) and (16), that cannot be determined, since it depends on the unknown parameter vector  $\Theta$ . Using the certainty equivalence approach, we replace  $F_d(x, t)$  given by (19) by its estimate  $\hat{F}_d(x, t)$ :

$$\hat{F}_d(x(t), t(t)) = \phi_d(t)^T \hat{\Theta}(t) \quad (20)$$

where  $\phi_d(t)$  is defined in (18), and  $\hat{\Theta}(t)$  is the estimate of  $\Theta$ .

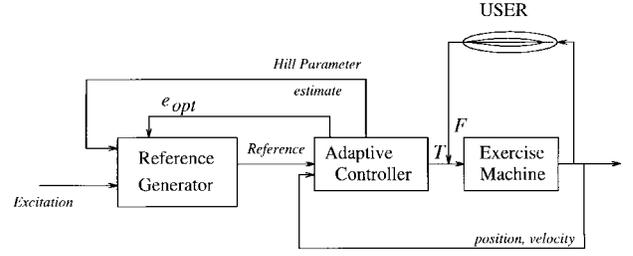


Fig. 1. Control scheme for a smart exercise machine.

The parameter estimate vector  $\hat{\Theta}$  is updated using the following parameter adaptation algorithm (PAA), which is a modification of the identifier proposed in [10]:

$$\dot{\hat{\Theta}}(t) = \mu \text{Proj}_{\Theta}[-\mathbf{P}(t)\hat{\Theta}(t) + \mathbf{d}(t) + \phi_d(t)e_1(t)] \quad (21)$$

$$\dot{\mathbf{P}}(t) = -\lambda \mathbf{P}(t) + \rho(t)\rho^T(t), \quad \mathbf{P}(0) = \mathbf{0} \quad (22)$$

$$\dot{\mathbf{d}}(t) = -\lambda \mathbf{d}(t) + \rho(t)\mathcal{F}(t), \quad \mathbf{d}(0) = \mathbf{0} \quad (23)$$

where  $e_1(t) := \dot{x}(t) - V_d(x(t), t)$  is the velocity field tracking error,  $\lambda > 0$  is a forgetting factor,  $\mu > 0$  is a gain constant, the regressors  $\phi_d(t)$  and  $\rho(t)$  are defined in (14) and (18),  $\mathcal{F}(t)$  is the force observer output given by (12), and  $\text{Proj}_{\Theta}$  is the projection operator that makes sure that  $\hat{\Theta}$  remains in the convex set  $\mathcal{P}_{\Theta} \subset \mathbb{R}^{2N}$ , so that the Hill parameter estimates  $\hat{a}(x, t)$ ,  $\hat{b}(x, t)$ , obtained from  $\hat{\Theta}$ , satisfy  $\hat{a}(x, t) \geq \underline{a}(x)$ ,  $\hat{b}(x, t) \in [\underline{b}(x), \bar{b}(x)]$ , as specified in *Assumption 1*.

The key feature of this PAA is the existence of a computable signal which is directly related to the parameter estimation error. Multiplying both sides of (22) by  $\Theta$ , we observe that  $\mathbf{P}(t)\Theta$  and  $\mathbf{d}(t)$  satisfy the same linear differential equation and initial conditions. Therefore, by the uniqueness of the solution of an ordinary differential equation (ODE),

$$\mathbf{d}(t) = \mathbf{P}(t)\Theta \quad \forall t \geq 0.$$

Therefore, the parameter error  $\tilde{\Theta}(t) := \hat{\Theta} - \Theta$  satisfies

$$\mathbf{P}(t)\tilde{\Theta}(t) = \mathbf{P}(t)\hat{\Theta} - \mathbf{d}(t). \quad (24)$$

This enables us to obtain a direct measurement of the parameter error  $\tilde{\Theta}$ , whenever the matrix  $\mathbf{P}(t)$  is invertible. Using this fact, (21) can be rewritten as

$$\dot{\hat{\Theta}}(t) = \mu \text{Proj}_{\Theta}[-\mathbf{P}(t)\tilde{\Theta}(t) + \phi_d(t)e_1(t)]. \quad (25)$$

*Theorem 1:* Suppose that for some  $\underline{V}(x) > 0$  and  $\underline{F}(x) > 0$ , the desired velocity field  $V_d(x, t)$  satisfies

$$\underline{V}(x) < V_d(x, t) < \frac{\underline{a}(x) - \underline{F}(x)}{\bar{b}(x)}$$

where  $\underline{a}(x)$ ,  $\bar{b}(x)$  are the bounds on the functions  $a(x)$  and  $b(x)$  in *Assumption 1*.

The adaptive dynamic damping controller (15) and (16) with  $F_d(x, t)$  replaced by  $\hat{F}_d(x, t)$  in (20) and  $\hat{\Theta}(t)$  updated by the PAA in (21)–(23) has the following properties.

- 1) The closed-loop system with the human force  $F(t)$  as the input and the velocity  $\dot{x}(t)$  as the output is passive with respect to the supply rate  $F(t)\dot{x}(t)$ .

- 2) Suppose that the user's force  $F(t)$  is given by a Hill surface which satisfies *Assumption 1*. Let  $e_1(t) := \dot{x}(t) - V_d(x(t), t)$ . Then,  $e_1(t) \rightarrow 0$ . Thus, the user asymptotically exercises according to the desired velocity field  $V_d(x, t)$ .
- 3) Let  $r(t)$  be the minimum eigenvalue of  $\mathbf{P}(t)$  in (22). If

$$\int_0^t r(\tau) d\tau \rightarrow \infty \quad \text{as } t \rightarrow \infty$$

then  $\tilde{\Theta}(t) \rightarrow 0$  as  $t \rightarrow \infty$ .

*Proof:* See Appendix I. ■

Notice that two types of error signals are used in (25) for adaptation, the ‘‘prediction’’ error signal  $\mathbf{P}(t)\tilde{\Theta}(t)$  and the ‘‘output’’ error signal  $e_1(t) = \dot{x}(t) - V_d(x(t), t)$ . The former, which is furnished by the force observer output  $\mathcal{F}(t)$ , is statically related to the parameter error  $\tilde{\Theta}$ , whereas the latter is only related to  $\tilde{\Theta}$  through the dynamics of the system. The result in *Theorem 1* that  $e_1 \rightarrow 0$  can be attributed to the ‘‘output’’ error and can be achieved using a simpler PAA, such as a gradient PAA. However, the use of the ‘‘prediction error’’ and the filtered gain matrix update law (22) is key to item 3) in *Theorem 1*. In particular, as will be shown in Section IV, this result enables the parameter error to vanish asymptotically under a less restrictive condition for the regressor vector  $\boldsymbol{\rho}(t)$  than the so-called *persistence of excitation* condition [13].

#### IV. REFERENCE GENERATOR

The adaptive controller in Section III enables the user to exercise as specified by the desired velocity field  $V_d(x, t)$ , despite uncertainty in  $F_d(x(t), t)$  in (17). There is still the remaining problem of specifying  $V_d(x, t)$  to be the optimal velocity profile (4). Unfortunately, the adaptive controller itself cannot guarantee that  $\hat{\Theta}(t) \rightarrow \Theta$ . Therefore, the optimal velocity field *estimate*

$$\hat{V}^*(x, t) = \frac{\rho}{\rho + 1} \frac{\hat{a}(x, t)}{\hat{b}(x, t)} \quad (26)$$

computed based on the parameter estimate  $\hat{\Theta}(t)$  does not necessarily converge to  $V^*(x)$ .

To resolve this difficulty, we now design a reference generator which specifies  $V_d(x, t)$  by time multiplexing an estimate of the optimal velocity field and a training velocity field  $V^{\text{tr}}(x, t)$ . The latter provides sufficient excitation so that  $\Theta$  can be identified. A binary state machine, called the excitation supervisor, determines which of the two types of velocity fields will be chosen, based on an error signal  $e_{\text{opt}}(t)$ , which monitors the accuracy of the estimated parameters. Switching is performed in such a way that the training velocity field is chosen less frequently as the estimate of  $\Theta$  improves.

##### A. Training Velocity Field

To qualify as a training velocity profile, the velocity field  $V^{\text{tr}}(x, t)$  must satisfy the following excitation condition.

*Excitation Condition:* Let  $\boldsymbol{\rho}(t)$  be the regressor function in (14), so that the force observer output is  $\mathcal{F}(t) = \boldsymbol{\rho}(t)^T \Theta$ . There exists a constant  $T_{\text{max}} > 0$  and a continuous function  $\alpha: \mathfrak{R} \rightarrow \mathfrak{R}$  with  $\alpha(0) > 0$  such that for any  $T_0 > 0$ , if  $|\dot{x}(t) - V_d(x(t), t)| \leq \epsilon$  for all  $t \in [T_0, T_0 + T_{\text{max}}]$ , then

$$\int_{T_0}^{T_0 + T_{\text{max}}} \boldsymbol{\rho}(\tau) \boldsymbol{\rho}(\tau)^T d\tau \geq \alpha(\epsilon) I_{2N}. \quad (27)$$

Thus, if a training velocity field  $V^{\text{tr}}(x, t)$  satisfies the excitation condition, then, whenever the actual velocity  $\dot{x}$  tracks  $V^{\text{tr}}(x, t)$  sufficiently closely, the regressor  $\boldsymbol{\rho}(\tau)$  will span  $\mathfrak{R}^{2N}$  in a finite time  $T_{\text{max}}$ . This ensures that there is sufficient information contained in the force observer output  $\mathcal{F}(\cdot)$  to reconstruct the unknown parameter  $\Theta$ .

Since the Hill surface is affine in the velocity  $\dot{x}$  at each position  $x$ , it should be intuitive that the training velocity field must at least visit sufficiently often two different velocities at each position  $x$ . In the following,  $V^{\text{tr}}(x, t)$  is defined to consist of alternately two constant velocities,  $V_{\text{high}}$  and  $V_{\text{low}}$ , with a smooth transition between them. Hence,

$$V^{\text{tr}}(x, t) := L_1(t)V_{\text{high}} + L_2(t)V_{\text{low}} \quad (28)$$

where  $V_{\text{high}} > V_{\text{low}} > 0$ . The weighting vector  $\mathbf{L}(t) = [L_1(t), L_2(t)]^T$  is determined by

$$\frac{d}{dt} \begin{pmatrix} L_1 \\ L_2 \end{pmatrix} = -\lambda_L \begin{pmatrix} L_1 \\ L_2 \end{pmatrix} + \lambda_L p(t); \quad \lambda_L \gg 0, \quad (29)$$

$$p(t) = \begin{cases} (1 \ 0)^T, & \text{if } t \bmod (T_1 + T_2) < T_1 \\ (0 \ 1)^T, & \text{if } t \bmod (T_1 + T_2) \geq T_1 \end{cases} \quad (30)$$

where  $T_1 > 2n\pi/V_{\text{high}}$  and  $T_2 > 2n\pi/V_{\text{low}}$ , and  $n > 1$  (chosen to be 3 in the implementation) is the number of cycles the exercise motion must go through before the desired velocity is switched from  $V_{\text{high}}$  to  $V_{\text{low}}$  or vice versa.

The training velocity field, as defined above, indeed satisfies the excitation condition.

*Proposition 1:* If  $V_{\text{high}} > V_{\text{low}} > 0$  and  $\lambda_L$  in (29) is sufficiently large, then the training velocity field  $V^{\text{tr}}(x, t)$  defined in (28)–(30) satisfies the excitation condition with  $T_{\text{max}} = T_1 + T_2$ .

*Proof:* See [14].

##### B. Excitation Supervisor

The excitation supervisor is a binary-state machine which switches between the `control` and `train` states based on the event symbols `go_control` or `go_train`, as shown in Fig. 2. Roughly speaking, an estimate of the optimal velocity field is selected when the state is `control`, and the ‘‘training’’ velocity field is selected when the state is `train`. The transition times  $T_k$ ,  $k \in \mathcal{Z}$  will be defined subsequently. Denote the state before  $T_k$  by

$$q_{k-1} \in \{\text{control}, \text{train}\}.$$

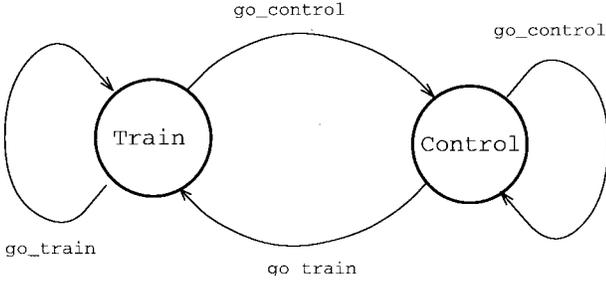


Fig. 2. Excitation supervisor state diagram.

Set the first transition time to be  $T_{k=0} = 0$ , and initialize the state of the supervisor to  $q_{k=0} = \text{train}$ . After each subsequent transition time  $T_k$ , the following events take place.

- 1) The state changes from  $q_{k-1} \rightarrow q_k$  based on the transition event symbol issued. These events will be specified later.
- 2) The estimate of the optimal velocity field is updated as follows:

$$V_k^*(x) = \begin{cases} V_{k-1}^*(x), & \text{if } q_{k-1} = \text{control} \\ \hat{V}^*(x, T_k), & \text{if } q_{k-1} = \text{train} \end{cases}$$

where  $\hat{V}^*(x, T_k)$  is the estimate of the optimal velocity field based on the parameter estimate  $\hat{\Theta}(T_k)$ , as defined in (26). Notice that  $\hat{V}^*(x, T_k)$  is updated only if  $q_{k-1} = \text{train}$ .

- 3) The velocity field to be selected after  $T_k$  is given by

$$V_k^{\text{com}}(x, t) = \begin{cases} V^{\text{tr}}(x, t), & \text{if } q_k = \text{train} \\ V_k^*(x), & \text{if } q_k = \text{control} \end{cases}$$

- 4) The reference velocity field  $V_d(x, t)$  is defined to be the smooth interpolation between  $V_{k-1}^{\text{com}}(x, t)$  and  $V_k^{\text{com}}(x, t)$ . This can be accomplished as follows. Define  $V_d(x, t)$  to be the value of a polynomial spline of sufficiently high order between  $V_{k-1}^{\text{com}}(x, t)$  and  $V_k^{\text{com}}(x, t)$  over the period  $[T_k, T_k + T_{\text{sp}}^k]$  and to be  $V_k^{\text{com}}(x, t)$  for  $t \in (T_{\text{sp}}^k, T_{k+1}]$ . When  $q_k = q_{k-1}$ , no smoothing is necessary, so that  $T_{\text{sp}}^k = 0$ .
- 5) Finally, the next transition time  $T_{k+1}$  is chosen:

$$T_{k+1} = \begin{cases} T_k + T_{\text{max}} + T_{\text{sp}}^k, & \text{if } q_k = \text{train} \\ T_k + T_{\text{dwell}}, & \text{if } q_k = \text{control} \end{cases}$$

where  $T_{\text{max}}$  is specified in the excitation condition which  $V^{\text{tr}}$  is assumed to satisfy, and  $T_{\text{dwell}}$  is set to the amount of time the exercise motion takes to complete  $n = 3$  turns. Notice that  $T_1, T_2, \dots$  is an unbounded sequence.

The following *Proposition* states that, if the transition symbol  $\text{go\_train}$  is chosen infinitely many times, then the parameter estimates converge to the true parameters.

*Proposition 2:* Let  $\{T_{j_1}, \dots, T_{j_k}\}$  be a subsequence of increasing transition times. Suppose that at each  $T_{j_i}$ ,  $i = 1, \dots, k$ , the training velocity field  $V^{\text{tr}}$  is chosen, then for  $T_{j_1}$  large enough, there exists  $\delta > 0$ , so that

$$\int_0^\infty r(\tau) d\tau \geq kc_1, \quad c_1 := \frac{\exp(-\lambda T_{\text{max}})}{\lambda} \delta \alpha(0) \quad (31)$$

where  $r(t)$  denotes the minimum eigenvalue of  $\mathbf{P}(t)$  in (22), and  $T_{\text{max}}$  and  $\alpha(e)$  are, respectively, the time interval and the scalar function in the lower bound in (27).

Moreover, if  $k \rightarrow \infty$ , then the parameter error  $\tilde{\Theta}(t) \rightarrow \mathbf{0}$  as  $t \rightarrow \infty$ .

*Proof:* See Appendix II. ■

*Proposition 2* states that if the training velocity field is chosen infinitely many times, then  $\tilde{\Theta}(t) \rightarrow \mathbf{0}$ . This does not, however, preclude the training velocity field from being selected at a vanishing frequency.

The proof of *Proposition 2* gives immediately the following result.

*Lemma 1:* For  $T_k$  sufficiently large, if  $q_{k-1} = \text{train}$ , the matrix  $\mathbf{P}(T_k)$ , where  $\mathbf{P}(\cdot)$  is given by (22), is invertible.

The conditions that  $T_{j_1}$  and  $T_k$  are sufficiently large in *Proposition 2* and *Lemma 1* are satisfied if the tracking error  $e_1(t) = \dot{x}(t) - V_d(x(t), t)$  is sufficiently small, as guaranteed by *Theorem 1*. In implementation, this is attained very quickly, and we can assume the conditions in *Proposition 2* and *Lemma 1* are satisfied if  $j_1 \geq 1$  and  $k \geq 1$ , respectively.

### C. State Transition

We still need to specify how the transition events ( $\text{go\_train}$  or  $\text{go\_control}$ ) which trigger the state transitions should be issued. They will be determined from the optimality error signal  $e_{\text{opt}}(k)$ , which we now define.

Let us denote the estimate of the objective function in (3) at the velocity  $V$ , based on the parameter estimate  $\hat{\Theta}$ , by  $J_\rho(x, V, \hat{\Theta})$ , i.e., if  $\hat{\Theta}$  gives rise to the estimate  $\hat{a}(x)$  and  $\hat{b}(x)$  of  $a(x)$  and  $b(x)$  in (2), then

$$J_\rho(x, V, \hat{\Theta}) := V^\rho [\hat{a}(x) - \hat{b}(x)V].$$

From (26), the estimate of the optimal value of the objective function based on the parameter estimate  $\hat{\Theta}$  is denoted by  $J_\rho^*(x, \hat{\Theta})$ :

$$J_\rho^*(x, \hat{\Theta}) = \left[ \frac{\rho}{(1+\rho)} \frac{\hat{a}(x)}{\hat{b}(x)} \right]^\rho \frac{\hat{a}(x)}{1+\rho}.$$

Define the optimality error signal  $e_{\text{opt}}$  at time  $T_k$  to be

$$e_{\text{opt}}(k) = \gamma_1 g_1(k) + g_2(k), \quad \gamma_1 > 0 \quad (32)$$

where

$$g_2(k) = \begin{cases} \|\hat{\Theta}(T_k) - \mathbf{P}^{-1}(T_k)\mathbf{d}(T_k)\|^2, & \text{if } q_{k-1} = \text{train} \\ g_2(k-1), & \text{otherwise} \end{cases} \quad (33)$$

and  $g_1(k) = 0$  when  $q_{k-1} = \text{Control}$ , and it is

$$g_1(k) = \max_{x,t} |J_\rho^*(x, \hat{\Theta}(t)) - J_\rho(x, V_{k-1}^*(x), \hat{\Theta}(t))| \quad (34)$$

if  $q_{k-1} = \text{Train}$ , where the max is taken over  $t \in [T_{k-1}, T_k]$  and  $x \in \mathcal{G}$ .

The purpose of  $g_1(k)$  is to detect instances when  $\Theta$  has changed (e.g., due to changes in motivation or fatigue) by checking if the current parameter estimate  $\hat{\Theta}(t)$  and the optimal velocity field estimate being specified  $\hat{V}_{k-1}^*$  are consistent. Its exact definition, however, is not critical to the

theoretical result. Notice that, because of *Lemma 1*,  $g_2(k)$  in (33) is well defined after the velocity field tracking error  $e_1(t) = \dot{x} - V_d(x(t), t)$  is sufficiently small. This is ensured by *Theorem 1* and is achieved very quickly in implementation. Moreover, from (24),  $g_2(k)$  is, in fact,  $\|\tilde{\Theta}(T_j)\|^2$ , where  $T_j$  is the transition time after the previous training velocity field has been selected. Therefore, it measures the convergence of the parameter estimation error.

The transition event can now be defined. Denote the transition event at  $T_k$  by  $S(k) \in \{\text{go\_train}, \text{go\_control}\}$ . Two methods are proposed to determine the transition events, a *deterministic* method and a *stochastic* method:

*Deterministic:* Choose  $\Delta^2\theta > 0$ , so that if  $\|\Theta_1 - \Theta_2\|^2 \leq \Delta^2\theta$ , then the estimated optimal velocity field  $\hat{V}_{\Theta_1}^*(x)$  and  $\hat{V}_{\Theta_2}^*(x)$  are such that  $\max_x |\hat{V}_{\Theta_1}^*(x) - \hat{V}_{\Theta_2}^*(x)| < \underline{z}$  where  $\underline{z}$  is some predefined tolerance for suboptimality. Define the transition event by

$$S(k) = \begin{cases} \text{control}, & \text{if } e_{\text{opt}}(k) \leq \Delta^2\theta \\ \text{train}, & \text{otherwise} \end{cases}. \quad (35)$$

*Stochastic:* Let  $w: [0, \infty) \rightarrow \mathfrak{R}$  be a continuous non-decreasing function with  $w(e) = 0$  iff  $e = 0$ . Define the sequence  $\zeta_k, k = 0, 1, \dots$  to be an independent identically distributed (i.i.d.) random process with a uniform probability density distribution between  $[0, 1]$ . Then, define the transition event by

$$S(k) = \begin{cases} \text{go\_control}, & \text{if } \zeta_k > w(e_{\text{opt}}(k)) \\ \text{go\_train}, & \text{otherwise} \end{cases}. \quad (36)$$

Hence, `go_control` is more probable to occur if  $e_{\text{opt}}(k)$  is small.

The properties of the combination of the adaptive damping controller described in Section III and the reference generator described above are summarized in the following theorem.

*Theorem 2:* Assume that the user's force satisfies *Assumption 1*, and let  $V^*(x)$  in (4) be the true optimal velocity field.

If the transition events in the excitation supervisor are defined using the deterministic formula given by (35), then there exists an integer  $\kappa < \infty$  such that  $\forall j > \kappa$  the following apply:

- 1)  $q_j = \text{control}$ ;
- 2)  $e_{\text{opt}}(j), \gamma_1 g_1(j), g_2(j) < \Delta^2\theta$ ;
- 3) there exists a continuous function  $e(t)$ , such that

$$|\dot{x}(t) - V^*(x(t), t)| \leq \underline{z} + e(t)$$

where  $\underline{z}$  is the tolerance for suboptimality used to determine  $\Delta^2\theta$  in (35), and  $e(t) \rightarrow 0$  as  $t \rightarrow \infty$ .

If the transition events are defined using the stochastic formula given by (36), then as  $k \rightarrow \infty$ , the following apply:

- 1) with probability one,  $e_{\text{opt}}(k), g_1(k), g_2(k) \rightarrow 0$ ;
- 2) the probability that  $q_k = \text{control}$  tends to 1;
- 3)  $\exists e_2: \mathfrak{R} \rightarrow \mathfrak{R}$ , and  $e_2(t) \rightarrow 0$ , such that

$$\lim_{t \rightarrow \infty} \text{Prob}\{|\dot{x}(t) - V^*(x(t), t)| \leq e_2(t)\} = 1.$$

*Proof:* See Appendix III. ■

*Theorem 2* states that, when the transition events are determined by the deterministic rule, only a finite number

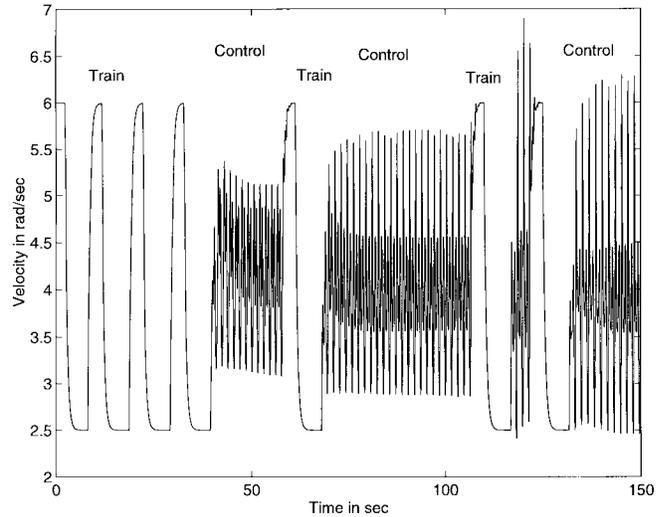


Fig. 3. The reference velocity profile  $V_d[x(t), t]$  alternates between training profiles (train) and estimated optimal profiles (control).

of `go_train` symbols occur. Thereafter, the user exercises according to a velocity profile close to the optimal velocity profile. If the transition events are determined by the stochastic rule, then, as  $t \rightarrow \infty$ , the probability that the `go_train` symbol occurs vanishes, and the user performs the optimal exercise with almost certainty.

In application, the user's Hill surface varies due to fatigue and other factors. By setting the parameters  $\gamma_1$  in (32),  $\Delta^2\theta$  in (35), and  $w(\cdot)$  in (36) appropriately, the excitation supervisor can be made insensitive to small variation of the Hill surface, but still be able to respond to sufficiently large changes in the Hill surface by transitioning to the `train` state, so as to learn the new Hill surface.

## V. EXPERIMENTAL RESULTS

The self-optimizing control strategy was implemented on the experimental setup described in [1]. The subject was instructed to exercise at a constant effort level. The mechanical power at each position in the exercise motion was to be maximized [i.e.,  $J_\rho$  in (3) with  $\rho = 1$ ]. The stochastic formula in (36) was used to determine the state transitions in the reference generator.

As shown in Fig. 3, the desired velocity field  $V_d(x(t), t)$  consisted of the training velocity fields which are alternate constant velocities at  $V_{\text{high}} = 6$  rad/s and  $V_{\text{low}} = 2.5$  rad/s and the estimated optimal velocity fields. Notice also that the training velocity field was less frequently applied as the experiment progresses, indicating that the control system was becoming confident of its estimate of the subject's Hill surface. From Fig. 4, where the actual velocity and the desired velocity are plotted, we see that the control was able to cause the subject to execute the desired profile during both the `train` state and the `control` state.

We now investigate whether the Hill surface was identified correctly. We have experimentally verified that the force observer output  $\mathcal{F}(t)$  was a good estimate of the user's force

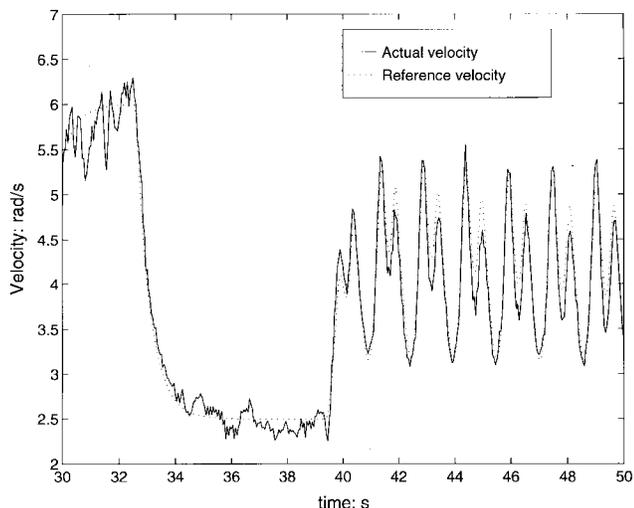


Fig. 4. Reference and actual velocity versus time.

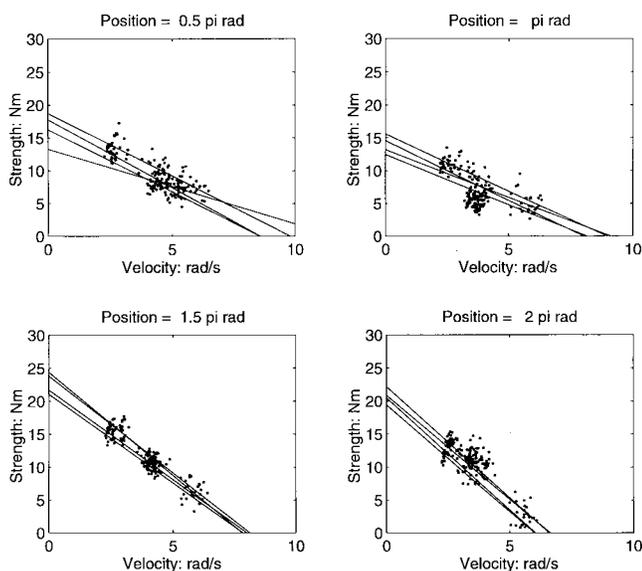
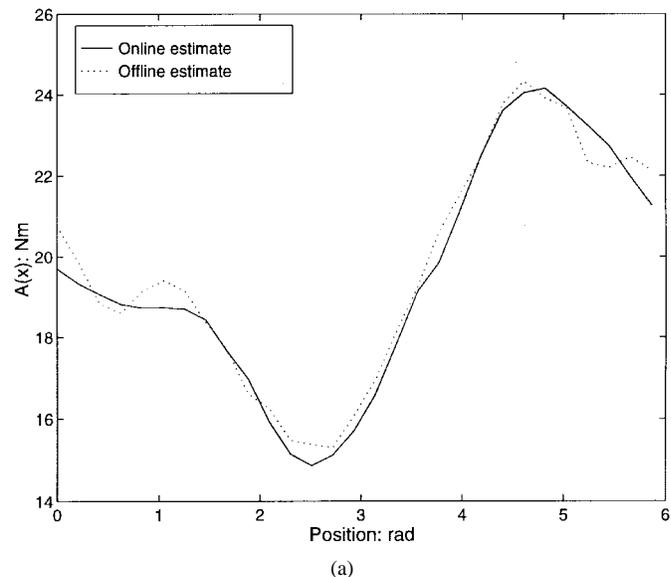
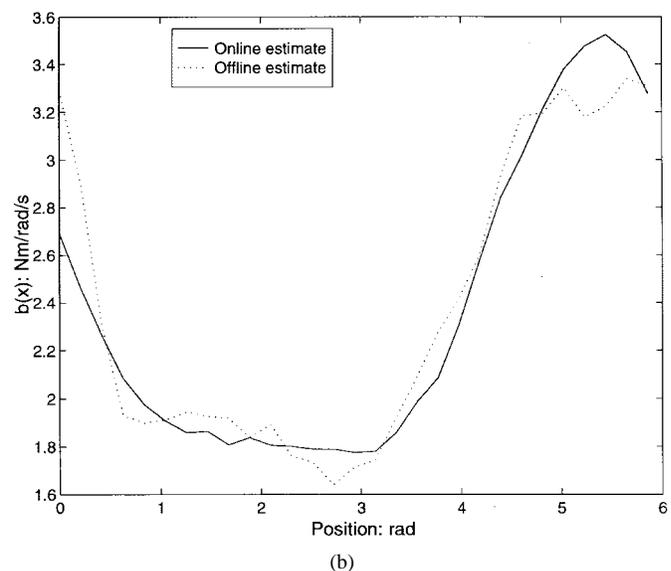


Fig. 5. Force versus velocity and estimated Hill curves at four positions.

$F(t)$ . This result was expected, since  $F(t)$  was slowly varying. Thus, using  $\mathcal{F}(t)$  as an estimate of the user's force  $F(t)$ , the actual Hill surface and the on-line estimates (approximately at  $t = 40$ ,  $t = 65$ ,  $t = 110$ , and  $t = 130$  s) at four positions are plotted in Fig. 5. Notice that the actual data and the estimates agree. An off-line least-square fit was generated for the force–velocity relation at each position. The resulting off-line estimates of the Hill parameters  $a(x)$  and  $b(x)$  were compared to the on-line estimates obtained at the second instance when the control was entered (at  $t \approx 70$  s). Notice that, in Fig. 6, the off-line estimate and the on-line estimate are also very similar. The Hill surfaces reconstructed using the off-line estimates and the on-line estimates of the Hill parameters are also very similar (Fig. 7). As shown in Fig. 8, the velocity at which the exercise was performed during the control state also follows the optimal velocity field computed based on the off-line estimate of the Hill surface, as desired.



(a)



(b)

 Fig. 6. Off-line and on-line estimates of the Hill parameters. (a) Estimates of  $a(x)$ . (b) Estimates of  $b(x)$ .

## VI. CONCLUSIONS

In this paper, we presented an intelligent, self-optimizing control system for a class of exercise machines. The control system optimizes the user's workout by causing the user to exercise according to a velocity field that optimizes a modified power criterion, dependent on the unknown user's biomechanical behavior—the Hill surface. Based on the dynamic damping controller developed in [1], we developed an adaptive controller which eliminates the need to know the Hill surface of the user beforehand. Since adaptive control alone does not guarantee that the unknown parameters are accurately identified, we also developed a reference generator to specify a desired velocity field which allows both the identification of the Hill surface (and, hence, the identification of the optimal exercise for the user) and the eventual execution of the optimal exercise. We achieve this by switching between the tasks of tracking the optimal velocity field and of training the control

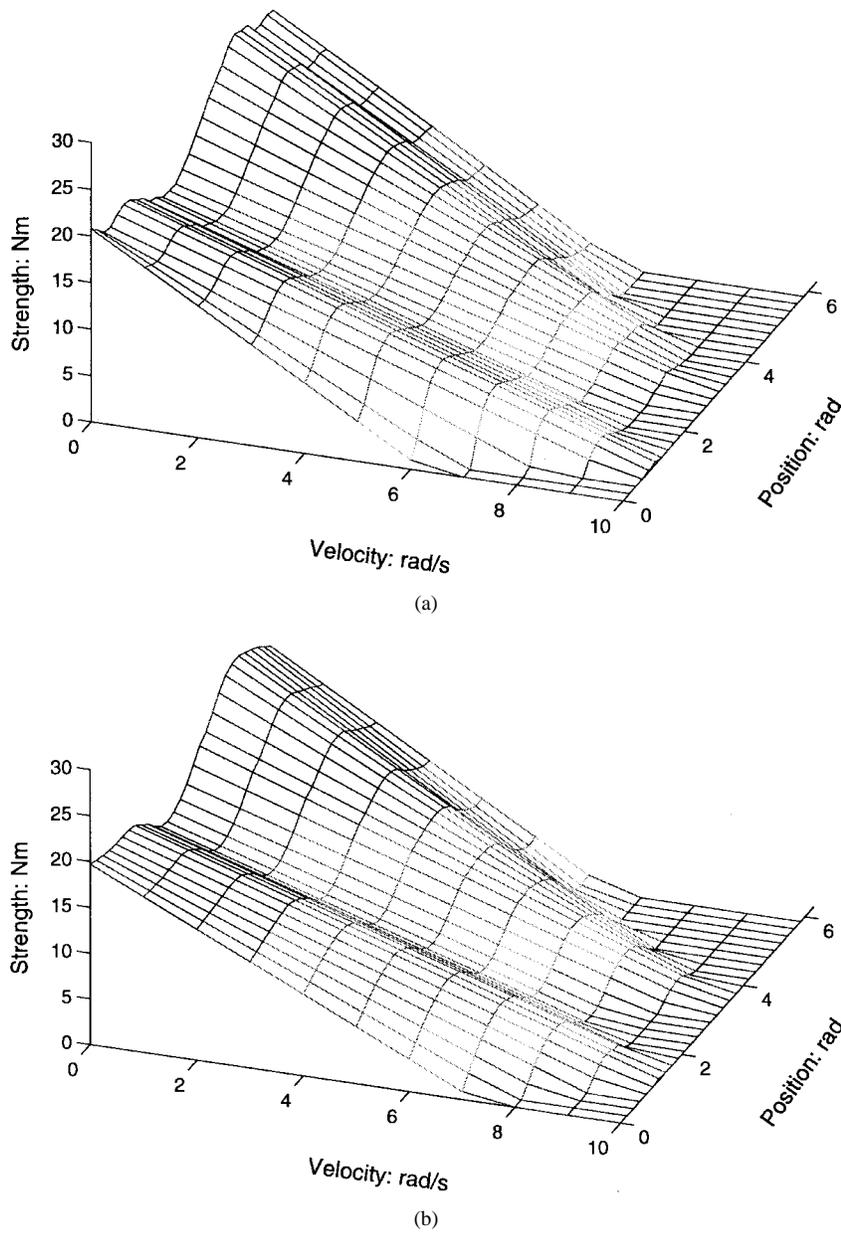


Fig. 7. Off-line and online estimates of the Hill surface. (a) Off-line estimate. (b) On-line estimate.

system to learn the Hill surface. The overall control scheme was implemented, and the experimental results verified the system's ability to obtain good estimates of the user's Hill surface and to cause the user to perform an exercise that maximizes the mechanical power output. A generalization of the self-optimizing control strategy to other situations can be found in [2].

APPENDIX I  
PROOF OF THEOREM 1

- 1) Following the proof of *Theorem 2* in [1], the I/O system is passive if the matrix  $\mathbf{D}(x, \mathbf{v}_a, t)$  is positive definite, which, in turn, is true if  $\hat{F}_d(x, t)V_d(x, t) > 0$ . Indeed,  $\hat{F}_d(x, t)V_d(x, t) > 0$  because

$$\hat{F}_d(x, t) = \hat{a}(x, t) - \hat{b}(x, t)V_d(x, t),$$

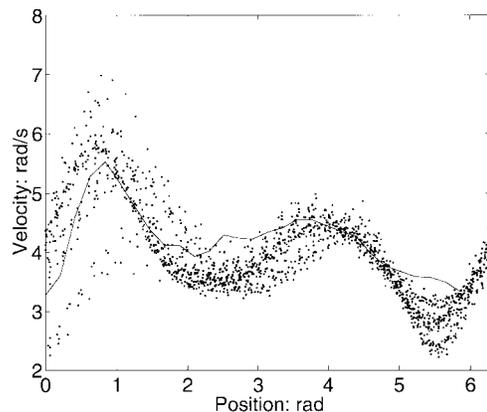


Fig. 8. Off-line computed optimal velocity profile (solid line) and achieved velocity profiles (dots).

$\hat{a}(x, t) \geq 0$  and  $\hat{b}(x, t) \in [\underline{b}, \bar{b}]$  as guaranteed by the parameter projection algorithm, and  $V_d(x, t) < \underline{a}/\bar{b}$  by *Assumption 1*.

- 2) Let  $W(t)$  be the Lyapunov function used in the proof of *Theorem 2* in [1] which has the property that  $W(t) \geq \frac{1}{2}M(x)e_1^2$ . Following the same proof, with  $F_d(x, t)$  substituted by  $\hat{F}_d(x, t)$  and using the fact that  $F_d(x, t)V_d(x, t) \geq \underline{FV}$ , we obtain

$$\dot{W}(t) \leq -\alpha W(t) + e_1(t)[F(t) - \phi_d(t)^T \hat{\Theta}(t)], \quad \alpha = \frac{\underline{FV}}{\bar{E}}$$

where  $\bar{E}$  is the constant defined in [1, eq. (13)]. Using the fact that  $F(t) = \phi_d(t)^T \Theta - b(x(t))e_1(t)$ , we obtain

$$\dot{W}(t) \leq -\alpha W(t) - b(x(t))e_1(t)^2 - e_1(t)\phi_d(t)^T \hat{\Theta}(t).$$

Define a new Lyapunov function given by

$$W_1(t) := W(t) + \frac{1}{2\mu} \hat{\Theta}(t)^T \hat{\Theta}(t).$$

Differentiating with respect to time, making use of (25), and the property of the projection [13] that  $\hat{\Theta}^T \text{Proj}_{\hat{\Theta}}[s] \leq \hat{\Theta}^T s$ , we obtain

$$\begin{aligned} \dot{W}_1(t) &\leq -\alpha W(t) - e_1 \phi_d(t)^T \hat{\Theta} - \hat{\Theta}(t)^T \\ &\quad \cdot [\mathbf{P}(t)\hat{\Theta}(t) - \phi_d(t)e_1(t)] \end{aligned} \quad (37)$$

$$\leq -\alpha W(t) - \hat{\Theta}(t)^T \mathbf{P}(t)\hat{\Theta}(t). \quad (38)$$

Standard Lyapunov and Barbalat's lemma arguments show that  $W(t) \rightarrow 0$ , from which  $e_1(t) \rightarrow 0$  is obtained.

- 3) By the definition of  $r(t)$  and (38)

$$\dot{W}_1(t) \leq -\alpha W(t) - r(t)\|\hat{\Theta}(t)\|^2.$$

Let  $\sigma(t) = \min[\alpha, 2\mu r(t)]$ , then  $\dot{W}_1(t) \leq -\sigma(t)W_1(t)$  and, therefore,

$$W_1(t) \leq W(0) \exp\left[-\int_0^t \sigma(\tau) d\tau\right].$$

Because the dynamics of  $\mathbf{P}(t)$  given in (22) is a stable linear filter with bounded input,  $r(t)$  is bounded. It is easy to show that, since  $r(t)$  is bounded, and  $\alpha > 0$ ,  $\int_0^t r(\tau) d\tau \rightarrow \infty$  implies that  $\int_0^t \sigma(\tau) d\tau \rightarrow \infty$ . The latter, in turn, shows that  $W_1(t) \rightarrow 0$ .  $\hat{\Theta}(t) \rightarrow 0$  follows.

## APPENDIX II

### PROOF OF PROPOSITION 2

Because of the tracking property of the adaptive controller (*Theorem 1*), for each  $\epsilon > 0$ , there is a  $T_{j1}$  sufficiently large, so that  $\forall t \geq T_{j1}$ ,  $|e_1(t)| \leq \epsilon$  where  $e_1(t) = \dot{x}(t) - V_d(x(t), t)$ . If  $V_d(x(t), t) = V^{\text{tr}}(x(t), t)$ , then, since  $\alpha(\cdot)$  in the excitation condition in (27) is continuous, there exists  $\delta > 0$ , so that  $\alpha(|e_1(t)|) \geq \delta\alpha(0)$  for all  $t \geq T_{j1}$ .

Let  $\mathbf{P}_i(t)$  be the solution to (22) with  $\boldsymbol{\rho}(t)$  replaced by  $\boldsymbol{\rho}_i(t)$ , the truncation of  $\boldsymbol{\rho}(t)$  to  $[T_{ji} + T_{sp}^{ji}, T_{ji} + T_{sp}^{ji} + T_{\max}]$  when  $V_d(x(t), t) = V^{\text{tr}}(x(t), t)$ . Thus,

$$\dot{\mathbf{P}}_i(t) = -\lambda \mathbf{P}_i(t) + \boldsymbol{\rho}_i(t)\boldsymbol{\rho}_i(t)^T, \quad \mathbf{P}_i(0) = \mathbf{0}.$$

Let  $r_i(t)$  be the minimum eigenvalue of  $\mathbf{P}_i(t)$ . Using the linearity of (22) and the fact that for any  $\mathbf{r} \in \mathfrak{R}^{2N}$ ,  $\mathbf{r}^T \mathbf{P}(t)\mathbf{r} \geq$

$\mathbf{r}^T \mathbf{P}_1(t)\mathbf{r} + \mathbf{r}^T \mathbf{P}_2(t)\mathbf{r} + \dots + \mathbf{r}^T \mathbf{P}_k(t)\mathbf{r}$ , we have  $r(t) := \min \text{eig} \mathbf{P}(t) \geq \sum_{i=1}^k r_i(t)$ .

We will compute  $r_i(t)$  for  $t \geq T_{ji} + T_{sp}^{ji} + T_{\max}$ . Let  $\mathbf{r} \in \mathfrak{R}^{2N}$  be an arbitrary vector. Let  $a_i = T_{ji} + T_{sp}^{ji}$  and  $b_i = T_{ji} + T_{sp}^{ji} + T_{\max}$ . By computing  $r_i(b_i)$  using the convolution formula, we obtain

$$\begin{aligned} \mathbf{r}^T \mathbf{P}_i(b_i)\mathbf{r} &= \exp(-\lambda T_{\max}) \\ &\quad \cdot \int_{a_i}^{b_i} \exp[\lambda(\tau - a_i)] \mathbf{r}^T \boldsymbol{\rho}_i(\tau)\boldsymbol{\rho}_i(\tau)^T \mathbf{r} d\tau \\ &\geq \exp(-\lambda T_{\max}) \int_{a_i}^{b_i} \mathbf{r}^T \boldsymbol{\rho}_i(\tau)\boldsymbol{\rho}_i(\tau)^T x d\tau \\ &\geq \exp(-\lambda T_{\max}) \alpha(\epsilon) \|\mathbf{r}\|^2 \\ &\geq \exp(-\lambda T_{\max}) \delta \alpha(0) \|\mathbf{r}\|^2. \end{aligned}$$

The last two inequalities result from (27) and the fact that if  $T_{j1}$  is sufficiently large,  $|e_1(\tau)| \leq \epsilon$  where  $\epsilon$  is small. Thus,

$$r_i(b_i) \geq \delta \exp(-\lambda T_{\max}) \alpha(0) = \lambda \cdot c_1.$$

For  $t \geq b_i$ , we apply the transition function for the dynamics of  $\mathbf{P}_i$  and obtain  $\forall t \geq b_i = T_i + T_{sp}^i + T_{\max}$

$$r_i(t) \geq \exp[-\lambda(t - b_i)] \lambda c_1.$$

Integrating the above and using the fact that  $r(t) \geq \sum_{i=1}^k r_i(t)$ , we obtain the desired result:

$$\int_0^\infty r(\tau) d\tau \geq \sum_{i=1}^k \int_{a_i}^\infty r_i(\tau) d\tau \geq k c_1.$$

As  $k \rightarrow \infty$ ,  $\int_0^\infty r(\tau) d\tau \rightarrow \infty$ . So, by item 3) in *Theorem 1*,  $\hat{\Theta} \rightarrow 0$ .

## APPENDIX III

### PROOF OF THEOREM 2

To establish *Theorem 2*, we need the following lemma.

*Lemma 2:* Let  $\mathcal{A} = A_1, A_2, \dots$  be an independent, identically distributed (i.i.d.) random process with  $A_k \in \{0, 1\}$ . Suppose that  $\text{Prob}(A_k = 1) \geq p > 0$ ,  $\forall k$ . Then,

$$\text{Prob}(\mathcal{A} \text{ having only a finite number of 1's}) = 0.$$

*Proof:* Note that  $\text{Prob}(A_j = 0) := q_j \leq (1 - p) \leq 1$ . For any finite  $k > 0$ , the joint probability of  $A_j = 0$  for  $j \geq k$  is given by

$$\lim_{n \rightarrow \infty} \prod_{j=k}^n q_j \leq \lim_{n \rightarrow \infty} (1 - p)^{(n-k+1)} = 0.$$

*Proof of Theorem 2:* Consider first the deterministic case. We shall prove that  $e_{\text{opt}}(k) \leq \Delta^2 \theta$  for  $k$  sufficiently large. Suppose not, then  $\exists n_1 < n_2 < n_3 \dots \rightarrow \infty$ , such that  $e_{\text{opt}}(n_i) > \Delta^2 \theta$ , and, so, the training task is chosen infinitely often. By *Proposition 2*,  $\int_0^t r(\tau) d\tau \rightarrow \infty$ . This would imply (using *Theorem 1*, conclusion 3) that  $\hat{\theta}(T_k) \rightarrow \theta$  as  $k \rightarrow \infty$ .  $g_1(k)$ ,  $g_2(k)$  and  $e_{\text{opt}}$  would, therefore, have to converge to 0, extracting a contradiction. Hence,  $e_{\text{opt}}(k)$  is ultimately bounded by  $\Delta^2 \theta$ . ■

Let  $\kappa$  be an integer such that  $e_{\text{opt}}(T_k) \leq \Delta^2\theta \forall k \geq \kappa$  and  $j$  be the largest integer such that  $j \leq \kappa$  and  $S(j-1) = \text{go\_train}$ . Then, for  $k > \kappa$ ,  $S(k) = \text{go\_control}$  and  $g_2(k) = g_2(j) = \|\hat{\Theta}(T_j)\|^2 \leq \Delta^2\theta$ . For  $t \geq T_{\kappa+1}$ , since  $V_d(x(t), t) = V_j^*(x(t))$

$$|\dot{x}(t) - V^*(x(t))| \leq |V_j^*(x) - V^*(x)| + |\dot{x}(t) - V_d(x(t), t)| \leq \underline{z} + |e_1(t)|$$

where  $|e_1(t)| = |\dot{x}(t) - V_d(x(t), t)| \rightarrow 0$  by *Theorem 1*.

Next, we consider the probabilistic case. For any realization for which  $e_{\text{opt}}(k) \not\rightarrow 0$ ,  $\exists e_2 > 0$  and an integer sequence  $n_1 < n_2 < n_3 \dots \rightarrow \infty$ , such that  $e_{\text{opt}}(n_i) \geq e_2$ . At each  $n_i$ , the probability that the training phase is applied is positive. Thus, by *Lemma 2*, the probability that only a finite number of training phases are applied is zero. On the other hand, if an infinite number of training phases are applied (this occurs with probability one), by *Proposition 2*,  $\|\hat{\theta}(t)\|^2 \rightarrow 0$ . Thus,  $e_{\text{opt}} \rightarrow 0$ . This would contradict, with probability one, that  $e_{\text{opt}} \not\rightarrow 0$ . Hence,  $e_{\text{opt}}(k), g_1(k), g_2(k) \rightarrow 0$  with probability one.

Assume that  $e_{\text{opt}}(k) \rightarrow 0$ . Thus, the estimated optimal velocity field  $V_k^*$  converges to the true optimal,  $V^*$ . Moreover,  $\text{Prob}[S(k) = \text{go\_control}]$  tends to one asymptotically. Thus, as  $t \rightarrow \infty$ , the probability that  $V_d(x(t), t) = V_k^*(x(t))$  tends to one. By *Theorem 1*,  $\exists e_3: \mathfrak{R} \rightarrow \mathfrak{R}, e_3(t) \rightarrow 0$ , such that  $|\dot{x}(t) - V_d(x(t), t)| \leq e_3(t)$ . Therefore, as  $k \rightarrow \infty$  and  $t \in [T_k, T_{k+1}]$ , the probability that

$$|\dot{x}(t) - V^*(x(t))| \leq e_3(t) + |\hat{V}_{\hat{\Theta}(T_{p_k})}^*(x(t)) - V^*(x(t))| =: e_2(t)$$

tends to one, where  $\hat{V}_{\hat{\Theta}(T_{p_k})}^*(x)$  is the estimated optimal velocity field based on  $\hat{\Theta}(T_{p_k})$  and  $T_{p_k}$  is the last instance before  $T_k$  that the supervisor switches from `train` to `control`. Since  $\hat{\Theta}(T_{p_k}) \rightarrow \Theta$  [because  $g_2(k) \rightarrow 0$ ] and  $e_3(t) \rightarrow 0$ ,  $e_2(t) \rightarrow 0$ . The desired result is, therefore, obtained, since  $e_{\text{opt}}(k) \rightarrow 0$  with probability one.

## REFERENCES

- [1] P. Y. Li and R. Horowitz, "Control of smart exercise machines—Part I: Problem formulation and nonadaptive control," this issue, pp. 237–247.
- [2] ———, "Self-optimizing control," in *Proc. 36th CDC*, 1997, to be published.
- [3] J. Slotine and W. Li, "On the adaptive control of robot manipulators," *Int. J. Robotics Res.*, vol. 6, no. 3, pp. 49–59, 1987.
- [4] N. Sadegh and R. Horowitz, "Stability and robustness analysis of a class of adaptive controllers for robotic manipulators," *Int. J. Robotics Res.*, vol. 9, no. 3, pp. 74–92, June 1990.
- [5] S. Arimoto, S. Kawamura, and F. Miyazaki, "Bettering operation of robots by learning," *J. Robot Syst.*, vol. 1, no. 2, pp. 123–140, 1984.
- [6] W. Messner, R. Horowitz, W. Kao, and M. Boals, "A new adaptive learning rule," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 188–197, Feb. 1991.
- [7] A. A. Feldbaum (translated by A. Kraiman), *Optimal Control Systems*. New York: Academic, 1965.
- [8] J. W. Polderman, "Adaptive LQ control: Conflict between identification and control," *Linear Algebra Applicat.*, vol. 122, pp. 219–244, Sept. 1989.
- [9] M. Zarrop and M. Rommens, "Convergence of a multi-input adaptive extremum controller," *Proc. Inst. Elect. Eng.*, vol. 140, pt. D, no. 2, pp. 65–69, 1993.
- [10] G. Kreisselmeier, "Indirect adaptive control with a self-excitation capability," *IEEE Trans. Automat. Contr.*, vol. 34, pp. 524–528, May 1989.
- [11] R. Cristi, "Adaptive control with finite time persistency of excitation," *IEEE Trans. Automat. Contr.*, vol. 31, pp. 982–985, Oct. 1986.
- [12] R. Courant and D. Hilbert, *Methods of Mathematical Physics, Volume I*. New York: Interscience, 1953.
- [13] S. Sastry and M. Bodson, *Adaptive Control: Stability, Convergence, and Robustness*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
- [14] P. Y. Li, "Self optimizing control and passive velocity field control of intelligent machines," Ph.D. dissertation, Dep. Mech. Eng., Univ. California, Berkeley, 1995.

**Perry Y. Li** (S'87–M'89), for a photograph and biography, see this issue, p. 247.

**Roberto Horowitz** (M'89), for a photograph and biography, see this issue, p. 247.