# Hebbian Covariance Learning and Self-Tuning Optimal Control

by

Daniel L. Young
B.S., Mechanical Engineering
MIT, 1995

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of

Master of Science in Mechanical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1997



Author .....
Department of Mechanical Engineering
May 20, 1997

Certified by .....
Dr. Chi-Sang Poon
Harvard-MIT Division of Health Sciences and Technology
Thesis Supervisor

Accepted by .....
Prof. Ain A. Sonin
Chairman, Departmental Committee on Graduate Students



Eng

# Hebbian Covariance Learning and Self-Tuning Optimal Control

by

Daniel L. Young

B.S., Mechanical Engineering

MIT, 1995

Submitted to the Department of Mechanical Engineering
on May 20, 1997, in partial fulfillment of the
requirements for the degree of
Master of Science in Mechanical Engineering

## Abstract

A novel adaptive optimal control paradigm is presented that is inspired by Hebbian covariance learning, the celebrated biological synaptic process thought to underlie learning and memory and other complex biological functions. The adaptation is driven by the spontaneous fluctuations in the system input and output, the covariance of which provides useful information about the changes in the system behavior. Theoretical foundations for the paradigm are derived using Lyapunov theory. In numerous computer simulations, the controller is shown to effectively optimize linear and non-linear systems of arbitrary order in the presence of noise disturbances. The on-line adaptation method is computationally simple to apply in comparison to other optimal control schemes which may require complete parameter estimation. Furthermore, the learning algorithm is applicable to a wide class of real-world optimal control problems.

This thesis also explores the plausibility that Hebbian covariance learning may underlie respiratory control, satisfying certain physiological and neurobiological constraints. The respiratory system has been hypothesized to be regulated in an optimal fashion by a specialized brainstem center. The positive results of these simulations lend themselves to future inquiries into the computational functions of synaptic plasticity in biological neurons and into the neuronal processes which may underlie respiratory control.

Thesis Supervisor: Dr. Chi-Sang Poon
Title: Harvard-MIT Division of Health Sciences and Technology

# Acknowledgements

This thesis has been a pleasure to prepare. I am very grateful to Dr. Poon for his continual guidance and support throughout this project. His steadfast optimism and helpful suggestions were always motivating forces.

The members of the lab, Zhongren, Marina, and Mike, were all very encouraging. Discussions with them helped focus my ideas.

My family and friends have also been especially supportive. Heather has been patient with my long work hours and has been encouraging over the last few years. Mom and Dad have always supported me and their well wishes always lighten my load.

# Contents

# List of Figures

# Chapter 1

# Introduction

Adaptiveness is an innate ability of many living organisms which must conform to a wide range of environments. In this regard, biological systems must continually undergo vast physiological adjustments in order to survive in ever changing climatic, edaphic, or biotic conditions. In many instances, such as in sensory adaptation, these modifications may occur automatically. Organisms in such cases do not require conscious thought processes to synthesize the appropriate response. This innate adaptive ability suggests that there may exist fundamental neural mechanisms in the brain which computationally give rise to adaptive solutions.

Adaptive optimal control is one advanced engineering control method that continually regulates a system to best suit its environment. Such optimizing systems are generally needed when the system characteristics are changing and uncertain, and when resources are costly and performance goals are high. The presence of adaptive optimal behavior has long been recognized in certain organisms which are constrained by limited resources vital for survival [29].

The first step in designing an optimal control system is relating qualitative system goals to quantitative system measures, often in the form of a scalar performance index. Secondly, for time-varying systems an adaptive controller is required to optimize these performance measures in the face of uncertainties. Such an adaptive optimal control design is the focus of this study in which a criterion function embodying some scalar performance measure is extremized in an on-line fashion. Process control and

trajectory planning are just two instances where such an adaptive controller may be useful. The class of problems of interest has no solution by conventional methods of adaptive control, reinforcement learning or artificial neural networks.

In this thesis, we present a novel adaptive optimal control paradigm for such problems by analogy to some "intelligent" computational mechanisms which may exist in certain brain structures. Following several illustrative examples, this control strategy is applied to the respiratory system, a vital physiological system whose regulation of homeostasis appears to be optimal across diverse conditions [30, 31].

Several forms of learning control have been proposed to model adaptation in the higher brain. The "adaptive critic" method [5] is one such model that learns the optimal action by means of a reinforcement signal. Similar neural reinforcement learning models have been proposed to simulate certain animal behavior such as maze navigation or other optimal path finding tasks [28]. Alternatively, model-based learning paradigms construct internal models of the external environment in order to adaptively formulate the control law. This type of adaptive control, similar to model-reference adaptive control [11], has been applied to motor learning tasks [24, 40, 50] where an internal model of the musculoskeletal system and external loads is thought to be learned and stored in some brain regions such as the primary motor cortex or the cerebellum.

Typically, adaptive learning models employ an extensive network of neurons whose connections are modifiable by some form of synaptic plasticity. An important mechanism is Hebbian synaptic plasticity, first postulated by Hebb nearly 50 years ago [20]. This neural mechanism has been linked to certain intelligent animal behaviors, such as classical and instrumental conditioning [12], making it a particularly appealing substrate for computational inquires. Furthermore, Hebbian synapses have been postulated to play a fundamental role in learning and memory in the hippocampus and other brain structures [25, 46]. Over the past five decades, numerous incarnations of Hebb's original learning rule have been proposed based on both theoretical and experimental grounds [10].

Recently, Poon [32, 34] proposed a new role for Hebbian covariance learning in

homeostatic control of certain physiological functions, such as respiration. This respiratory brainstem controller is hypothesized to continually minimize the total cost of breathing (a function of the energy consumed by respiratory muscles and the chemical imbalance in the arterial circulation) despite continual changes in physiological and environmental states. In this respect, the respiratory regulator constitutes a brain model of an adaptive optimal controller.

Inspired by the postulated role of Hebbian covariance learning in respiratory control, we propose in this thesis a generalized paradigm for self-tuning optimal control employing certain Hebbian adaptation rules. The controller may be viewed as a reinforcement learning system in which spontaneous, random perturbations in the system states are used to advantage in probing the environment for surveillance. By weighing the resulting feedback signals against the perturbations applied, the controller effectively tunes the system to satisfy the control objective. More precisely, a single synapse, representing the system feedback gain, is adaptively modified based on the covariance between the controller output and the reinforcement signal as well as the autovariance of the controller output. As the unknown environment changes, the Hebbian controller continually conforms optimally to its new surroundings by exploiting spontaneous fluctuations in the system.

We propose a theoretical framework for the Hebbian covariance learning paradigm for dynamic optimal control and introduce the notion of long-term and near-term objective functions. A general Hebbian adaptation rule is derived which is applicable to a wide class of optimal control problems. In computer simulations of both linear and non-linear systems, we show that the Hebbian covariance controller may adapt optimally to its environment in a robust fashion despite the presence of uncertainties and noise disturbances.

This thesis also assesses the role of Hebbian covariance learning in the dynamic control of respiration. This notion was developed from the underling neural structure of the brainstem and the physiological responses demonstrated by the respiratory system. Computer simulations support this hypothesis and yield interesting predictions concerning the behavior of the respiratory brainstem controller.

The newly discovered intrinsic optimization ability of Hebbian covariance synapses is interesting from two standpoints. Firstly, it suggests a new computational role for the Hebbian synapses in the brain. As well, certain testable hypotheses may be drawn from the respiratory Hebbian covariance learning model. Secondly, this study advances the efforts to reverse engineer the body's remarkable capability as a robust and intelligent controller. Current engineering applications of this paradigm are being investigated in hopes of yielding optimal learning systems.

# Chapter 2

# Hebbian Covariance Learning for Adaptive Optimal Control

Hebbian adaptation is a common form of synaptic plasticity that is generally thought to play an important role in many cognitive functions of the brain such as learning and memory [7], vision [17, 44], motor control [22], and development [26, 54]. The classical Hebbian model [20] postulated that the synaptic connections between two neurons may be strengthened in time if the pre- and post-synaptic neural activity coincide with each other within some short time interval:

$$\frac{dW}{dt} = k\,(x \cdot y)\,, \tag{2.1}$$

where $W$ is the synaptic weight; $x$ and $y$ represent the mean firing rate of the input and output neurons, respectively; and $k$ is an adaptation constant. This type of associative synaptic modification, also called conjunctional Hebbian learning [10], forms the basis for NMDA receptor mediated synaptic long-term potentiation (LTP) [7, 9].

Despite its simplistic appeal, the classical Hebbian model of LTP has several theoretical limitations such as irreversible saturation by continued coactivity and random runaway instability of the interacting neurons. To circumvent these difficulties, a covariance Hebbian rule was formulated [38, 39] which allows both up and down

regulation of synaptic strength based on the degree of correlation between pre- and post-synaptic activity. A simple form of the Hebbian covariance rule is given by the following equation:

$$\frac{dW}{dt} = k\,\delta x \cdot \delta y, \tag{2.2}$$

where $\delta x$ and $\delta y$ are respectively the temporal variations of the pre- and post-synaptic activities about their mean values over a given time interval. Thus the synapse is strengthened on average if $\delta x$ and $\delta y$ are positively correlated, weakened if negatively correlated, and maintained at a constant average strength when changes are uncorrelated in time [38]. This Hebbian covariance algorithm describes synaptic LTP when $k > 0$ and LTD (long-term depression) when $k < 0$. An important feature of the above learning rule is that it responds to the changes in neural activity as opposed to the mean activity. Such a synaptic adaptation rule has been demonstrated in various systems including area 17 of the visual cortex [17], the CA1 region of the hippocampus [46, 45] and certain neuromuscular junctions [14].

Although Eq. 2.2 satisfies many of the requirements for LTP and LTD, a synapse of this form may still be saturated given a persistently excitatory or inhibitory connection [1]. Several attempts have been made to include a decay term which would prevent runaway instability or saturation. One such form is [32]:

$$\frac{dW}{dt} = k_1\,\delta x \cdot \delta y - k_2\,W\,g(\delta x, \delta y; x, y)\,, \tag{2.3}$$

where $g(\cdot)$ is some positive-definite function. In this formulation, the second term on the right hand side acts as the decay term. Consequently, the rate of potentiation will decrease with increasing $W$. Note that both terms on the right hand side are generally associative (i.e. dependent on pre- and post-synaptic activity). By suitably choosing the adaptation rates $k_1$ and $k_2$, one may describe either long-term plasticity or short-term plasticity.

To apply the above generalized Hebbian covariance rule to an adaptive control paradigm, we first consider a learning system having input ($x$) and output ($y$) which are connected in a closed loop via an external environment (Fig. 2-1). For simplicity

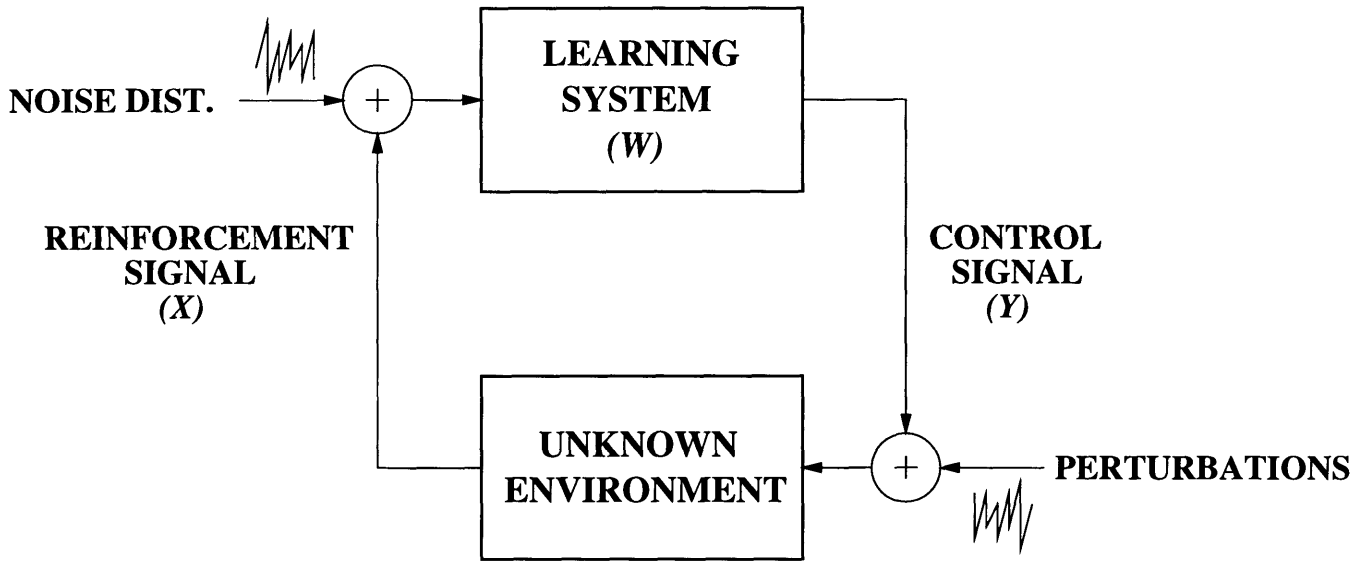


Figure 2-1: The Hebbian adaptive paradigm as a reinforcement learning system.

we neglect any intermediate connections and any other inputs which may be interposed between them. Assuming a linear input-output relationship in the learning system, the mean activity of the input and output signals can be related to a first approximation by the following equation:

$$y = Wx. \tag{2.4}$$

Now suppose the output signal is spontaneously fluctuating around some mean value. Such fluctuations ($\delta y$) in the output may occur, for example, as a result of random variations of neurotransmitter release in the presynaptic terminals or periodic variations of activity in a pacemaker cell or an oscillatory network. It has been suggested that persistent perturbations in learning systems may be beneficial in preventing spurious equilibrium states [2]. In what follows, we show that such spontaneous fluctuations in neural activity may provide a means of adaptive optimal control.

Due to the closed-loop structure of the system, the fluctuations in the output signal

of the learning system will result in corresponding fluctuations in the input ($\delta x$). The ratio $\delta x/\delta y$ is then a measure of the gradient of the input-output relationship of the environment. Hence, substitution of Eq. 2.4 into Eq. 2.3 results in a learning rule which is a function of the fluctuations in the input and output as follows:

$$\frac{dW}{dt} = k_1\,\delta x \cdot \delta y - k_2\,\frac{y}{x}\,g(\delta x, \delta y; x, y)\,. \tag{2.5}$$

A Hebbian covariance learning algorithm of the form in Eq. 2.5 may function as a neural optimizer. Thus, if changes in the output ($\delta y$) produce favorable changes in the input ($\delta x$), $W$ will be potentiated. Due to the decay term, this potentiation will be checked by the penalty associated with increases in the output. With a proper choice of the function $g(\cdot)$, the adaptation may serve to optimize a specific objective function of the learning system such that $dW/dt = 0$ at the optimal operating point.

## 2.1 Optimization in Static Systems

### 2.1.1 Theory

First consider the simple case in which the input-output relationship of the environment is given by a static, perhaps non-linear, continuously differentiable function:

$$x = f(y) \tag{2.6}$$

$$\frac{dx}{dy} = f'(y). \tag{2.7}$$

Suppose the goal of the learning system is to control the environment in some optimal fashion. The problem is then to minimize, by adaptive adjustment of the gain (i.e., synaptic weight $W$), a long-term objective function of the form:

$$J = J(x, y). \tag{2.8}$$

For convenience of discussion, we will assume that $J$ is continuous and has a unique

global minimum (i.e. lower bounded). A maximization problem may be formulated in a similar fashion by reversing the sign of $J$. Note that in our formulation, $J$ is generally a function of both the input ($x$) and the output ($y$). This is in contrast to certain adaptive control problems (such as adaptive stabilization or tracking problems) where the objective function is dependent on the output signal only.

To obtain an adaptation rule that would minimize $J$, we first differentiate Eq. 2.8 with respect to $W$:

$$\frac{dJ}{dW} = \frac{\partial J}{\partial x} \cdot \frac{dx}{dW} + \frac{\partial J}{\partial y} \cdot \frac{dy}{dW}. \tag{2.9}$$

Hence, using the chain rule we obtain the rate of change of the objective function during adaptation as:

$$\frac{dJ}{dt} = \frac{dJ}{dW} \cdot \frac{dW}{dt} = \left[\frac{\partial J}{\partial x} \cdot \frac{dx}{dy} + \frac{\partial J}{\partial y}\right] \cdot \frac{dy}{dW} \cdot \frac{dW}{dt}. \tag{2.10}$$

From Eqs. 2.4, 2.6, and 2.7 we have:

$$\frac{dy}{dW} = \frac{x}{1 - Wf'(y)}, \tag{2.11}$$

where the term $Wf'(y)$ in Eq. 2.11 is the loop gain of the feedback system linearized about the output $y$. A necessary condition for stability of the feedback system is:

$$Wf'(y) < 1. \tag{2.12}$$

Substitution of Eq. 2.11 into Eq. 2.10 yields:

$$\frac{dJ}{dt} = \left[\frac{\partial J}{\partial x} \cdot \frac{dx}{dy} + \frac{\partial J}{\partial y}\right] \cdot \frac{x}{1 - Wf'(y)} \cdot \frac{dW}{dt}. \tag{2.13}$$

The objective function $J$ will decrease continuously in time provided:

$$\frac{dJ}{dt} \leq 0, \text{ for all t}, \tag{2.14}$$

i.e., the time rate of change of $J$ is negative semi-definite. Since $J$ is lower bounded,

Eq. 2.14 implies that $J$ will converge to a limit as $t \to \infty$. Hence, from Eqs. 2.12-2.14 a sufficient condition that will guarantee the minimization of $J$ is:

$$\frac{dW}{dt} = -k_o\, x \cdot \left[\frac{\partial J}{\partial x} \cdot \frac{dx}{dy} + \frac{\partial J}{\partial y}\right], \tag{2.15}$$

where $k_o > 0$ is a proportionality constant. Noting that $dx/dy \approx \delta x/\delta y$, the above equation may be rewritten as:

$$\frac{dW}{dt} = -kx \cdot \frac{\partial J}{\partial x} \cdot \left[\delta x\, \delta y + \frac{\partial J}{\partial y} \Big/ \frac{\partial J}{\partial x} \cdot \delta y^2\right], \tag{2.16}$$

where $k \equiv k_o\, \delta y^2$.

The convergence of $\dot{J} \to 0$ may be guaranteed by Barbalat's Lemma [43]. Namely, because $J$ has a finite limit as $t \to \infty$, then if $\dot{J}$ is uniformly continuous, $\dot{J}(t) \to 0$ as $t \to \infty$.

Equation 2.16 is a modified Hebbian covariance rule (compare with Eq. 2.5) which effectively serves to optimize the objective function. There are two key components in this adaptation rule. The first is the covariance term $\delta x \delta y$ which computes the gradient of the input-output relationship of the environment, accounting for the resultant contribution of $x$ to the objective function due to the fluctuations in $y$. In this formalism, the signal $x$ is the reinforcement signal which acts to reward the learning system if the changes are beneficial. The second component, $\delta y^2$, computes the autovariance of the output $y$ and hence evaluates $y$'s changing contribution to the objective function due to its own fluctuations. This term is needed whenever the output variable is included in the objective function. The adaptation reaches steady state ($dW/dt = 0$) when these two terms balance each other. We illustrate the implementation of this algorithm with the following examples.

### 2.1.2 Examples

#### Criterion Dependent on Output Only

Let $J = -x^2/2$ and $x = f(y) = \theta_0 - (y - \theta_1)^2$ where $\theta_0$ and $\theta_1$ are unknown positive constants. Note that the environment system is static or 'memory-less'. We derive the perturbation equation as:

$$\delta x \, \delta y = -2 \, (y - \theta_1) \, \delta y^2. \tag{2.17}$$

From Eq. 2.16 the adaptation rule that will minimize $J$ is:

$$\frac{dW}{dt} = k \, x^2 \, \delta x \, \delta y. \tag{2.18}$$

Because $x^2 \geq 0$ at all times, this term only affects the rate of the adaptation and thus may be eliminated without affecting the optimal solution. The resulting adaptation rule is identical to the conventional Hebbian covariance rule (Eq. 2.2) suggesting the intrinsic optimizing ability of basic Hebbian covariance synapses. This adaptation rule in conjunction with the perturbation equation (Eq. 2.17) verifies the optimal solution $x = \theta_0$, $y = \theta_1$ and $W = \theta_1/\theta_0$.

#### Criterion Dependent on Both Input and Output

The foregoing example optimizes an objective function that is dependent only on the input to the controller, $x$. Many practical problems, however, are characterized by the need to extremize the system based on some criterion that is a function of both the input and the output. For instance, in designing a controller for a robotic arm, one may have a desired minimum error yet be constrained by the total energy consumption of the actuators. The design goal of such a problem could be stated as: choose an action which minimizes the sum of the divergence from the minimal error *and* the total energy expended. To illustrate, let us consider the same static system

as in Sect. 4.2.1 but with the following objective function:

$$J = J(x, y) = \frac{x^2}{2} + \frac{(2y)^2}{2}. \tag{2.19}$$

In this formulation, the first term reflects the departure of the system output from zero and the second term is some measure of the input energy.

By applying Eq. 2.16, we obtain the following adaptation rule which would minimize $J$:

$$\frac{dW}{dt} = -k \left[ \delta x \, \delta y + 2 W \, \delta y^2 \right]. \tag{2.20}$$

This equation is similar to the general Hebbian covariance rule with decay term (cf. Eq. 2.3). In this example, the intrinsic optimization character of the general Hebbian covariance rule is dramatic.

## 2.2 Optimization in Dynamic Systems

The optimization problem presented in Sect. 2.1 focuses on systems whose current states are related by static relationships with no explicit dependence on time and no 'memory'. In such cases, the long-term objective $J$ is satisfied simply by the proper transformation of the static, steady-state relationship into a Hebbian covariance rule. In most practical problems, however, the output and the input are governed by a dynamical relationship. For example, such dynamics may arise from inherent sluggishness (e.g. slow time constants) or time delays. In these systems, the current states which are sensed by the controller are not accurate indications of the steady-state values. One approach to solving this problem [34] is to introduce a near-term objective function, $Q$, which embodies the dynamic relations between the inputs and outputs. With some suitable transformation between $J$ and $Q$, the learning system may make short-term decisions extremizing $Q$ in the near-term and, hence, $J$ in the long-term.

### 2.2.1 Theory

To demonstrate the above general approach, we begin with a simple first-order dynamic environment of the form:

$$\tau \frac{dx}{dt} = q(x, y) = f(y) - x. \tag{2.21}$$

where $\tau$ is a time constant and $q(x, y)$ can be any non-linear function of the input and output signals, whereby in the steady-state, $x = f(y)$. (Note that this approach is extendible to higher order systems [34].) To simplify the analysis, we first discretize this equation, resulting in:

$$\frac{\tau \left[x(n+1) - x(n)\right]}{T} = q(x(n), y(n)), \tag{2.22}$$

where $T$ is the time step of the integration and $n$ is the time index.

Assuming random, uncorrelated noise in the controller output ($\delta y(n)$) and minimal recirculation of this output noise through the closed-loop system such that $\delta x(n)\delta y(n) = 0$ on average, Eq. 2.22 becomes:

$$\frac{\tau \delta x(n+1)\delta y(n)}{T} = \frac{\partial f(y(n))}{\partial y(n)} \delta y(n)^2, \tag{2.23}$$

after taking the partial derivatives and multiplying by the small perturbations where all terms with $\delta x(n)\delta y(n)$ have been removed.

In cases where the foregoing assumptions do not hold (for instance, if the system time constants are much slower than the period of the perturbations in $y$) all the covariance terms must be retained. Intuitively, the correlation $\delta x(n)\delta y(n)$ is significant because both the current state, $x(n)$, and the control signal, $y(n)$, are correlated with $x(n-1)$. In such cases, we replace $\delta x(n+1)$ in Eq. 2.23 with $\delta x^*(n+1)$ defined as:

$$\begin{aligned} \delta x^*(n+1) &= \delta x(n+1) - \left( \frac{T}{\tau} \frac{\partial q(x(n), y(n))}{\partial x(n)} + 1 \right) \delta x(n) \\ &= \delta x(n+1) + \left( \frac{T}{\tau} - 1 \right) \delta x(n), \end{aligned} \tag{2.24}$$

where all perturbation terms are now included. Note that two time steps of the state ($x(n)$ and $x(n+1)$) are incorporated in the perturbation equation (Eq. 2.24), thereby accounting for the structure of the 1st-order closed-loop system. In general, for an $n^{th}$ order system with closed-loop feedback, n+1 time steps in the state will be needed to form the augmented perturbation equation.

To formulate the near-term objective function $Q$, the dynamic variables $x^*(n+1)$ and $y(n)$ must be related to the static variables $x$ and $y$. We recall that in the static problem, the perturbation equation is:

$$\delta x \delta y = \frac{\partial f(y)}{\partial y} \delta y^2. \tag{2.25}$$

By comparing Eqs. 2.23-2.25, the near-term objective is formed by substituting $y(n)$ and $(\tau/T)x^*(n+1)$ for $y$ and $x$, respectively into Eq. 2.8 yielding:

$$Q = Q(x^*(n+1), y(n)). \tag{2.26}$$

For instance, if $J = (1/2)(x^2 + 4y^2)$ as in Sect. 2.1.2, the near-term function becomes $Q = (1/2)[(\tau/T)^2 x^*(n+1)^2 + 4y(n)^2]$. This type of transformation applies to any long-term objective $J$ where the plant equations are defined in the form of Eq. 2.21.

Finally, the static Hebbian covariance rule (Eq. 2.16) is transformed into a dynamic one by substituting $Q$, $x^*(n+1)$, and $y(n)$ for $J$, $x$, and $y$, respectively:

$$\delta W(n+1) = -k x^*(n+1) \cdot \frac{\partial Q}{\partial x^*(n+1)} \cdot \left[ \delta x^*(n+1)\, \delta y(n) + \frac{\partial Q}{\partial y(n)} \Big/ \frac{\partial Q}{\partial x^*(n+1)} \cdot \delta^2 y(n) \right]. \tag{2.27}$$

This transformation, while not unique, illustrates one systematic method for mapping long-term into near-term criterion functions.

### 2.2.2 Linear Examples

#### 1st-Order Linear System

We present a 1st-order linear example to illustrate the algorithm and demonstrate its performance, including its convergence and robustness properties. The dynamic equation is as follows:

$$\frac{dx}{dt} = \theta_o - ax + b\,(y - \theta_1)\,, \tag{2.28}$$

where $\theta_o$, $\theta_1$, $a$ and $b$ are all unknown, possibly slow time-varying parameters. In this case, we choose to minimize the long-term objective function $J = (1/2)\,(x^2 + y^2)$. Recognizing that the time constant, $\tau$, is $1/a$, one may substitute dynamic variables into $J(x, y)$ (cf. Eqs. 2.21 and 2.26) to derive the near-term objective function:

$$Q = \left(\frac{1}{Ta}\right)^2 \frac{x^*(n+1)^2}{2} + \frac{y(n)^2}{2}, \tag{2.29}$$

where from Eq. 2.24, $x^*(n+1)$ is:

$$x^*(n+1) = x(n+1) + (Ta - 1)x(n). \tag{2.30}$$

From Eqs. 2.27 and 2.29, a stable adaptation rule is:

$$\delta W(n+1) = -k\left[\delta x^*(n+1)\delta y(n) + (Ta)^2 \frac{y(n)}{x^*(n+1)}\delta y(n)^2\right]. \tag{2.31}$$

Setting this equation to zero, we obtain the steady-state solutions $W = -b/a$, $y = (b^2 - b\theta_o)/(b^2 + a^2)$ and $x = y/W$ which minimize both $Q$ and $J$.

In deriving this adaptation rule, it was assumed that all the parameters are unknown and possibly time-varying. The adaptation rule, however, includes one unknown quantity, $Ta$. The time step for the discretization, $T$, is given, so this poses no barrier to the implementation of the adaptation rule. On the other hand, the system parameter $a$, which is related to the time constant of the system, must be known. This explicit dependence is a direct result of the transformation from the long-term to the near-term objective function. There are several ways to deal with

this dependence (see the following section); for now we simply constrain this system to have a known, though possibly time-varying parameter, $a$.

A simulation of this Hebbian covariance adaptation rule is shown in Fig. 2-2. In this example, the parameter $b$ is stepped from -0.5 to -0.75 at time=10 sec. The asymptotic convergence to the new optimal values is demonstrated with and without the augmented adaptation equations. With all the covariance terms included, the system settles to the optimal steady states within 40 sec while the simplified adaptation rule (without all the covariance terms) converges to the wrong optimum, presumably due to a constant error associated with recirculation in the system. In Fig. 2-3, the system response to square-wave variations in the unknown parameter $b$ is illustrated.

We examine the robustness of the Hebbian covariance adaptation rule by adding uniform distributed noise to the state variable $x$. Simulations show that two adaptation parameters, the adaptation rate, $k$, and the maximum perturbation amplitude, $\Delta y$, control the rate of the asymptotic convergence as well as the fluctuations about the optimum. As seen in Fig. 2-4, the Hebbian controller effectively optimizes the system despite the added noise while $k$ determines the rate of convergence and the size of the fluctuations. In fact, the convergence can be made arbitrarily fast by increasing $k$, but large $k$'s also lead to increased sensitivity to noise disturbance. Hence, a trade-off with respect to $k$ exists between the convergence rate and the robustness to noise perturbations. As seen in Fig. 2-5, a similar trade-off also exists with respect to $\Delta y$.

### 2nd-Order Linear System

In the above example, the transformation from the long-term to the near-term cost function requires knowledge of some system parameters, or at least some combination of the parameters, in order to adapt to the optimal state. If these parameters are unknown and/or change with time, they must be estimated on-line. We show one such method in a 2nd-order linear system in which all the parameters are time-varying and unknown.

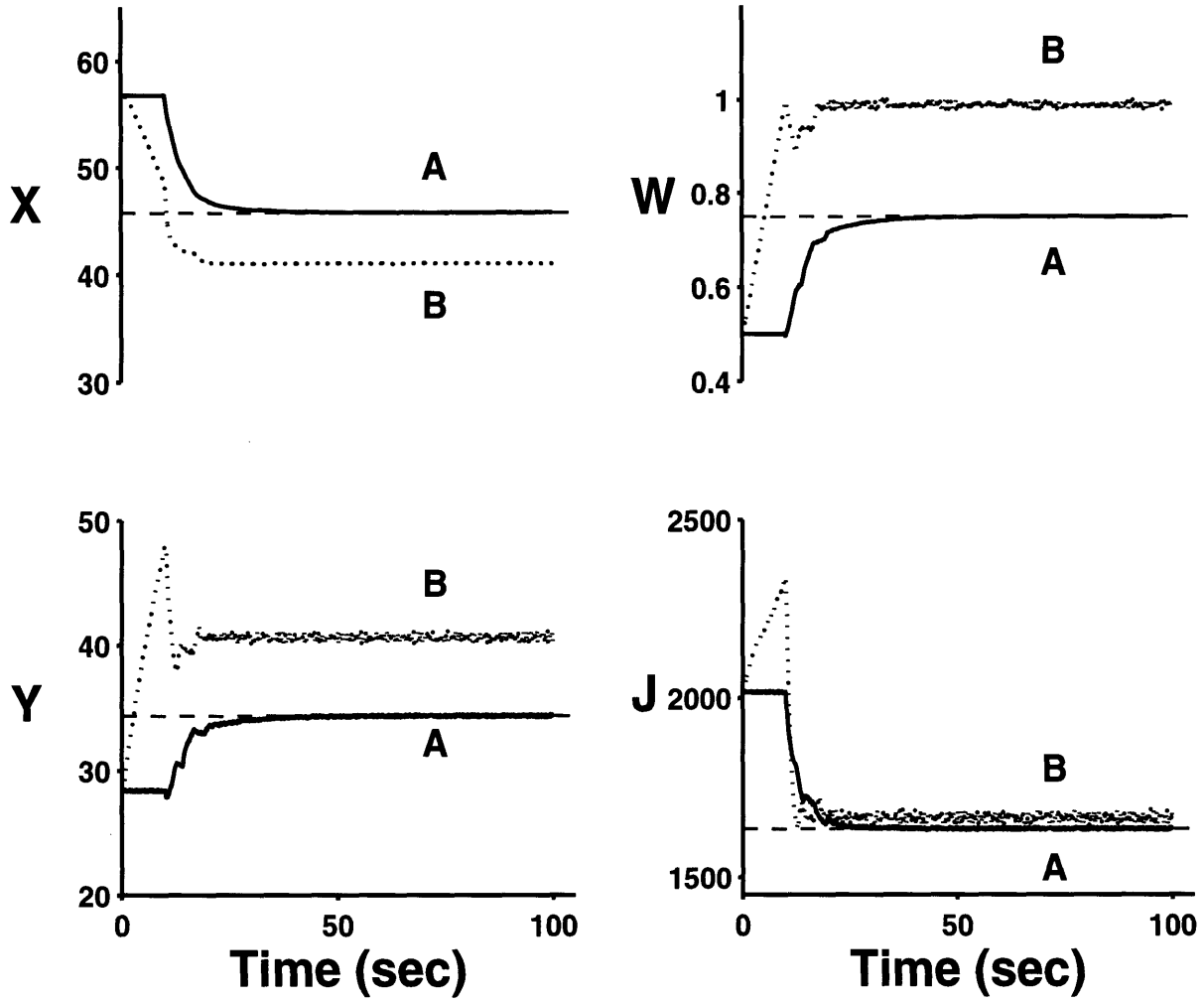


Figure 2-2: Simulation of the 1st-order linear example. These plots compare the performance of the Hebbian covariance rule with (A) and without (B) the augmented $x^*(n+1)$ term. The horizontal dashed lines indicate the optimal steady-state solutions. Unknown parameter $b$ is stepped from -0.5 to -0.75 at time=10 sec. Due to its simplified covariance rule, the (B) simulation is adapting even before the step change in $b$ because the initial state is not a steady-state solution. In this simulation, the parameters $\theta_o, \theta_1, a$ and $k$ are 70, 2, 1 and 0.4, respectively. Both simulations employ perturbations in $y$ of 0.1 maximum amplitude at every simulation time step (0.1 sec).

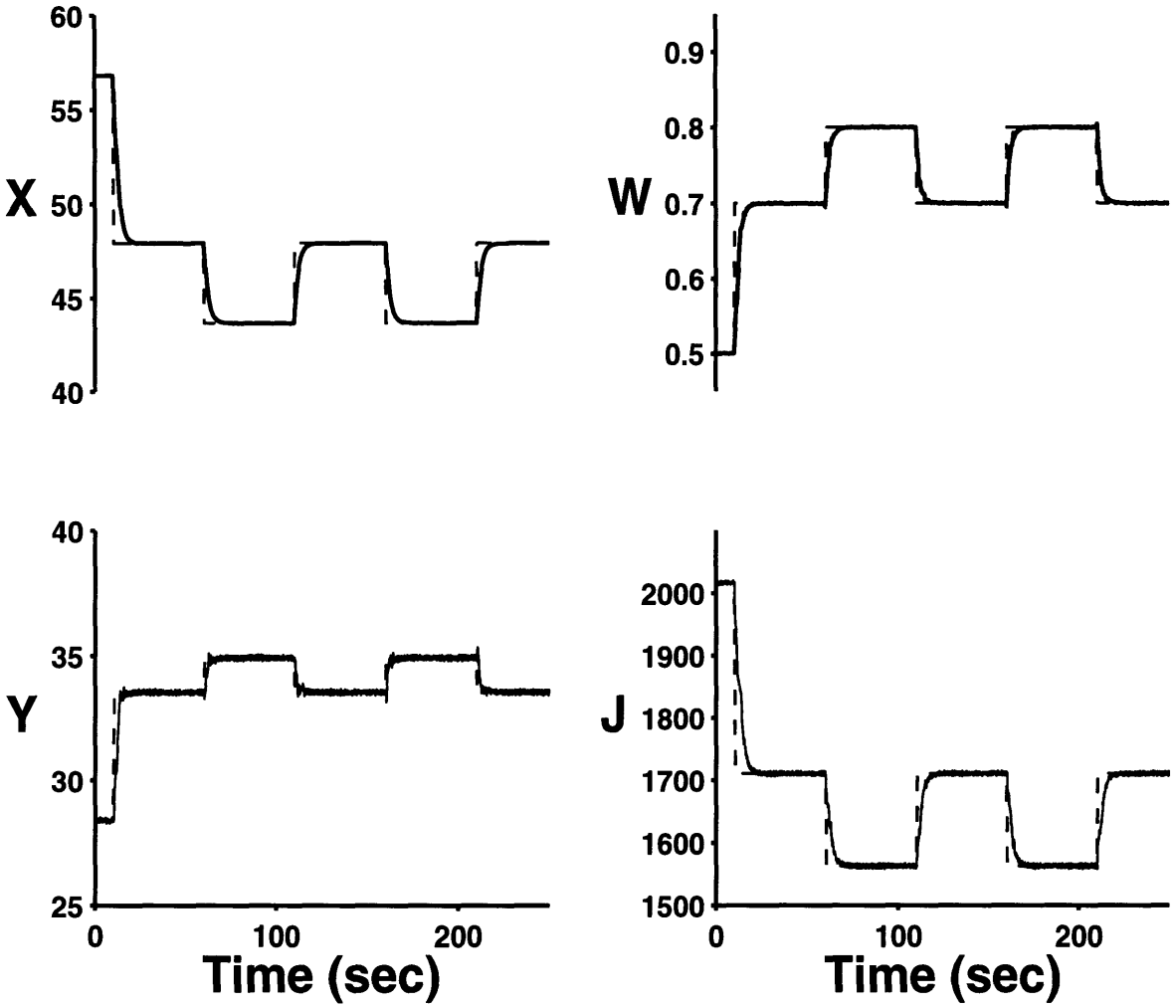


Figure 2-3: Simulation of the 1st-order linear adaptive system. In this case, after time=10 sec, $b$ is set to a square wave with mean -0.75, amplitude 0.05, and period of 100 sec. The dashed lines are the optimal values. The adaptation gain, $k$, is 1.

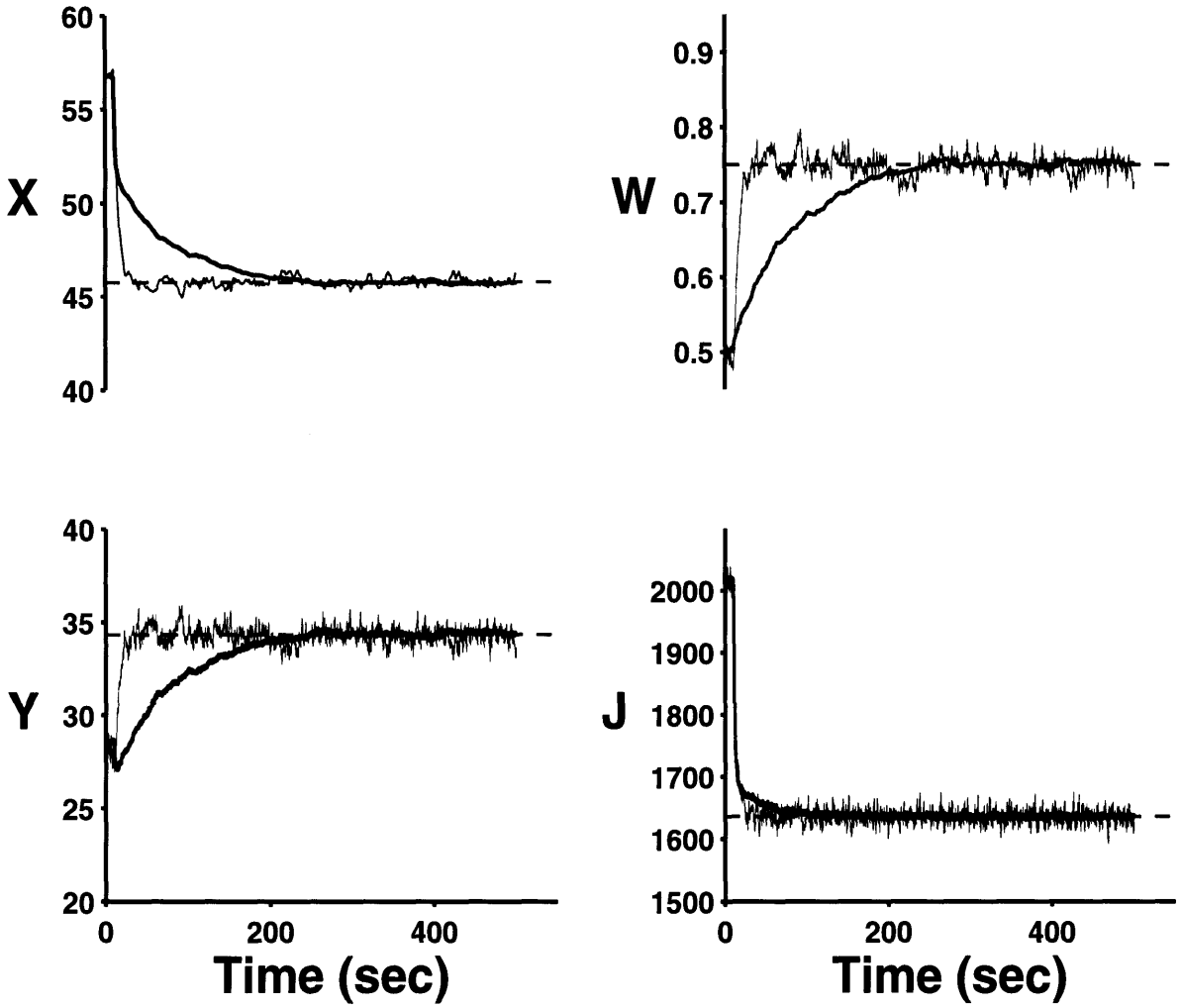


Figure 2-4: These simulations of the 1st-order system demonstrate the robustness of the adaptive paradigm with noise disturbance added to the state variable $x$. Here we show the role of the adaptation gain, $k$. One can see clearly the trade-off between the speed of convergence and the amount of fluctuations about the optimal state. The heavier and lighter traces in each panel correspond to $k = 0.1$ and $k = 50$, respectively, while the dashed lines represent the optimal steady-state values after a step change in $b$. The remaining parameters are the same as those defined for Fig. 2-2.

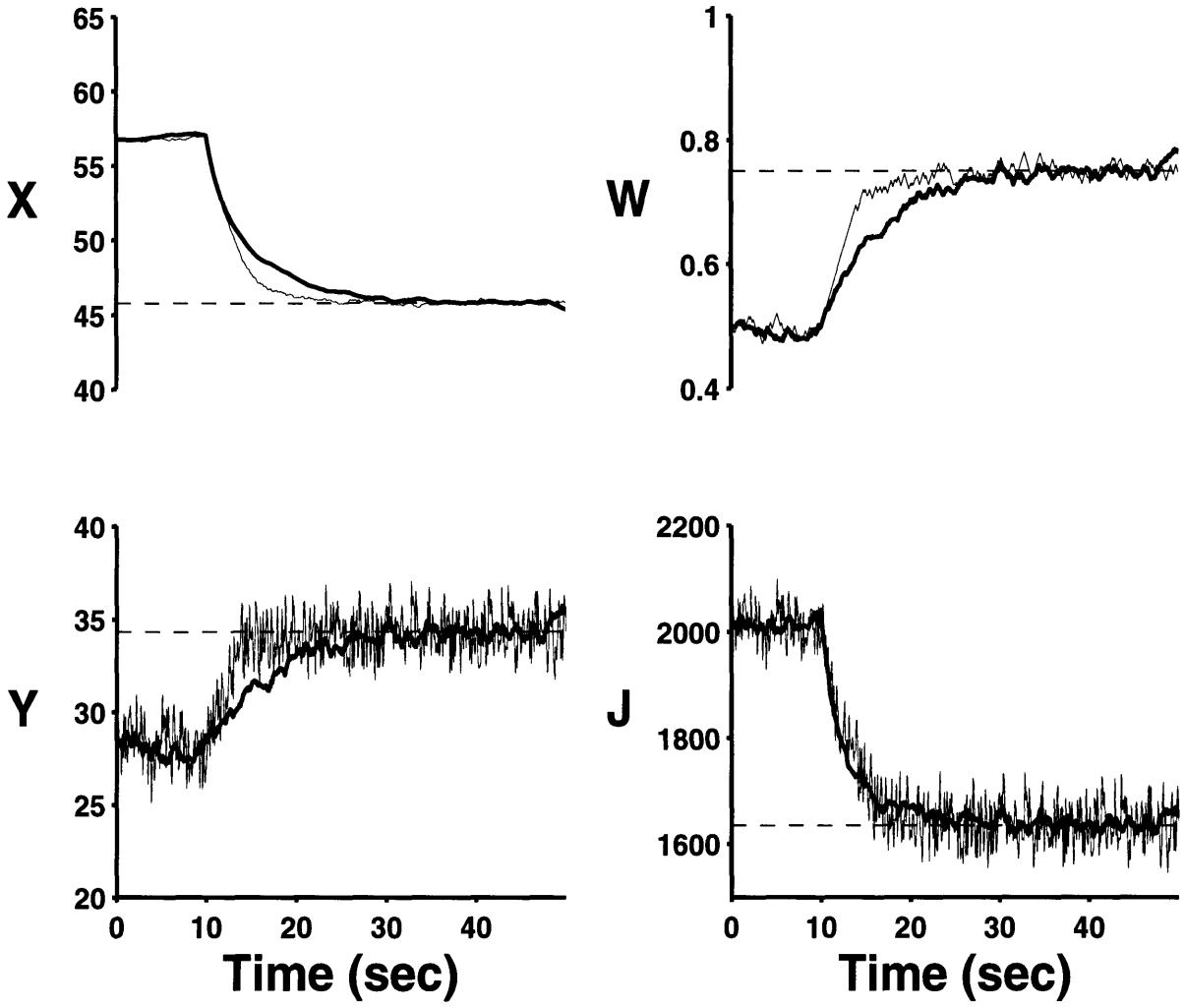


Figure 2-5: Simulation of the 1st-order linear system with noise disturbance corrupting the reinforcement signal $x$. This figure illustrates the role of the perturbation amplitude, $\Delta y$, on the rate of convergence and fluctuations about the optimal values (dashed lines). Two different maximum perturbation amplitudes are used; the lighter traces in each panel correspond to $\Delta y = 2.0$ and the darker to $\Delta y = 0.1$. All other parameters are the same as defined in Fig. 2-2.

Let the dynamic system be defined as follows:

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}u + \mathbf{S} \\ y &= \mathbf{C}\mathbf{x} \end{aligned} \tag{2.32}$$

with the following matrix definitions:

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \mathbf{S} = \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 0 & 1 \end{bmatrix}. \tag{2.33}$$

In this formulation, we have two state variables ($x_1$ and $x_2$), one feedback input ($u$) and two external disturbances ($S_1$ and $S_2$). The feedback signal is given by $u = Wy$. Assuming a long-term objective function, $J = (1/2)\,(x_2^2 + u^2)$, and using the method outlined in Eqs. 2.22-2.27, we derive the following near-term objective function:

$$Q = \left(\frac{1}{T^2(ad - cb)}\right)^2 \frac{x_2^*(n+2)^2}{2} + \frac{u^2(n)}{2} \tag{2.34}$$

and the corresponding Hebbian covariance adaptation rule:

$$\delta W(n+2) = -k\left[\delta x_2^*(n+2)\delta u(n) + T^4(ad - cb)^2 \frac{u(n)}{x_2^2(n+2)}\delta u(n)^2\right]. \tag{2.35}$$

In the above equations, $x^*(n+2)$ is defined as:

$$\begin{aligned} x^*(n+2) &= x_2(n+2) + x_2(n+1)\,[T(a+d) - 2] + \\ &\quad x_2(n)\left[T^2(ad - cb) - T(a+d) + 1\right]. \end{aligned} \tag{2.36}$$

Solving Eq. 2.35 when $\delta W(n+2) = 0$, we find the optimal solutions to be $u = [c^2(S_1 d + S_2 b) + cS_2(ad - cb)]/[d(ad - cb)^2 + c^2 d]$ and $W = c/(ad - cb)$.

To implement this adaptive paradigm, several system quantities must be given or estimated. In particular, from Eqs. 2.35 and 2.36, two terms must be estimated: 1) the determinant of the system $(ad - cb)$ and 2) the trace $(a + d)$. Note that while the individual parameters $a$, $b$, and $d$ are not identifiable, the determinant and the trace

can be estimated. We begin by deriving the following discrete relationship:

$$\delta x_2(n+2) = \Theta_1 \delta u(n) + \Theta_2 \delta x_2(n+1) + \Theta_3 \delta x_2(n), \tag{2.37}$$

where the parameters $\Theta_i$ define the determinant, the trace, and $c$ in the following ways:

$$|\mathbf{A}| = \frac{-\Theta_2 - \Theta_3 + 1}{T^2} \qquad trace(\mathbf{A}) = \frac{\Theta_2 - 2}{T} \qquad c = -\frac{\Theta_1}{T^2}. \tag{2.38}$$

Thus, the goal is to estimate these three parameters on-line using a covariance estimation rule. By applying the acclaimed MIT rule [3]:

$$\frac{d\boldsymbol{\Theta}}{dt} = -\gamma e \frac{\partial e}{\partial \boldsymbol{\Theta}}, \tag{2.39}$$

where $\gamma$ is the adaptation gain and $e$ in the error in the predicted state, Eq. 2.37 yields the following Hebbian-like covariance estimation rules:

$$\begin{aligned}
\delta\hat{\Theta}_1(n+2) &= -\gamma_1 \left[\delta\hat{x}_2(n+2)\delta u(n) - \delta x_2(n+2)\delta u(n)\right] \\
\delta\hat{\Theta}_2(n+2) &= -\gamma_2 \left[\delta\hat{x}_2(n+2)\delta x_2(n+1) - \delta x_2(n+2)\delta x_2(n+1)\right] \\
\delta\hat{\Theta}_3(n+2) &= -\gamma_3 \left[\delta\hat{x}_2(n+2)\delta x_2(n) - \delta x_2(n+2)\delta x_2(n)\right],
\end{aligned} \tag{2.40}$$

where $\hat{\Theta}_i$ are the estimated parameters; $\gamma_i$ are the adaptation gains; and $\delta\hat{x}_2(n+2)$ is the estimated variance in $x_2$ calculated from Eq. 2.37 and the current estimates for $\Theta_i$.

This adaptive system is simulated with and without random noise disturbance added to the state variable $x_2$. Figure 2-6 shows the convergence of both the system states to the optimal solution and the parameter estimates to the actual model parameters, $\Theta_i$, in a disturbance-free environment.

After a step change in $c$ at time=100 sec, the system's optimal state is changed. Clearly, $W$ is appropriately tuned by the adaptation so that the objective function, $J$ is minimized. Note that although the trace of the system has not yet converged, the Hebbian covariance rule is still able to find the optimal operating states. With

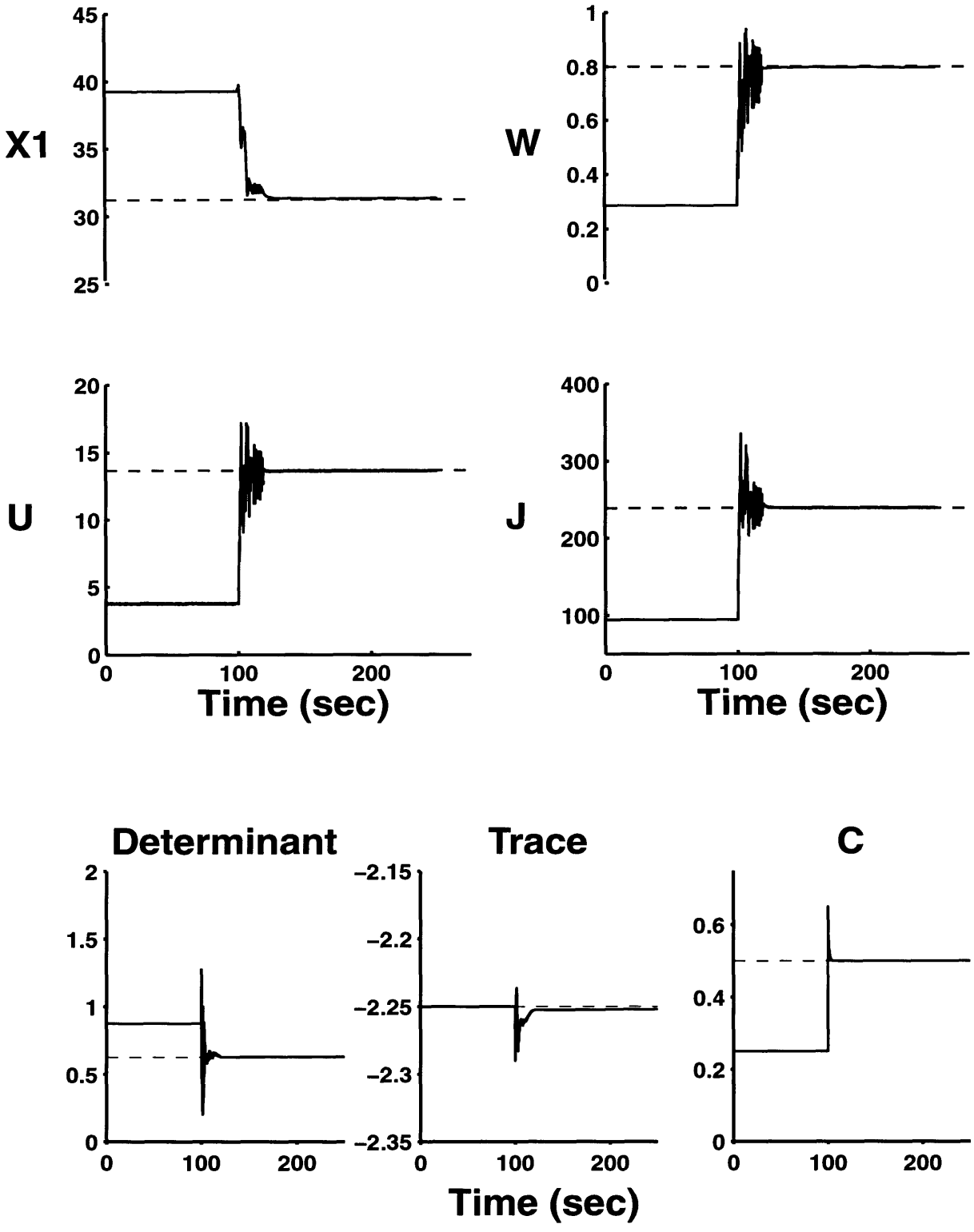


Figure 2-6: Simulation of the 2nd-order linear example with Hebbian covariance optimization and a Hebbian-like MIT rule used for parameter estimation. Convergence to the optimal states (dashed lines) is illustrated. In this simulation, the parameter $c$ is stepped from 0.25 to 0.5 at time = 100 sec while the other parameters were defined as $a = 0.75$, $b = 1$, $d = 1.5$, $k = 150$, $\Delta y = 0.05$, $\gamma_1 = 0.1$, $\gamma_2 = 10$ and $\gamma_3 = 10$.

disturbance added in the environment (Fig. 2-7), the adaptive system is still able to estimate the model parameters and move the system to the optimum.

### 2.2.3 Non-Linear Examples

#### Non-Linear System 1

Using the same steady-state objective function as in a prior static example (Sect. 2.1.2), $J = -x^2/2$, the static relation $x = f(y)$ is now replaced by the dynamic non-linear equation:

$$\tau \frac{dx}{dt} = \theta_0 - x - (y - \theta_1)^2, \tag{2.41}$$

where $\tau$ is a time constant.

To find the near-term objective function, we first rewrite Eq. 2.41 in discrete form as:

$$\frac{\tau}{T}[x(n+1) - x(n)] = \theta_0 - x(n) - (y(n) - \theta_1)^2. \tag{2.42}$$

Assuming random perturbations in the controller output $y$, Eq. 2.42 yields:

$$\frac{\tau}{T}\delta x^*(n+1)\,\delta y(n) = -2\,(y(n) - \theta_1)\,\delta y(n)^2, \tag{2.43}$$

where $x^*(n+1) = x(n+1) + (T/\tau - 1)x(n)$. Using the theory presented in Sect. 2.2.1, the following near-term objective function is found:

$$Q(n+1) = -\left(\frac{\tau}{T}\right)^2 \cdot \frac{x^*(n+1)^2}{2}. \tag{2.44}$$

Application of Eq. 2.27 yields the following near-term adaptation rule:

$$\delta W(n+1) = k \cdot \left(\frac{\tau}{T}\right)^2 \cdot x^*(n+1)^2 \cdot \delta x^*(n+1) \cdot \delta y(n), \tag{2.45}$$

which, in steady state, will satisfy the long-term objective function $J$. This is seen by comparing Eqs. 2.18 and 2.45 and by noting that $x(n+1) = x(n) = x$ in the steady state. Likewise, we find the same steady-state solutions as in the static case: $x = \theta_0$,

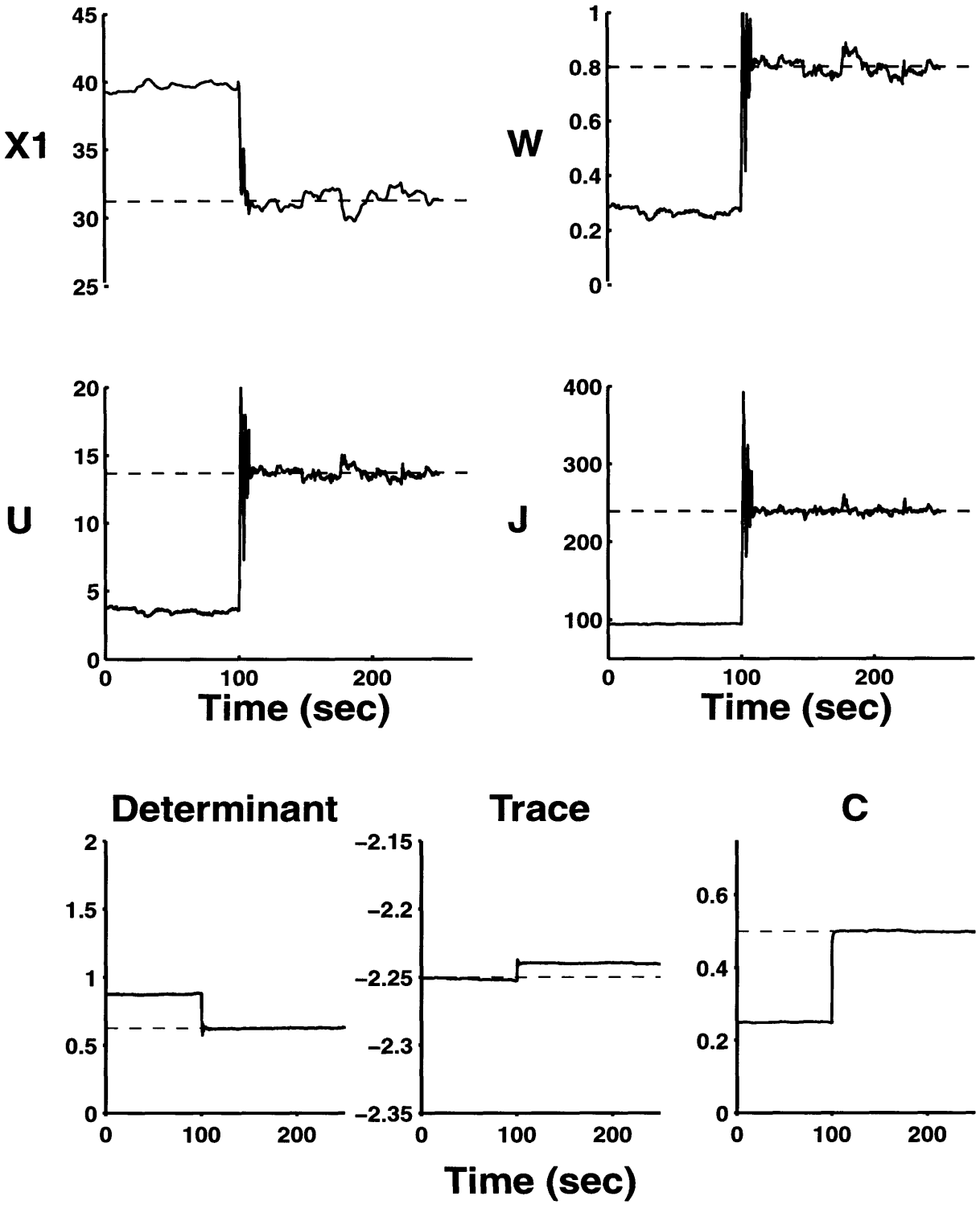


Figure 2-7: Robust analysis of the Hebbian covariance adaptation rule and the Hebbian-like covariance estimation rules. The 2nd- order linear system was simulated with noise added to the reinforcement signal $x$. The adaptation rates for the adaptive estimator were $\gamma_1 = 0.1$, $\gamma_2 = 1$ and $\gamma_3 = 1$ and the Hebbian adaptation gain $k = 100$. All other parameters were identical to the simulation shown in Fig. 2-6.

$y = \theta_1$, and $W = \theta_1/\theta_0$.

This non-linear adaptive system is simulated assuming that only $\tau$ is known. In Fig. 2-8, simultaneous step changes in $\theta_o$ and $\theta_1$ are made from 70→80 and 2→5, respectively. The system adapts quickly to the new conditions to find the optimal states (dashed lines).

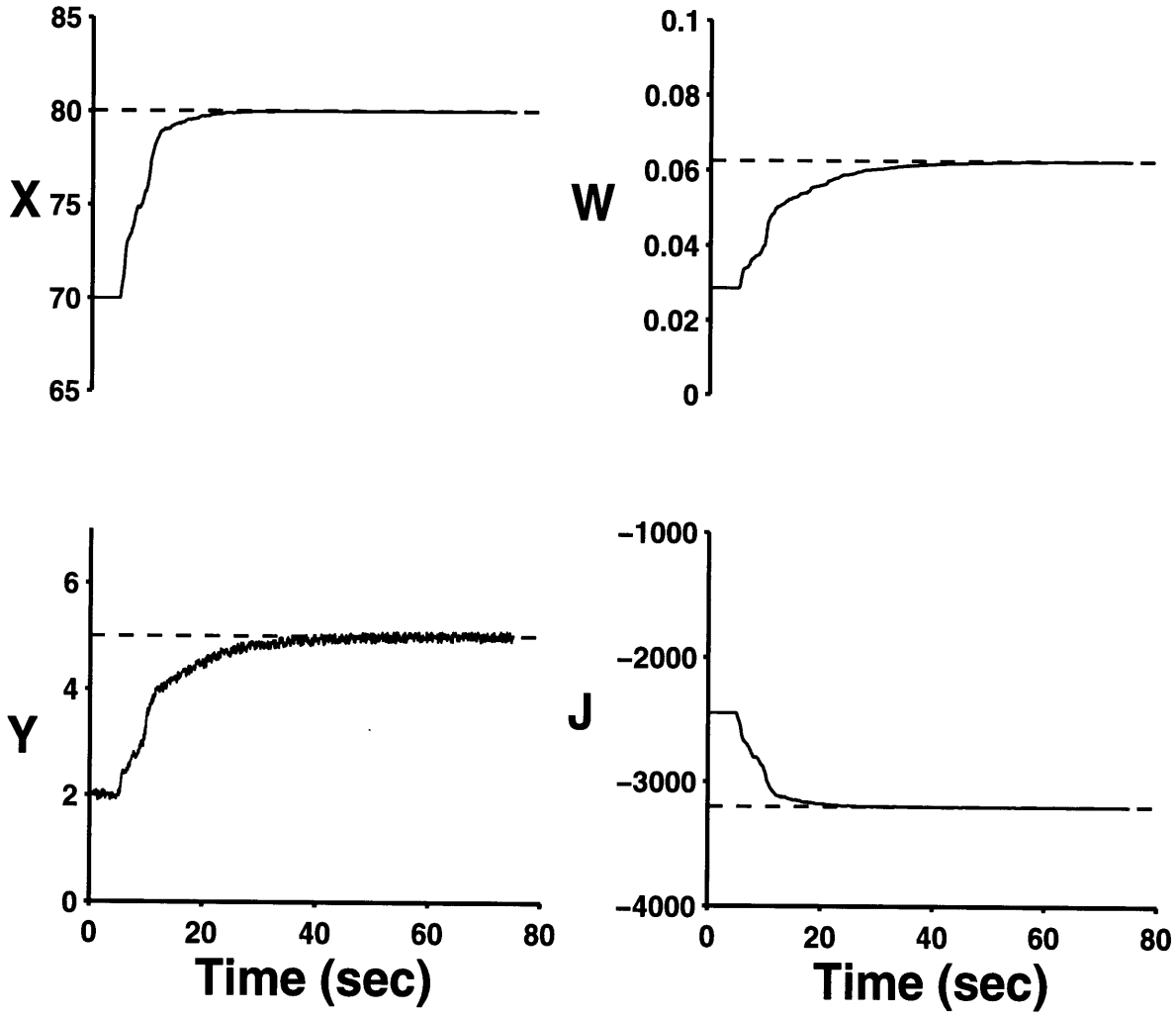


Figure 2-8: Simulation of the first dynamic non-linear system. Step changes in parameters $\theta_o$ and $\theta_1$ were made at time = 5 sec (from 70 to 80 and from 2 to 5, respectively). The new optimal solution is indicated by the dashed lines. The adaptation gain, $k$, equals 0.025 and the maximum perturbation amplitude, $\Delta y$, equals 0.1.

### Non-Linear System 2

In the above non-linear example the steady-state optimal solution was unique and attainable if both parameters $\theta_1$ and $\theta_2$ are known. However, this may not always be the case. For example consider the function $x = f(y) = \theta_o/y + \theta_1$ and the objective function $J = (1/2)(x^2+y^2)$. A corresponding 1st-order non-linear differential equation is:

$$\tau \frac{dx}{dt} = -xy + \theta_1 y + \theta_0, \tag{2.46}$$

where the time constant of the system is now $\tau/y$. From the theory in Sect. 2.2.1, the corresponding near-term objective function is derived as:

$$Q = \left(\frac{\tau}{T}\right)^2 \frac{x^*(n+1)^2}{2} + \frac{y(n)^2}{2}, \tag{2.47}$$

where

$$x^*(n+1) = \frac{x(n+1)}{y(n)} + \left(\frac{T}{\tau} - \frac{1}{y(n)}\right) x(n). \tag{2.48}$$

After deriving the adaptation rule using Eq. 2.27, we solve for the following steady-state solutions:

$$y^4 - \theta_0\theta_1 y - \theta_0^2 = 0 \tag{2.49}$$

$$x^4 - 3\theta_1 x^3 + 3\theta_1^2 x^2 - \theta_1^3 x - \theta_0^2 = 0. \tag{2.50}$$

Clearly there are multiple solutions satisfying these equations. As shown in Fig. 2-9, two real-valued minima (one local and the other global minimum) exist for certain values of $\theta_0$ and $\theta_1$. Simulation of this system (Fig. 2-10) illustrates that the local minimum is found after the step changes in the parameters at time=50 sec. The Hebbian covariance system solves this complex minimization problem in a direct manner without estimating the parameters $\theta_0$ and $\theta_1$. Although only the local minimum is found, the global minimum in this case is an unstable solution due to the formation of a positive eigenvalue when $y < 0$.

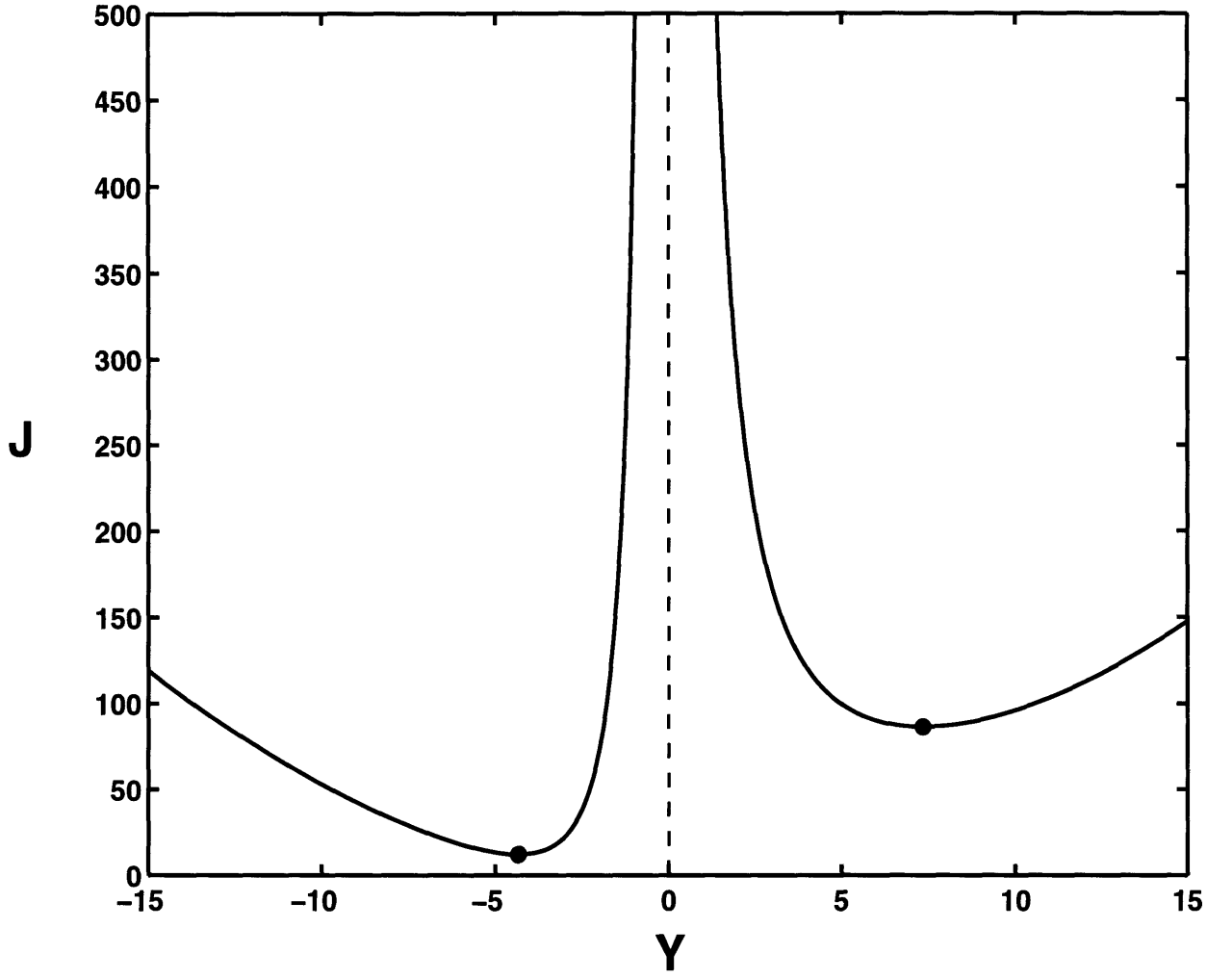


Figure 2-9: This figure illustrates the existence of two local minima for the objective function, $J = (1/2)(x^2 + y^2)$, in the second dynamic non-linear example. Here parameters $\theta_o$ and $\theta_1$ are 36 and 6, respectively. The two local minima of $J$ exist at $y$ equal to 7.32 and -4.35 (solid circles).

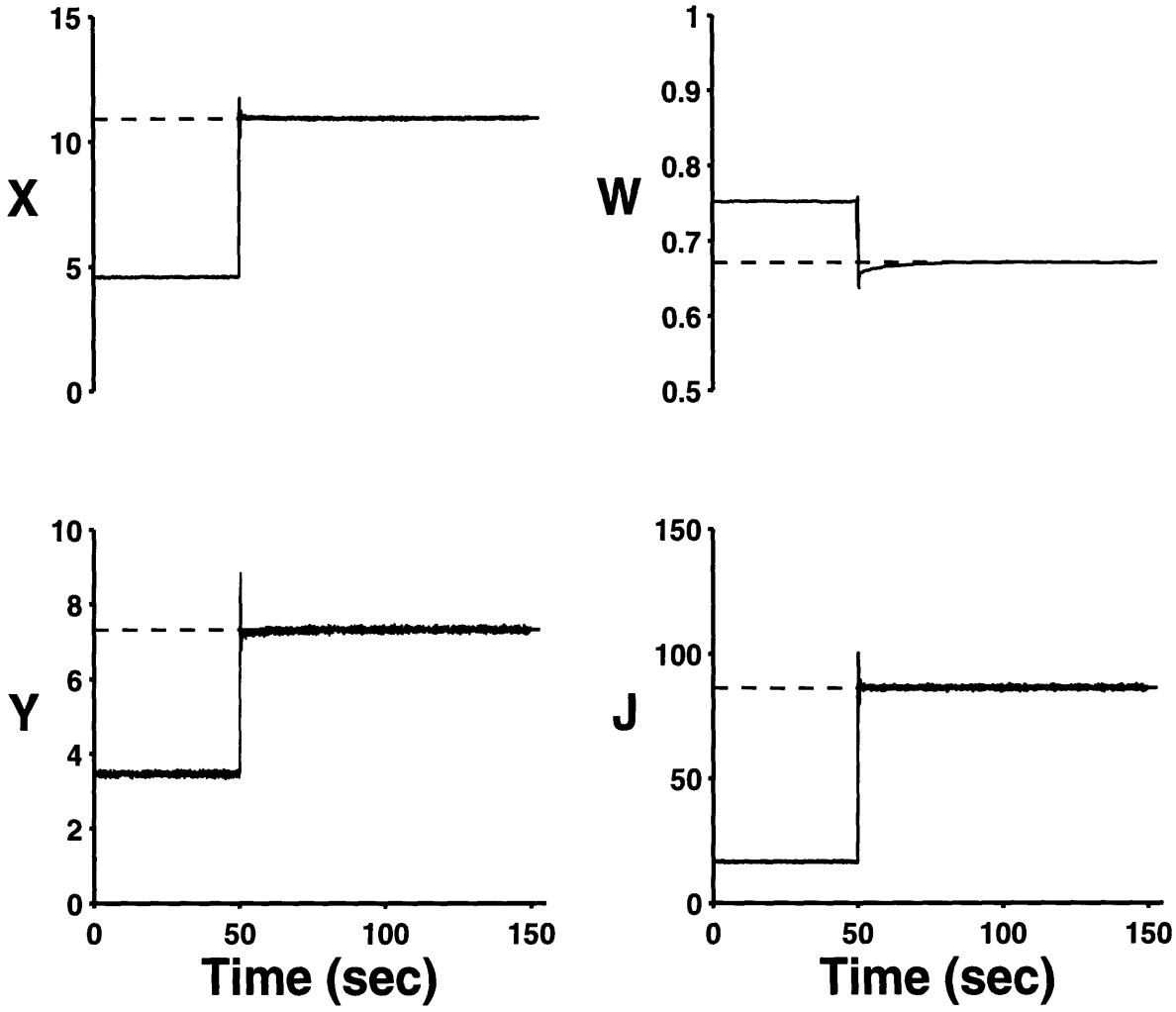


Figure 2-10: Simulation of the second dynamic non-linear system. Step changes in parameter $\theta_o$ and $\theta_1$ were made at time = 50 sec (from 9 to 36 and from 2 to 6, respectively). The new optimal values are illustrated by the dashed lines. The adaptation gain and the perturbation amplitude are 0.5 and 0.1, respectively.

# Chapter 3

# Hebbian Covariance Learning for Optimal Respiratory Control

## 3.1 Background

### 3.1.1 The Brainstem Controller

Respiration in mammals is a vital autonomic system that maintains homeostasis despite varying physiological and environmental conditions. The specialized brainstem center underling neural respiratory control is not solely autonomous, for voluntary control seizes command during speech, mastication and other behavioral tasks. Nevertheless, the respiratory controller can function even when deprived of input from the higher brain [49]. This brainstem controller continually regulates the blood pH and blood-gas tensions by integrating various afferent inputs, including chemosensitive, pulmonary stretch, and proprioceptive signals. As will be examined below, recent studies have hypothesized that this brainstem controller may intelligently adapt to different conditions whereby an optimal operating point is always maintained.

The respiratory control center is located throughout the medulla in discrete nuclei. Several specific regions have now gained consensus as being critical for setting the rate and depth of respiration as well as maintaining a regular respiratory pattern. Each neuronal group may generally be classified by their salient features: either

pattern generation/modulation centers, thereby comprising part of the central respiratory pattern generator (CRPG), or integrative/output centers. The CRPG is essentially composed of inhibitory neuronal groups whose activity rhythmically alternates with the inspiratory and expiratory cycles. Conversely, the close-loop behavior, often studied separately from pattern generation, is controlled by distinct nuclei which determine the total respiratory output.

The respiratory brainstem nuclei are primarily separated into the dorsal and ventral groups. The dorsal respiratory group (DRG) located throughout the nucleus tractus solitarius (NTS) is largely inspiratory related. Afferent inputs converge on the NTS from peripheral and central chemoreceptors, somatic pain receptors, pulmonary stretch receptors, and cortical and pontine centers. The DRG efferent output projects contralaterally via the phrenic and intercostal motoneurons to drive inspiration while other projections inhibit expiratory neurons of the ventral respiratory group (VRG).

The VRG comprises several nuclei involved in both inspiration and expiration. The nucleus ambiguus (NA) contains primarily inspiratory neurons while the nucleus retro-ambigualis (NRA) is separated into a caudal expiratory region and a rostral inspiratory region. The Botzinger's complex, the most rostral VRG nuclei, is solely expiratory and has been shown to inhibit both inspiratory NTS and phrenic motoneurons.

Mounting neurobiological evidence suggests that the neural mechanisms underling respiratory regulation are not hard wired but may in fact contain modifiable synaptic connections and even memory. Such synaptic modification is considered the primary mechanism for learning and memory in invertebrates [23] and vertebrates [8]. Early studies of respiratory memory focused on "afterdischarge" [41], a phenomena resulting from peripheral nerve stimulation [19]. Such persistent nerve activity was thought to result from a network of reverberating respiratory neurons [15].

Recently, however, short-term potentiation (STP) of respiratory drive has been linked to synaptic plasticity [18, 51]. New evidence suggests that NMDA receptors in the medulla may mediate STP [16] as well as long-term depression (LTD) [55].

The observed LTD is consistent with Hebbian covariance plasticity [10, 4] in which the pairing of pre- and post-synaptic activity may augment or diminish the efficacy of synaptic transmission. Hebbian covariance learning (see Section 2) is a neural mechanism heavily studied in the hippocampus and is thought to underlie classical associative and conditional learning [46].

The importance of NMDA receptors in respiratory control was confirmed vividly in newborn mice whose NMDA receptors were abolished by NMDAR1 gene knockout. Poon et al. (1994) showed that these neonate mice, healthy at birth, experienced respiratory failure and death within the first day of life.

Various disorders of respiratory control exist including two forms of sleep apnea which may lead to death. Perhaps most widely recognized, sudden infant death syndrome (SIDS) is characterized by a sudden cessation of breath during sleep in newborns. This phenomena may result from either a decreased central sensitivity to $CO_2$ in the arterial circulation ($P_{aCO_2}$) or a decreased number of chemosensitive cells [48]. Ondine's curse, however, typically occurs in adults during sleep. Patients with this condition must voluntarily increase ventilation during bouts of apnea. Less critical respiratory control disorders include Cheyne-Stokes breathing in which periodic increases and decreases in ventilation are caused by head injury or cardiac failure.

### 3.1.2 Respiratory Control Hypotheses

Traditionally, black-box reflex control theories [41] have dominated studies of homeostatic control, including respiration, due to their conformity with classical control theories. However, recent evidence has weakened these feedback/feedforward models of respiration propelling researches to search for more intelligent brain control strategies.

It is important to recognize how the classical respiratory control models fail in explaining critical respiratory phenomena. Two phenomena are particularly illuminating: exercise hyperpnea and the hypercapnic response to $CO_2$ inhalation. The basic ventilatory responses to these conditions are shown in Fig. 3-1. The steady-state exercise response is characterized by a large increase in the ventilation rate, $\dot{V}_E$,

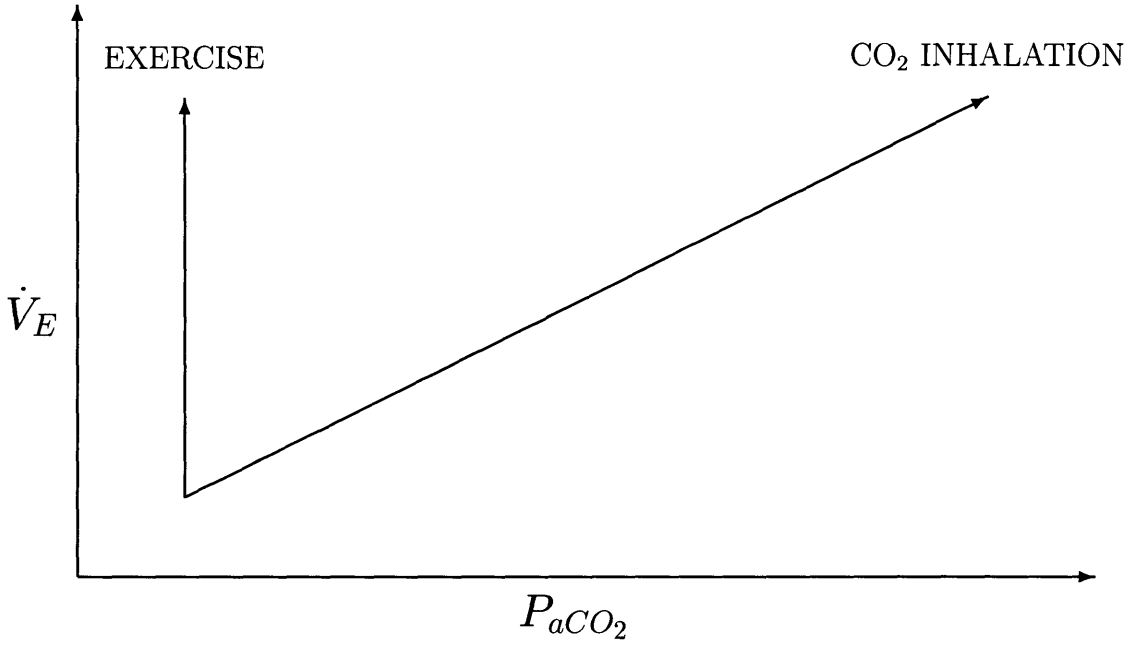


Figure 3-1: The dilemma of respiratory regulation. With increased metabolic $CO_2$ load during muscular exercise, arterial $CO_2$ tension ($P_{aCO_2}$) is regulated by the respiratory controller about a nearly constant operating point. Thus, homeostasis is maintained by the respiratory controller which increases respiratory output ($\dot{V}_E$) in the absence of any chemical feedback. With increased exogenous $CO_2$ load during inhalation, respiratory output increases only with increased chemical feedback and arterial homeostasis is abolished.

and essentially no deviation in blood $CO_2$ levels ($P_{aCO_2}$). This response effectively counteracts the increased metabolic $CO_2$ load which results from muscular activity (as well as maintaining proper blood pH levels). During exogenous $CO_2$ loading, the response is substantially different. Namely, $\dot{V}_E$ increases with the imbalance in $P_{aCO_2}$, thus never restoring the quiescent $P_{aCO_2}$ level. The slope of this curve is one measure of chemoreceptor $CO_2$ sensitivity. What type of controller can explain these divergent responses?

A classical theory for homeostasis asserts that set-points for biological parameters determine the body's behavior. For respiration, this model predicts that the body should maintain certain nominal states of blood gas concentrations. This model suffices to explain the respiratory response to exercise if there is a high feedback gain. However the set-point hypothesis fails to account for the hypercapnic response during $CO_2$ inhalation. Furthermore, such high-gain feedback-loops are often unstable in system with intrinsic time delays such as in respiration.

Reflexogenic models are another widely accepted theory for respiratory control. These hypotheses maintain that a finite gain in the feedback is responsible for the hypercapnic response during $CO_2$ inhalation. This is a reasonable assumption given that the slope of the response would be determined by this finite gain. However, this model fails to account for exercise hyperpnea. To overcome this dilemma, certain feedforward exercise stimuli have been proposed. Unfortunately, no signal has yet been confirmed [52].

Recent hypotheses entertain the notion that the respiratory controller may adapt to varying conditions. Such premises extend Cannon's farsighted theories of the "wisdom of the body"[13], in which intelligent control is proposed to underlie autonomic behavior. Various forms of adaptive control have been suggested to underlie homeostatic regulation [21, 37], however the neural mechanisms for such schemes remain uncertain.

Alternatively, optimal respiratory control theories have gained recent attention in explaining the body's apparent ability to maximize performance over varying physiological and environmental conditions [30, 31]. This hypothesis suggests that the aim

of respiration is to minimize an implicit objective function which can be expressed as the cost of the $CO_2$ divergence from a set-point and the energy consumed by the mechanical act of breathing. Such an optimal control objective conforms to exercise hyperpnea and hypercapnic respiratory responses [32, 34].

## 3.2 Theory

As discussed above, the respiratory system is responsible for maintaining chemical homeostasis of the blood by controlling a mechanical musculoskeletal system. The steady-state behavior of the chemical plant is described by the following nonlinear relationship [33]:

$$P_{aCO_2} = P_{iCO_2} + \frac{K \cdot \dot{V}_{CO_2}}{\dot{V}_E} \tag{3.1}$$

where $P_{aCO_2}$ and $P_{iCO_2}$ are the partial pressures of $CO_2$ in the arterial circulation and in the inspired air, respectively, $\dot{V}_{CO_2}$ is the production rate of $CO_2$ due to body metabolism, $\dot{V}_E$ is the ventilation rate, and $K$ is a constant.

Based on the theory of optimal respiratory regulation, the body is faced with minimizing the cost of the aberrant blood gas composition as well as the energy due to the motor act of breathing. Thus a long-term, physiological objective function $J$ may be defined as follows [30, 31, 36]:

$$J = J_c + J_m = [\alpha(P_{aCO_2} - \beta)]^2 + \ln \dot{V}_E^2, \tag{3.2}$$

where $J_c$ is the chemical cost of respiration expressed as the squared deviation of the $CO_2$ tension from the desired level, $J_m$ represents the mechanical cost of breathing expressed as a logarithmic function of the respiratory ventilation, and $\alpha$ and $\beta$ are sensitivity and threshold parameters for chemosensitivity, respectively.

The self-tuning Hebbian covariance learning system, as seen in Fig. 3-2, consists of a chemoafferent sensory neuron which senses $P_{aCO_2}$, an interneuron with a plastic synaptic weight, $W$, and an output motoneuron which drives the respiratory muscles. The interneuron input-output behavior is:

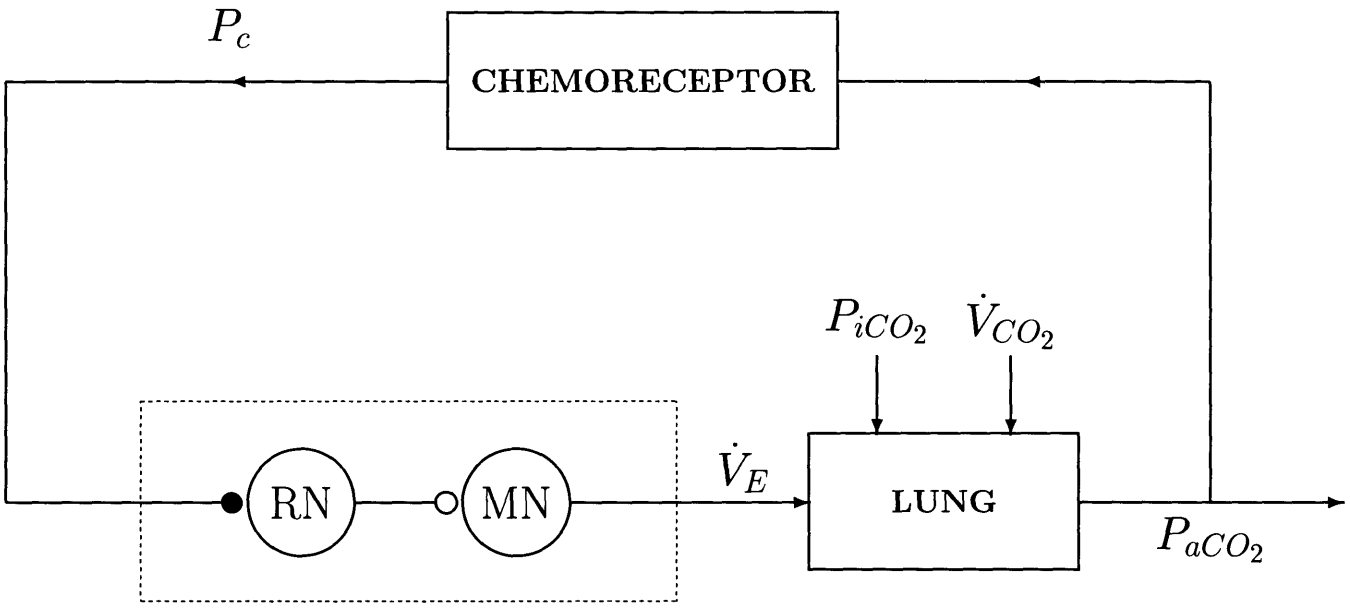


Figure 3-2: Self-tuning optimal regulator model of respiratory control using Hebbian covariance learning. The controller gain is adaptively regulated by synaptic plasticity in some respiratory interneuron (RN) which drives the motor neuron (MN). Synaptic potentiation and depression are governed by a Hebbian covariance rule which computes the optimal controller gain based upon correlated fluctuations in the motor outflow and chemoafferent feedback. The plastic synapse is represented by the filled circle, -●.

$$\dot{V}_E = W \cdot P_c, \tag{3.3}$$

where $P_c$ is the chemoafferent neural output.

Because of the complex transformations between static, physiological events ($P_{aCO_2}$) and dynamic, neural events ($P_c$), a near-term neural objective function $Q$ must be formulated that, in the steady state, will lead to the same optimal solution as $J$. This same transformation problem was illustrated in Section 2.2. The dynamic equations for the nonlinear respiratory system may be expressed as:

$$V_L \cdot \frac{dP_a}{dt} = P_{iCO_2} - P_a + \frac{K\dot{V}_{CO_2}}{\dot{V}_E} \tag{3.4}$$

and

$$\tau \cdot \frac{dP_c}{dt} = -P_c + \alpha[P_a - \beta], \tag{3.5}$$

where $V_L$ and $\tau$ are constants and $P_a$ is the instantaneous arterial $P_{aCO_2}$ (with $P_a = P_{aCO_2}$ in the steady state). Applying the dynamic theory presented in Sect. 2.2.1, we discretize these equations in time, include the spontaneous perturbation in the output ($\delta\dot{V}_E(n)$), and derive the resultant perturbation relationship:

$$\delta P_c^*(n+2)\delta\dot{V}_E(n) = -\left(\frac{T^2}{\tau V_L}\right)\frac{\alpha K\dot{V}_{CO_2}}{\dot{V}_E(n)^2}\delta\dot{V}_E(n)^2, \tag{3.6}$$

where the augmented term, $P_c^*(n+2)$, is defined as follows:

$$P_c^*(n+2) = P_c(n+2) + \left[\frac{T}{\tau} + \frac{T}{V_L} - 2\right]P_c(n+1) + \left[\frac{T}{V_L}\left(\frac{T}{\tau} - 1\right) - \frac{T}{\tau} + 1\right]P_c(n). \tag{3.7}$$

By comparing Eq. 3.6 to the following static perturbation equation:

$$\delta P_c \delta\dot{V}_E = -\frac{\alpha K\dot{V}_{CO_2}}{\dot{V}_E^2}\delta\dot{V}_E^2, \tag{3.8}$$

the near-term neural objective equation is derived as (c.f. Eq. 2.26):

$$Q(n+2) = \left(\frac{\tau V_L}{T^2}\right)^2 P_c^*(n+2)^2 + \ln\dot{V}_E(n)^2. \tag{3.9}$$

Finally the near-term adaptation rule which will minimize Eq. 3.9 is found by applying Eq. 2.27:

$$\delta W(n+2) = -k\left[\delta P_c^*(n+2)\delta\dot{V}_E(n) + \left(\frac{T^2}{\tau V_L}\right)^2 \frac{\delta\dot{V}_E(n)^2}{P_c^*(n+2)\dot{V}_E(n)}\right]. \tag{3.10}$$

To derive the steady-state near-term solution, we substitute Eq. 3.6 into Eq. 3.10, set $\delta W(n+2) = 0$ and substitute into the dynamic equations noting that $P_a(n+1) = P_a(n) = P_{aCO_2}$ at steady state. After some rearrangement, the following steady-state respiratory response is found:

$$\dot{V}_E = \alpha^2(P_{aCO_2} - \beta)K \cdot \dot{V}_{CO_2}, \tag{3.11}$$

where the optimal weight is $W = \alpha K\dot{V}_{CO_2}$. Equation 3.11 is identical to the optimal response corresponding to the physiological static objective function [34]. Thus this dynamic Hebbian covariance learning model should be compatible with both critical phenomena depicted in Fig. 3-1. Examination of Eq. 3.11 during exercise reveals that metabolic increases of $\dot{V}_{CO_2}$ directly leads to increases in $\dot{V}_E$. Thus no divergence in $P_{aCO_2}$ is experienced. During $CO_2$ inhalation, the response is less direct. Namely, a divergence in blood $CO_2$ must develop in order for ventilation to increase.

## 3.3 Results

This section presents the simulations of the self-tuning Hebbian covariance controller in the respiratory system. The results demonstrate the optimal behavior of the system in various conditions (exercise and $CO_2$ inhalation) and compare these results to known respiratory behavior. The robustness of the controller is also explored.

To simulate exercise we provide a step change in the metabolic load ($\dot{V}_{CO_2}$) from rest conditions. By examination of Eq. 3.11, a step change in $\dot{V}_{CO_2}$ from 0.2 to 1.0 l/min (at 20 seconds) should produce a similar five fold increase in the ventilation rate ($\dot{V}_E$), while $P_{aCO_2}$ should remain at the control level ($\sim$ 35 mmHg) in the steady

state. The controller increases ventilation by tuning the synaptic weight in the feedback loop until the optimal solution is attained. A typical simulation of the exercise response is shown in Fig. 3-3 where the optimal steady states are illustrated by the dashed line. The ventilation rate reaches roughly 95% of its optimal response within 3 minutes and continues to adapt asymptotically approaching the optimal solution. Small spontaneous random fluctuations in the output ($\Delta y$=0.015 l/min applied at every time step, $t_{step}$=0.05 sec) and the discrete adaptations in the synaptic weight are not readily visible at this scale. For this simulation, the adaptation rate $k$ was chosen as $5 * 10^6$, while time constants $V_L$ and $\tau$ were chosen to be 40 and 0.1, respectively.

$CO_2$ inhalation was simulated by a step increase in the inspired air ($P_{iCO_2}$). (The "intelligence" of the controller lies partly in its ability to distinguish between the two sources of $CO_2$.) The optimal respiratory response is to increase ventilation in response to the chemoafferent drive as illustrated by Eq. 3.11. As a result, the synaptic weight should remain constant in the steady state and steady-state $P_{aCO_2}$ should increase. After the step increase from 0 to 50 mmHg in $P_{iCO_2}$ at 20 seconds, the system responds to increased arterial $CO_2$ levels by increasing the ventilation rate from 5 to roughly 28 l/min (Fig. 3-4). By assessing the input-output relationship of the respiratory environment, after an initial increase, the synaptic weight optimally adapts, returning to the control level. Once again, roughly 95% convergence is reached in under 3 minutes. The simulation parameters are the same as in the previous exercise simulation.

We demonstrate that the model performance closely mirrors the body's responses to both exercise and exogenous $CO_2$ loading over a broad range of conditions. Figure 3-5 illustrates the optimal steady-state response of the model (cf. Fig. 3-1). The exercise response demonstrates the body's tendency to increase ventilation and maintain a nearly constant $P_{aCO_2}$ operating point without any additional chemical error feedback. On the other hand, during increased inspired $CO_2$, respiratory output increases with increased chemoafferent drive and homeostasis is abolished. Both of these optimal responses are predicted by the steady-state solution of the near-term Hebbian covariance learning rule (Eq. 3.11).

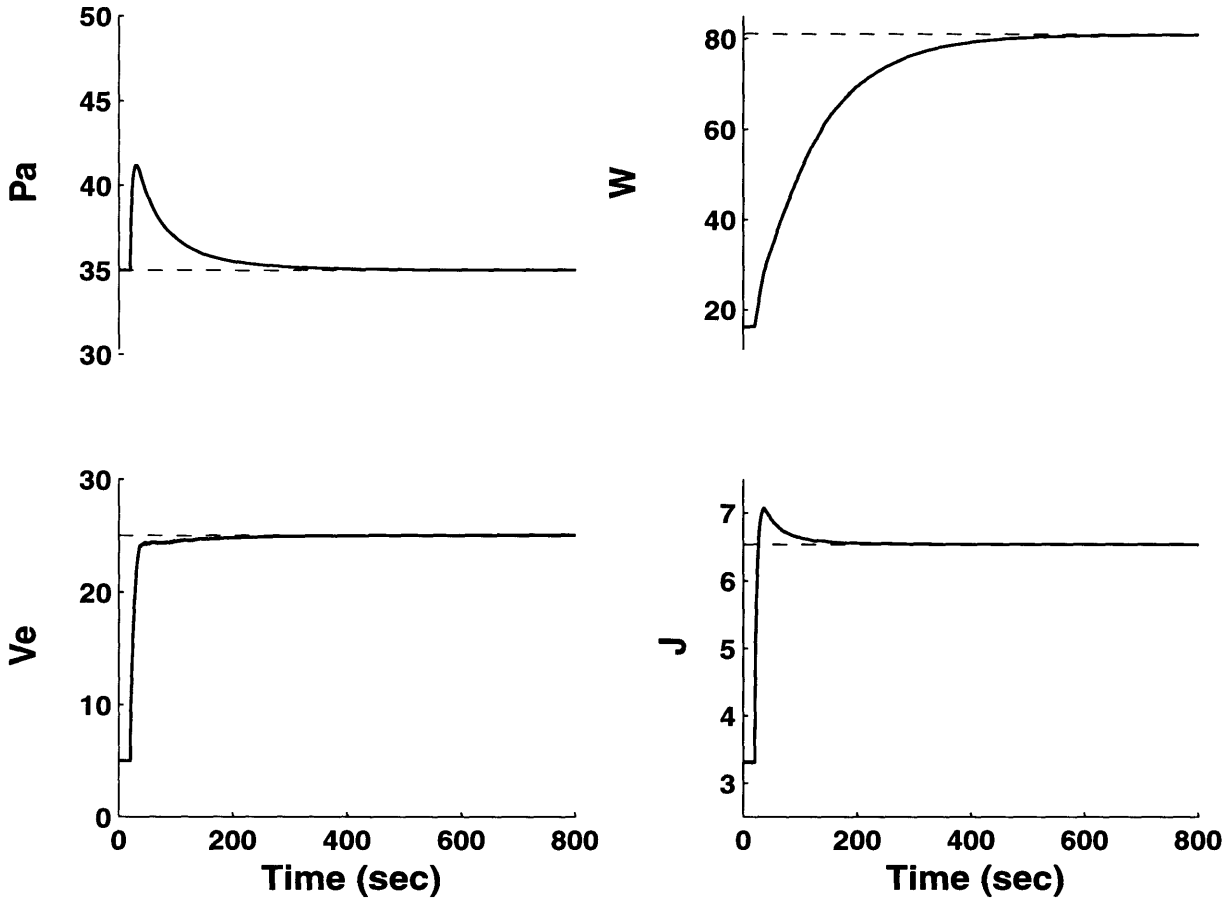


Figure 3-3: Computer simulation of exercise hyperpnea. A step change in $\dot{V}_{CO_2}$ was made from the quiescent level (0.2 l/min) to a moderate exercise state (1.0 l/min) at time=20 sec. The dashed lines illustrate the optimal weights which minimize the long-term criterion function. Random pulse perturbations were applied to the ventilation output with a maximum amplitude of 0.015 l/min at every time step (0.05 sec). The adaptation rate, $k$, was chosen to be $5*10^6$, while time constants, $V_L$ and $\tau$, were chosen to be 40 and 0.1, respectively. A maximum rate of change in $W$ was set to 0.8 units/sec.

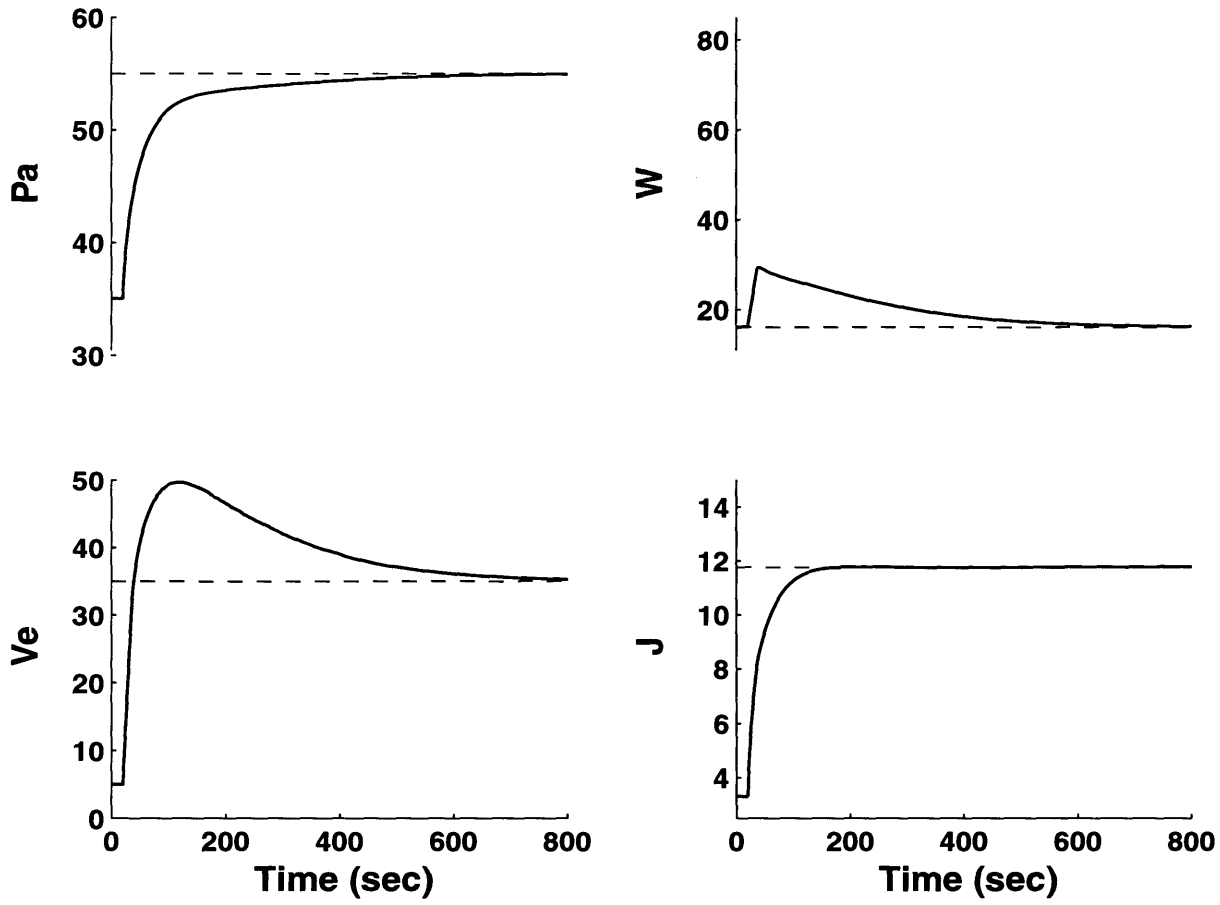


Figure 3-4: Computer simulation of $CO_2$ inhalation. A step change in $P_{iCO_2}$ was made from 0 to 50 mmHg at time=20 sec. The dashed lines illustrate the optimal weights which minimize the long-term criterion function. All the parameters were kept the same as in Fig. 3-3.

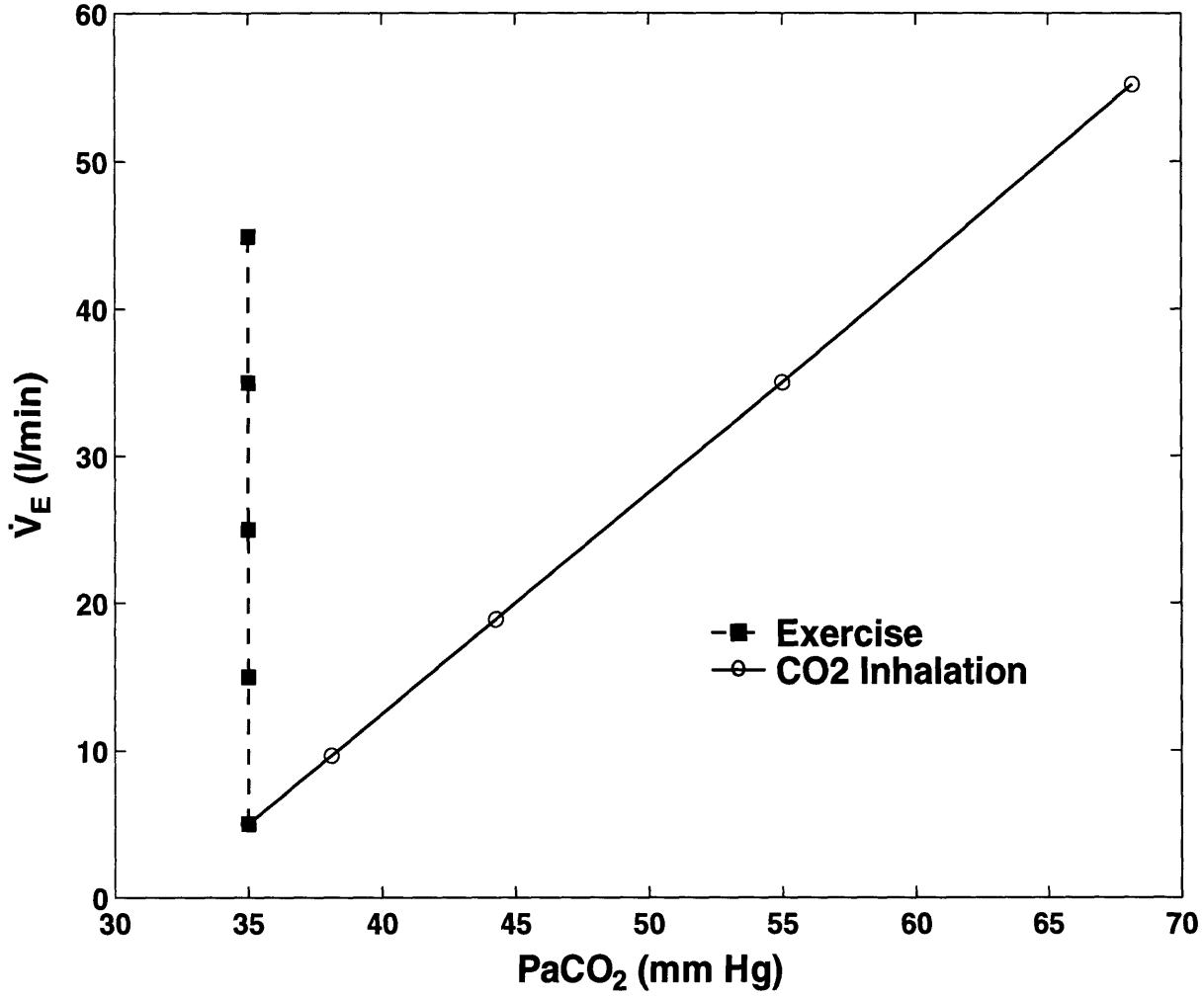


Figure 3-5: Optimal steady-state response of the self-tuning respiratory model under varying degrees of exercise hyperpnea and increased exogenous $CO_2$. The results closely resemble actual physiologic behavior shown in Fig. 3-1. The values for $\dot{V}_{CO_2}$ and $P_{iCO_2}$ were varied between 0.2 and 1.8 l/min and between 0 and 65 mmHg, respectively.

The physiological ventilatory response to exercise is tri-phasic [52]. Phase I is characterized by a rapid, almost instantaneous increase in the ventilation rate. Phase II is a slowly increasing transient lasting 4-5 minutes until a plateau level is reached (phase III). The previous exercise simulations employed a step change in $\dot{V}_{CO_2}$, however, the increase in metabolic $CO_2$ is likely more gradual in the body. This may be simulated by a simple ramp increase in $\dot{V}_{CO_2}$ over a 60 sec period. Figure 3-6 shows the ventilation rate increasing during exercise where trace 1 and 2 result from a step change and a ramp change, respectively. Trace 3 represents the difference between

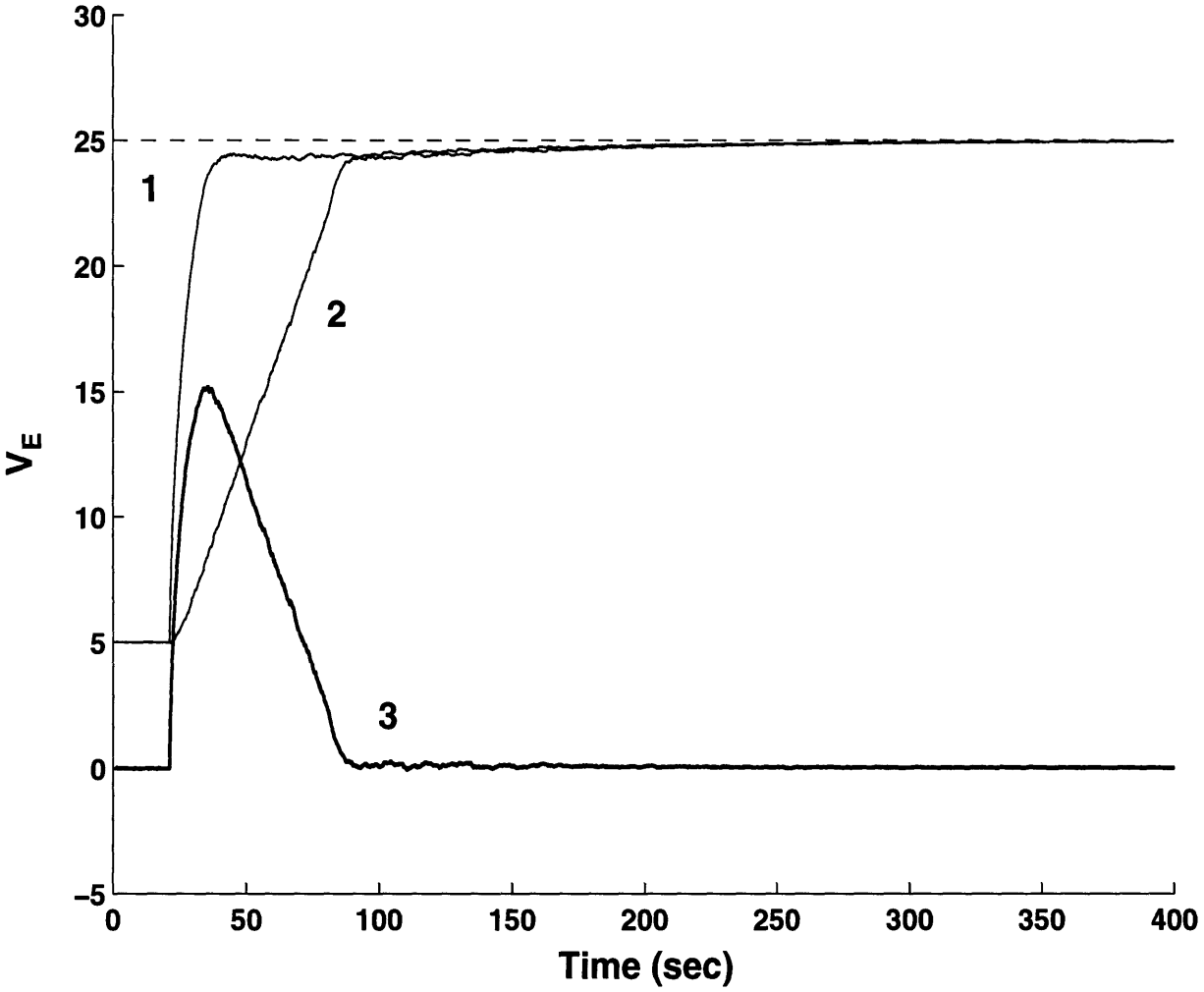


Figure 3-6: These simulations represent the ventilation responses during exercise for step (trace 1) and ramp (trace 2) changes in $\dot{V}_{CO_2}$. Trace 3 is the difference between the two responses and may represent an additional signal which may be available to the brainstem controller.

traces 1 and 2. We note that trace 1 possesses a characteristic phase I exercise response. To achieve such a rapid increase in ventilation during a ramp increase in $\dot{V}_{CO_2}$, the Hebbian covariance learning model requires an additional drive proportional to trace 3. This signal may be a learned response in the higher brain. Such a preemptive signal may improve the dynamic output of the Hebbian covariance learning system during the onset of exercise when $O_2$ demands are especially high.

As shown by the examples in Sections 2.2.2 and 2.2.3, the Hebbian covariance rules for optimization are robust to noise disturbances. We reexamine the robustness properties in the respiratory model to understand how similar disturbances in the body may result in changes in the physiological responses. Figure 3-7 illustrates the role of the perturbation amplitude ($\Delta y$) in determining the exercise response during random noise disturbance in $P_c$. Trace 1 corresponds to a ratio of 150/1 (perturbation to disturbance amplitude). Trace 2 maintains the same $\Delta y$, but now the ratio of the amplitudes has been increased by 3.3 by decreasing the disturbance amplitude. As may be expected, the system is now more apt at reaching the proper steady-state values. Trace 3 corresponds to increasing $\Delta y$ by 20 times its value in trace 1 while maintaining the same disturbance amplitude. Not only does the system adapt to the optimal values, but convergence is achieved much faster. While there is substantial differences in the rate of the weight adaptations, the changes in the signal to noise ratio do not dramatically affect the ventilation rate output or the final operating cost. This feature is due to the closed-loop feedback which keeps $P_{aCO_2}$ near its quiescent level.

As seen in Sect. 2.2.2, the adaptation gain, $k$, has a significant role in both the speed of convergence and robustness to noise. Figure 3-8 illustrates this effect for three different values of $k$ during exercise. As the adaptation gain is increased, there is a significant increase in the convergence rate. However, we also note an increase in fluctuations about the optimal values as $k$ increases. This trade-off also appears in the examples of Section 2.2.2.

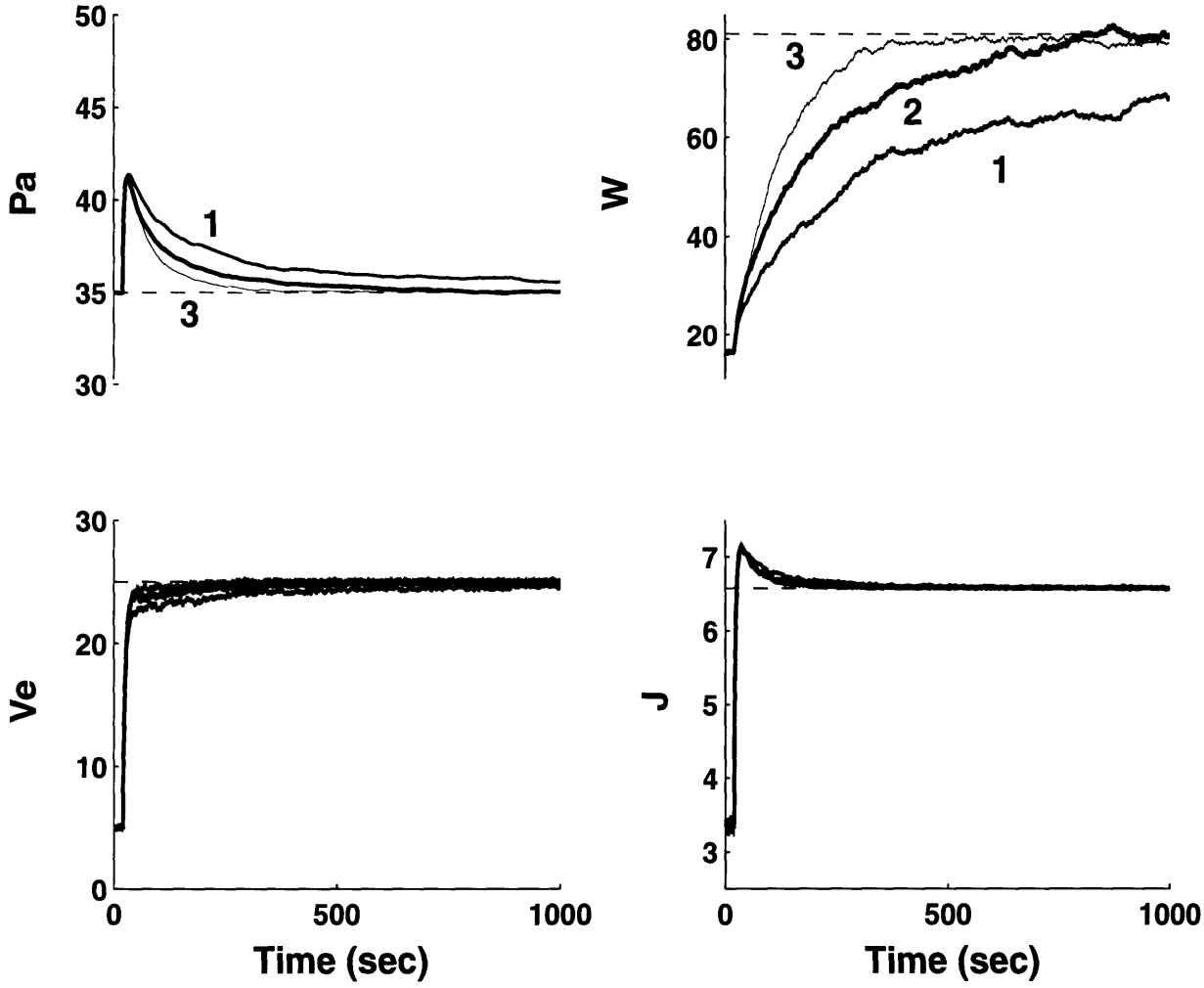


Figure 3-7: These simulations illustrate the role of the perturbation amplitude, $\Delta y$, in disturbance rejection. Trace 1 corresponds to a ratio of 150/1 (perturbation to disturbance amplitude), trace 2 maintains the same $\Delta y$ while decreasing the disturbance amplitude by 3.5, and trace 3 increases $\Delta y$ by 20 times that in trace 1.

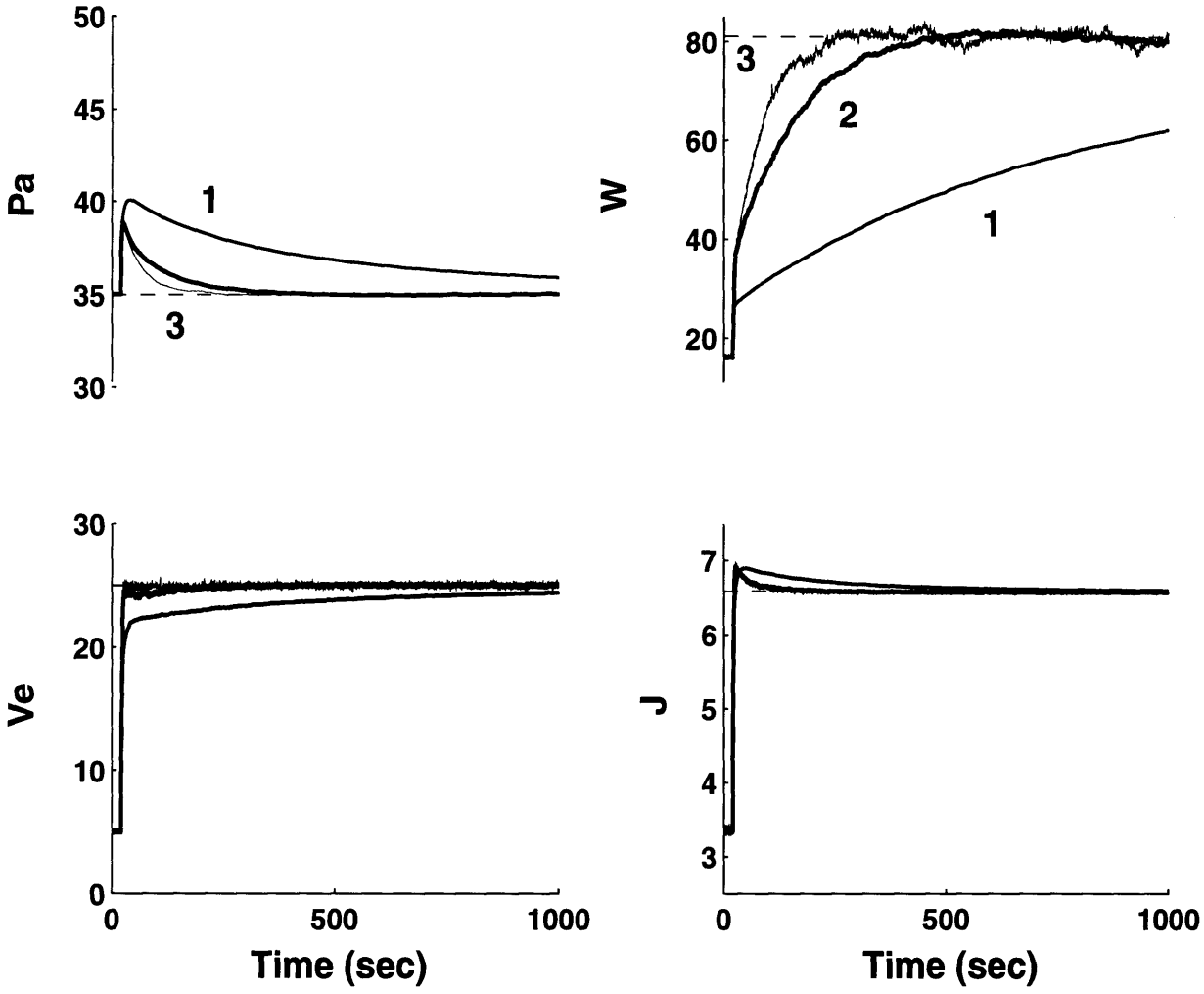


Figure 3-8: These simulations illustrate the role of the adaptation gain, $k$, in disturbance rejection. Trace 1, 2, and 3 correspond to increasing adaptation gains of $3*10^4$, $3*10^5$ and $3*10^6$. The perturbation to disturbance amplitude ratio was 1667/1 in all three traces.

# Chapter 4

# Discussion

## 4.1 Hebbian Covariance Learning and Optimization

The primary goal of adaptive optimal control is to continually extremize a scalar performance measure of a system despite model uncertainties. If the objective function is expressed as a long-term or asymptotic criterion, the transient trajectory that the system follows when approaching the optimum is not critical. In this respect, the adaptive optimal control structure being considered here is similar to tracking control problems, whose primary goal is to approach a given reference trajectory asymptotically.

The adaptive optimal control paradigm we have presented uses Hebbian covariance learning rules to optimize uncertain dynamic systems in an on-line manner. As outlined in Section 3 Hebbian learning has a long and expansive history in behavioral and neurobiological studies, yet its ability to optimize dynamic systems has never been fully explored. The conceptual framework behind our paradigm was first proposed by Poon (1993) to meet the physiological and neurobiological observations of respiratory control. The adaptive optimal control paradigm is designed so that a single Hebbian covariant synapse can tune the system feedback gain to continually extremize the near-term objective function. The Hebbian covariance law in our paradigm compares

the variance in post-synaptic activity with the resulting variance in pre-synaptic activity. This relationship differs from conventional covariance laws which relate changes in pre-synaptic input to the resulting changes in the post-synaptic output.

It has been hypothesized that the brain may employ various forms of adaptive control in certain physiological systems. For example, in motor control [27, 24, 50], it has been demonstrated that some form of adaptive control may underlie its behavior. Furthermore, the notion of optimal brain behavior has been suggested in the cognitive and behavioral domains where various computational models have emerged, such as "adaptive critic" [5], temporal difference [47] and Q-learning [53] forms of reinforcement learning. These learning models often incorporate Hebbian rules of adaptation and embody the results of dynamic programming [6], a widely used method for optimal path finding.

Our adaptive structure resembles a reinforcement learning system in that perturbations are invoked from the controller. In turn, the resulting signal from the environment serves to positively reinforce a "beneficial" change and likewise inhibit an "undesirable" change. Our paradigm, however, differs from other learning models in several manners. First, the structure of the system is comparatively simple, requiring only a single modifiable synapse. Second, some form of long-term memory is generally not required. This feature results because, unlike other adaptive methods, a complete model of the environment is typically unnecessary. Nevertheless, as demonstrated in the 2nd-order linear example (Sect. 2.2.2), some knowledge of the system dynamics may be required to form the adaptation law. Another major difference is that our objective is steady-state (asymptotic) optimal regulation whereas generally reinforcement learning is concerned with dynamic optimization in the non-steady state.

To implement the Hebbian covariance paradigm in dynamical systems, we have presented a general method to transform a static, long-term objective function into a dynamic, near-term form. In a biological context, the long-term objective may be viewed as a physiological goal whereas the near-term objective is its transformation into neural coordinates [34]. Thus time-dependent processes such as sluggishness,

closed-loop recirculation, or time delays, may be accounted for in the neural, near-term objective.

Through simulations, we have demonstrated that the Hebbian covariance paradigm is a robust optimizer of both linear and non-linear environments. The systems simulated were up to 2nd-order, although the general theory is applicable to higher dimensions. We also demonstrated the dilemmatic roles of both the adaptation gain and perturbation amplitude in determining the rate of convergence and the sensitivity to external disturbances. This behavior is reflected in a trade-off between the robustness to noise and the convergence rate. Similar roles for these parameters have been recognized in other adaptive systems where a balance must be sought between the transient errors and the steady-state oscillations about the goals.

Several caveats should be noted. First, some systems may be highly damped, thereby precluding the effect of high-frequency perturbations. To avoid aliasing in the adaptation, the frequency of perturbations must remain within the system bandwidth. As well, high-frequency perturbations, while being sufficiently rich to avoid spurious optima, may excite unmodelled dynamics in certain unknown environments. Although we did not demonstrate it explicitly, it should also be emphasized that the paradigm does not reject static disturbances in the classical sense. Rather, such disturbances are sensed as components in the environment thereby leading to new optimal solutions.

## 4.2 Optimal Respiratory Control

The Hebbian covariance learning model of respiratory control conforms well to existing neurobiological and physiological data. Recent evidence suggests that synaptic plasticity in the brainstem may participate in respiratory control [18, 51]. The observed forms of plasticity (e.g. LTD [55]) are consistent with the proposed Hebbian covariance learning rules. Evidence that NMDA mediated plasticity plays a significant role in respiratory control is mounting [35]. A recent *in-vivo* study reports that NMDA mediated synaptic plasticity may partake in habituation to vagal stimulation

[42], a process previously considered solely reflexogenic in nature.

The Hebbian covariance learning model not only agrees with the growing neurobiological data, it is also harmonious with the dilemmatic physiological responses to both exercise and $CO_2$ inhalation. Previous studies have shown that these responses are compatible with an optimal respiratory response [30, 31]. Poon (1996) first proposed a model of self-tuning optimal respiratory control by Hebbian covariance learning, where static and dynamic models were derived. The respiratory simulations here provide additional evidence that a Hebbian covariance learning system can robustly optimize the dynamic, nonlinear respiratory system.

The typical ventilatory response to exercise is recognized by a tri-phasic transient. Ideally our simulations should agree with the transient response as well as the steady-state optimal solution. The simulation results reveal that the transient response contains phases I-III after a step change in metabolic $CO_2$ but not after a ramp change. This discrepancy may suggest that an additional signal may be available to the brainstem controller which our model is lacking. While the Hebbian covariance learning paradigm is well suited for steady-state optimal solutions, it lags behind any short-term dynamic requirements. For instance, at the onset of exercise, one might imagine that the need for oxygen in the muscles is quite severe. Thus, the dynamic optimal solution would call for a more rapid increase in ventilation than would normally be provided. This type of short-term need could be incorporated in the objective function by adding an additional term expressing the sudden expectation of metabolic activity.

Several testable predictions may be made from the Hebbian covariance model of respiratory control. First, while the model does not entertain any parameter estimation, it does hinge on the explicit dynamic adaptation rule (Eq. 2.27.) The two time constants of the dynamic equations appear in this equation, namely, $V_L$ and $\tau$. Therefore manipulation of these time constants could lead to interesting behavior. The system may quickly adapt to the new parameters or may operate at entirely different levels.

The model also requires precise knowledge of the inherent time delays in the

system. One would expect that an increased circulatory time-delay would therefore befuddle the reinforcement learning system. Adaptations to match the new closed-loop structure would be needed to restore stability. This type of problem is related to Cheyne-Stokes breathing in which weak cardiac output increases the time delay between changes in pulmonary $CO_2$ tension and cerebral spinal fluid bathing the chemoreceptor neurons [48]. The symptoms which these patients experience, namely periodic increases and decreases in ventilation, may result from such abnormal time delays.

## 4.3 Future Research

The Hebbian covariance learning structure we presented adapts based on the immediate reinforcement signal from the environment. While this structure is suitable for the many problems, there may be advantages to learning over longer time scales. This approach may serve to improve robustness. Furthermore, it may be needed in higher-order systems where fast sampling times may overlook meaningful changes in the environment. Longer reinforcement windows may also be employed in respiratory system. It is conceivable that the system adapts using both spontaneous perturbations as well as changes from single or multiple respiratory cycles. One may hypothesize that different sensory time constants, such as in the peripheral and central chemoreceptors, could serve to provide different integration windows useful for adaptation.

Current engineering applications for this learning paradigm are under investigation. The simple adaptive structure of Hebbian covariance learning that uses spontaneous changes in the input and output may prove useful in systems whose conditions are uncertain and when optimal solutions are desired. The scalar nature of the objective function allows additional criterion, such as tracking errors, to be easily added to the objective function. Furthermore, we hope this paradigm can be extended to meet dynamical constraints such as pole placement or other frequency response characteristics.

In conclusion, we have shown that Hebbian covariance learning may act as an effective computational tool subserving adaptive optimal control. This computational paradigm is generally applicable to a wide class of optimal regulation problems. By using computer simulations, we have also shown that the Hebbian covariance controller is robust to both model uncertainties and noise disturbances. These theoretical results may set the stage for further experimental exploration into the novel optimization role for Hebbian covariance learning in neural systems.

# Appendix A

# Example of a Simulation Script

This section illustrates an example script used in the simulation of the 1st-order linear system (Sect. 2.2.2). The script was written for and run in MATLAB.

```
% TIME Variables
start_time=0;
sim_time=500;
step_size=.1;

% BASIC Variables
theta_0=70;
theta_1=2;
A=1;
B=-.5;
T=step_size;
t=2;
k_ic=.4;

adap_always=1;          % when =1, always adapting
                        % otherwise waiting for change in "var"
max_dw=.005;            % set maximum rate of change for W

using_sum_x=1;          % if =1, use x_sum values instead of x

% VARIATIONS STEP INPUTS to y1
step_amp=2;             % step amplitude perturbation
step_hold=.1;
step_seperation=step_hold/T;
```

```
steps=0;                    % if =1, steps, otherwise impulses
noise_yes=1;                % if =1, noise input

% INITIAL CONDITIONS
w_ic=-B/A;
y_ic=((B^2)*theta_1-B*theta_0)/((A^2)+(B^2));
x_ic=(1/A)*(theta_0+B*(y_ic-theta_1));

if noise_yes==1
        disp('Generating Random Input')
        noise_at_n=rand((sim_time/step_size)+1,1);
        noise_sign=rand((sim_time/step_size)+1,1);
end

% DISTURBANCES TO X
dist_yes=1;                 % if =1, add disturbances
dist_amp=.001;
dist=dist_amp.*(2.*rand((sim_time/step_size+2),1)-ones((sim_time/step_size+2),1));

if sim_time>-.1
        disp('simulating...')
end

time=zeros(sim_time/step_size,1);
x=zeros(sim_time/step_size,1);
y=zeros(sim_time/step_size,1);
w=zeros(sim_time/step_size,1);
J=zeros(sim_time/step_size,1);
var=zeros(sim_time/step_size,1);
dx=zeros(sim_time/step_size,1);
dy=zeros(sim_time/step_size,1);
dw=zeros(sim_time/step_size,1);
dx_sum=zeros(sim_time/step_size,1);
x_sum=zeros(sim_time/step_size,1);
k=zeros(sim_time/step_size,1);
avg_dw=zeros(sim_time/step_size,1);
extra_input=zeros(sim_time/step_size,1);

% STEP CHANGES
B_new=-.75;
A_new=A;
B_time=10;
A_time=10;

x(1,1)=x_ic;
```

```
x_sum(1,1)=x_ic;
y(1,1)=y_ic;
w(1,1)=w_ic;
J(1,1)=((x_ic^2)+(y_ic^2))/2;
k(1,1)=k_ic;
var(1,1)=0;
time(1,1)=start_time;
extra_input(1,1)=0;
sign=-1;
B_sign=1;
last_B=B_time;
n=0;

% START iterations
for current_time=time(1,1):step_size:sim_time+start_time,
        n=n+1;
% Advance TIME
        time(n+1,1)=time(n,1)+step_size;
% STEP inputs
        if time(n+1,1)>=B_time
                B=B_new;
        end
        if time(n+1,1)>=A_time
                A=A_new;
        end
% Controlling Variable INPUT Disturbance
        if noise_yes==0
        if n==1
                var(n+1,1)=var(n,1)+sign*step_amp;
                last=n;
                sign=sign*(-1);
        elseif n-last>=step_seperation
                if steps==1
                        var(n+1,1)=var(n,1)+(sign)*step_amp*2;
                else
                        var(n+1,1)=var(n,1)+(sign)*step_amp;
                end
                sign=sign*(-1);
                last=n;
        else
                if steps==1
                        var(n+1,1)=var(n,1);        % STEPS
                else
                        var(n+1,1)=0;               % IMPULSES
                end
```

```
        end
        else                            % NOISE
        if n==1
                var(n+1,1)=noise_at_n(n,1)*sign^(round(noise_sign(n,1)+1))*step_amp/2;
                last=n;
        elseif n-last>=step_seperation
                var(n+1,1)=noise_at_n(n,1)*sign^(round(noise_sign(n,1)+1))*step_amp;
                last=n;
        else
                if steps==1
                        var(n+1,1)=var(n,1);        % HOLDING VALUES (STEPS)
                else
                        var(n+1,1)=0;               % IMPULSES
                end
                end
        end
% The governing EQUATIONS
        if dist_yes==1
                x(n+1,1)=(T/t)*(theta_0+x(n,1)*((t/T)-A)+B*(y(n)-theta_1))+dist(n+1,1);
        else
                x(n+1,1)=(T/t)*(theta_0+x(n,1)*((t/T)-A)+B*(y(n)-theta_1));
        end;
        y(n+1,1)=x(n,1)*w(n,1)+var(n+1,1);
        J(n+1,1)=((x(n+1,1)^2)+(y(n+1,1)^2))/2;
% Calculate X_sum for recirculation problems
        if n==1
                x_sum(n+1)=x(n+1);
        else
                x_sum(n+1)=x(n+1)+(((T*A)/t)-1)*x(n);
        end;
% Calculate DELTAS
        if n<=0
                dx(n+1,1)=0;
                dy(n+1,1)=0;
        else
                dx(n+1,1)=x(n+1,1)-x(n,1);
                dy(n+1,1)=y(n+1,1)-y(n,1);
        end
% Calculate dx_sum
        if n==1
                dx_sum(n+1)=dx(n+1);
        else
                dx_sum(n+1)=dx(n+1)+(((T*A)/t)-1)*dx(n);
        end
% Calcultate Delt- W
```

```
        dw(n+1,1)=-k*((t/(T*A))^2*dx_sum(n+1,1)*dy(n,1)
        +(y(n,1)/x_sum(n,1))*(dy(n,1)*dy(n,1)));
% Set Max dw
        if abs(dw(n+1))>max_dw
                dw(n+1)=(abs(dw(n+1))/dw(n+1))*max_dw;
        end
% CHANGE W if warranted
        if n>2
                if adap_always==1
                        w(n+1,1)=dw(n+1,1)+w(n,1);
                else
                        if n>3
                                if abs(var(n-1,1)-var(n-2,1))>0 & abs(var(n-1,1)-var(n-3,1))>0
                                        w(n+1,1)=dw(n+1,1)+w(n,1);
                                else
                                        w(n+1,1)=w(n,1);
                                        dw(n+1,1)=0;
                                end
                        else
                                if abs(var(n-1,1)-var(n-2,1))>0
                                        w(n+1,1)=dw(n+1,1)+w(n,1);
                                else
                                        w(n+1,1)=w(n,1);
                                        dw(n+1,1)=0;
                                end
                        end
                end
        else
                w(n+1,1)=w(n,1);
                dw(n+1,1)=0;
        end
% SATURATION of W
        if -B/A<0
                if w(n+1,1)>0
                        w(n+1,1)=-0.01;
                end
        else
                if w(n+1,1)<0
                        w(n+1,1)=0.01;
                end
        end
        if w(n+1,1)>1
                w(n+1,1)=.98;
        end
        if w(n+1,1)<-1
```

```
                w(n+1,1)=-.98;
        end
end
```

# Bibliography


[1] E. Ahissar and M. Ahissar. Plasticity in auditory cortical circuitry. *Current Opinion in Neurobiology*, 4:580–587, 1994.

[2] K. J. Åström and T. Bohlin. Numerical identification of linear dynamic systems from normal operating records. In P. H. Hammond, editor, *Theory of Self-Adaptive Control Systems.* Plenum Press, New York, 1966.

[3] K. J. Åström and B. Wittenmark. *Adaptive Control.* Addison-Wesley Publishing Company, Inc., New York, 1995.

[4] C. A. Barnes, L. J. Bindman, Y. Dudai, Y. Frégnac, M. Ito, T. Knöpfel, S. G. Lisberger, R. G. M. Morris, M. Moulins, J. A. Movshon, W. Singer, and L. R. Squire. Group Report: Relating activity-dependent modifications of neuronal function to changes in neural systems and behavior. In A.I. Selverston and P. Ascher, editors, *Cellular and Molecular Mechanisms underlying Higher Neural Functions*, pages 81–110. Wiley, New York, 1994.

[5] A. G. Barto. Reinforcement learning and adaptive critic methods. In D. A. White and D. A. Sofge, editors, *Handbook of Intelligent Control: Neural, Fuzzy, and Adaptive Approaches*, pages 469–492. Van Nostrand Reinhold, New York, 1992.

[6] R. E. Bellman. *Dynamic programming.* Princeton University Press, Princeton, 1972.

[7] T. V. Bliss and G. L. Collingridge. A synaptic model of memory: long-term potentiation in the hippocampus. *Nature*, 361:31–39, 1993.

[8] T. V. P. Bliss and T. Lømo. Long-lasting potentiation of synaptic transmission in the dentate area of the anesthetized rabbit following stimulation of the perforant path. *J. Physiol.*, 232:331–335, 1973.

[9] T. H. Brown, P. F. Chapman, E. W. Kairiss, and C. L. Keenan. Long-term synaptic potentiation. *Science*, 242:724–728, 1988.

[10] T. H. Brown, E. W. Kairiss, and C. L. Keenan. Hebbian synapses: biological mechanisms and algorithms. *Ann. Rev. Neurosci.*, 13:475, 1990.

[11] H. Butler. *Model Reference Adaptive Control.* Prentice Hall, New York, 1992.

[12] J. H. Byrne. Cellular analysis of associative learning. *Physiol. Rev.*, 67:329–439, 1987.

[13] W. B. Cannon. *The Wisdom of the Body.* Norton, New York, 1932.

[14] Y. Dan and M. M. Poo. Hebbian depression of isolated neuromuscular synapses in vitro. *Science*, 256:1570–3, 1992.

[15] F. L. Eldridge and D. E. Millhorn. Oscillation, gating, and memory in the respiratory control system. In N. S. Cherniack and J. G. Widdicombe, editors, *Handbook of Physiology*, volume 2, section 3, pages 93–114. American Physiological Society, Bethesda, MD, 1986.

[16] S. J. England, J. E. Melton, P. Pace, and J. A. Neurbauer. NMDA receptors mediate respiratory short-term potentiation in the nucleus tractus solitarius. *FASEB J.*, 6:A1826, 1992.

[17] Y. Frégnac, D. Shulz, S. Thorpe, and E. Bienenstock. A cellular analogue of visual cortical plasticity. *Nature*, 333:367–370, 1988.

[18] R. F. Fregosi. Short-term potentiation of breathing in humans. *J. Apl. Physiol.*, 71:892–899, 1991.

[19] R. Gessel and M. A. Hamilton. Reflexogenic components of breathing. *Am. J. Physiol.*, 133:694–719, 1941.

[20] D. O. Hebb. *The Organization of Behavior.* Wiley, New York, 1949.

[21] J. C. Houk. Control strategies in physiological systems. *FASEB J.*, 2:97–107, 1988.

[22] M. Ito. Synaptic plasticity in the cerebellar cortex and its role in motor learning. *Can. J. Neurol. Sci.*, 20:S70–4, 1993.

[23] E. R. Kandel. *A cell biology approach to learning.* Grass Lecture Monograph 1, Bethesda, MD, 1978.

[24] M. Kawato and H. Gomi. A computational model of four regions of the cerebellum based on feedback-error learning. *Biol. Cybern.*, 68:95–103, 1992.

[25] S. R. Kelso, A. H. Ganong, and T. H. Brown. Hebbian synapses in hippocampus. *Proc. Natl. Acad. Sci. USA*, 83:5326–5330, 1986.

[26] R. Linsker. From basic network principles to neural architecture: emergence of orientation columns. *Proc. Natl. Acad. Sci.*, 83:9779–8783, 1986.

[27] S. G. Lisberger. The neural basis of learning of simple motor skills. *Science*, 242:728–735, 1988.

[28] P. R. Montague, P. Dayan, Person C., and T. J. Sejnowski. Bee foraging in uncertain environments using predictive Hebbian learning. *Nature*, 377:725–8, 1995.

[29] G. A. Parker and J. Maynard Smith. Optimality theory in evolutionary biology. *Nature*, 348:27–32, 1990.

[30] C.-S. Poon. Optimal control of ventilation in hypoxia, hypercapnia and exercise. In B. J. Whipp and D. W. Wiberg, editors, *Modelling and Control of Breathing*, pages 189–196. Elsevier, New York, 1983.

[31] C.-S. Poon. Ventilatory control in hypercapnia and exercise: optimization hypothesis. *J. Appl. Physiol.*, 62:2447–2459, 1987.

[32] C.-S. Poon. Adaptive neural network that subserves optimal homeostatic control of breathing. *Annals of Biomed. Engr.*, 21:501–508, 1993.

[33] C.-S. Poon. Respiratory models and control. In J.D. Bronzion, editor, *Biomedical Engineering Handbook*, page 2404. CRC Press, Boca Raton, Fl., 1995.

[34] C.-S. Poon. Self-tuning optimal regulation of respiratory motor output by Hebbian covariance learning. *Neural Networks*, 8:1–17, 1996.

[35] C.-S. Poon, Y. Li, S. X. Li, and S. Tonegawa. Respiratory rhythm is altered in neonatal mice with malfunctional NMDA receptors. *FASEB J.*, 8:A389, 1994.

[36] C.-S. Poon, S. L. Lin, and O. B. Knudson. Optimization character of inspiratory neural drive. *J. Appl. Physiol.*, 73:2005–2017, 1992.

[37] I. P. Priban and W. F. Fincham. Self-adaptive control and the respiratory system. *Nature (London)*, 208:339–343, 1965.

[38] T. J. Sejnowski. Statistical constraints on synaptic plasticity. *J. Theor. Biol.*, 69:385–389, 1977.

[39] T. J. Sejnowski. Storing covariance with nonlinearly interacting neurons. *J. Math. Biol.*, 4:303–321, 1977.

[40] R. Shadmehr and F. A. Mussa-Ivaldi. Adaptive representation of dynamics during learning of a motor task. *J. Neurosci.*, 14:3208–3224, 1994.

[41] C. S. Sherrington. *The Integrative Action of the Nervous System.* C. Scribner's Sons, New York, 1906.

[42] M. S. Siniaia, D. L. Young, and C.-S. Poon. Habituation and rebound of rhythmic respiratory motor responses to vagal stimulation. *Society for Neuroscience Abstract*, (in press), 1997.

[43] J.-J. E. Slotine. *Applied nonlinear control.* Prentice Hall, Englewood Cliffs, N. J., 1991.

[44] R. E. Soodak. Reverse-Hebb plasticity leads to optimization and association in a simulated visual cortex. *Vis. Neurosci.*, 6:507–518, 1991.

[45] P. K. Stanton. LTD, LTP, and sliding threshold for long-term synaptic plasticity. *Hippocampus*, 6:35–42, 1996.

[46] P. K. Stanton and T. J. Sejnowski. Associative long-term depression in the hippocampus induced by Hebbian covariance. *Nature*, 339:215–218, 1989.

[47] R. S. Sutton. *Temporal credit assignment in reinforcement learning.* Ph.D. diss., Dept. Computer and Info. Sci., Univ. Massachusetts, Amherst, 1984.

[48] A. E. Taylor, K. Rehder, R. E. Hyatt, and J. C. Parker. *Clinical Respiratory Physiology.* W. B. Saunders Company, Philidelphia, 1989.

[49] S. M. Tenney and L. C. Ou. Ventilatory response of decorticate and decerebrate cats to hypoxia and $CO_2$. *Respir. Physiol.*, 29:81–92, 1977.

[50] M. Tryfonidis and C.-S. Poon. Model reference adaptive control modeling of motor learning. *Society for Neuroscience Abstract*, (in press), 1997.

[51] P. G. Wagner and F. L. Eldridge. Development of short-term potentiation of respiration. *Respiratory Physiol.*, 83:129–140, 1991.

[52] K. Wasserman, B. J. Whipp, and R. Casaburi. Respiratory control during exercise. In N. S. Cherniack and J. G. Widdicombe, editors, *Handbook of Physiology*, pages 595–620. American Physiological Society, Bethesda, MD, 1986.

[53] C. J. C. H. Watkins. *Learning with delayed rewards.* Ph.D. thesis, Psychology Dept., Cambridge University, 1989.

[54] A. L. Yuille, D. M. Kammen, and D. S. Cohen. Quadrature and the development of orientation selective cortical cells by Hebb rules. *Biol. Cybern.*, 61:183–194, 1989.

[55] Z. Zhou, J. Champagnat, and C.-S. Poon. Phasic longterm depression in brainstem NTS neurons: differing roles of AMPA receptor desensitization. *J. Neuroscience*, (in press), 1997.