REVIEW

# An overview of adaptive model theory: solving the problems of redundancy, resources, and nonlinear interactions in human movement control

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF NEURAL ENGINEERING

J. Neural Eng. **2** (2005) S279–S312 doi:10.1088/1741-2560/2/3/S10

**REVIEW**

# An overview of adaptive model theory: solving the problems of redundancy, resources, and nonlinear interactions in human movement control

**Peter D Neilson and Megan D Neilson**

Neuroengineering Laboratory, School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, Australia

E-mail: p.neilson@unsw.edu.au

Received 2 February 2005
Accepted for publication 22 July 2005
Published 31 August 2005
Online at stacks.iop.org/JNE/2/S279

**Abstract**

Adaptive model theory (AMT) is a computational theory that addresses the difficult control problem posed by the musculoskeletal system in interaction with the environment. It proposes that the nervous system creates motor maps and task-dependent synergies to solve the problems of redundancy and limited central resources. These lead to the adaptive formation of task-dependent feedback/feedforward controllers able to generate stable, noninteractive control and render nonlinear interactions unobservable in sensory–motor relationships. AMT offers a unified account of how the nervous system might achieve these solutions by forming internal models. This is presented as the design of a simulator consisting of neural adaptive filters based on cerebellar circuitry. It incorporates a new network module that adaptively models (in real time) nonlinear relationships between inputs with changing and uncertain spectral and amplitude probability density functions as is the case for sensory and motor signals.

## 1. Introduction

We have been exploring the role of internal models in human movement control and its disorders since the 1980s. At that time many engineers were familiar with the concept of an adaptive internal model as the basis of a feedforward/feedback controller of a dynamic system (e.g., Goodwin and Sin (1984)). In contrast, argument among motor control theorists still persisted from earlier decades about 'open loop' motor programming versus 'closed loop' feedback control of movement. Some, disenchanted with a computational approach, moved towards ecological psychology that as a basic tenet eschews the notion of models or representations. And some in turn became involved with a theory of nonlinear dynamics that had developed from the study of oscillatory physical systems. This fuelled what became a potential paradigm crisis in the 1990s (Abernethy and Sparrow 1992). While this dissension has not entirely resolved, the computational approach has moved far from the old days of unrealistic motor programs. Internal models are alive and well and an account of skilled human performance in terms of the control of a multivariable, nonlinear, time-varying, dynamic system is no longer unrealistic.

In search of such an account we developed adaptive model theory (AMT), a computational account of movement control in the spirit of Marr and Poggio's paper 'From understanding computation to understanding neural circuitry' (1977). The theory stemmed from two recognitions that were relatively unexplored in movement control at that time. First was the idea that sensory feedback could potentially establish an adaptive inverse model as the basis of a feedforward controller of movement (Neilson 1982, Neilson and Neilson 1985).

Second was the realization that the circuitry of the cerebellum provided an adaptive filter network capable of carrying out the computations necessary to produce adaptive models from sensory feedback (Neilson *et al* 1985, 1992). Because of our involvement with movement disorders, in particular with cerebral palsy and stuttering, much of our work found its way to that literature (e.g., Neilson and Neilson (1987), Neilson *et al* (1992, 1997), Neilson and O'Dwyer (1984)) rather than to the publication stream of computational control and neural network modeling of movement. Nevertheless, over 25 years AMT has evolved as a model of central nervous system (CNS) processing, guided by behavioral and neurophysiological research and informed by the mathematics of control. In parallel, neural engineering has emerged in its own right and AMT is now one of various computational theories that seek to elucidate the mechanisms of human movement control (see Doya *et al* (2001), Flash and Sejnowski (2001), Jordan and Wolpert (1999), Kawato (1999), Mussa-Ivaldi (1999), Wolpert *et al* (2003)).

For this special issue on internal models we bring together in a single framework the essence of how AMT solves three fundamental problems of movement control: the problems of redundancy, resources and nonlinear interactions. In this context we present the design and performance of a newly developed adaptive filter network for nonlinear dynamical modeling. From earliest days AMT has always been couched in terms of multivariable, nonlinear, adaptive control but until recently our simulations were confined to linear neural adaptive filter processes applied to signals with tractable amplitude and spectral density probability functions. We chose to develop our ideas step by step, starting with adaptive control of a linear single input/single output linear dynamical system, then progressing to control of a multiple input/multiple output system and finally to control of a multiple input/multiple output nonlinear system. Each step involved many issues that had to be resolved by computer simulations and experimental testing. The most recent difficulty involved identification of a biologically realistic process for identifying nonlinear models of relationships between biological signals with time-varying, non-Gaussian statistical characteristics. As we report here, we have recently overcome this difficulty. This allows us to give a blueprint for a comprehensive simulation of human movement planning and control based on the formation of nonlinear internal models by means likely to exist in the CNS.

In the following section we give an account of the dynamical systems and signals that define the musculoskeletal system in interaction with the environment. We then introduce the three problems of redundancy, resources and nonlinear interactions associated with the control of this system. These are set out in section 2.4, then each is presented along with its proposed solution in sections 3, 4 and 5 respectively. In section 6 we bring the solutions together and incorporate the processing structure established for each into the design of an overall structure for planning and control of movement. Throughout this structure we require adaptive internal models of multivariable nonlinear dynamical relationships. Thus in section 7 we review the basic adaptive filter module of AMT and present the new adaptive filter network that can 'plug-in' to accomplish this. We then test its performance (section 7.4) as a validation of the viability of structures developed in the earlier sections. In section 8 we discuss how the proposal relates to experimental evidence and to other work.

There are two consistencies in what follows that we deal with in an appendix. First, throughout the paper we deal with discrete-time signals and discrete-time processing rather than with continuous-time systems. We set out the reasons for this in appendix A.1. Second, throughout the paper we use a consistent set of symbols for the many signals and systems involved. We define each symbol in the text and/or legend where it first occurs (and sometimes subsequently) but to assist the reader we include a comprehensive reference table in appendix A.2. This lists all symbols and other abbreviations in alphabetical order.

## 2. The plant

In AMT control of human movement is seen in terms of a feedforward/feedback controller of a 'plant' that consists of the entire musculoskeletal system in interaction with its environment. There are several hundred (say 700) functional muscles controlling about 110 elemental movements. By definition a functional muscle is a set of muscles fibers that always change length in a perfectly correlated way. All fibers in a functional muscle have the same biomechanical action on elemental movements. We define elemental movements to be the simplest movements that can be made independently one at a time. Some consist of rotations about a single axis at a single joint (e.g., elbow flexion) but many involve coupled rotations of multiple joints and/or movements (e.g., flexing the spine or raising an eyebrow). Coordinated combinations of elemental movements provide the skilled synergetic actions that interact with the external world to comprise the vast human movement repertoire. This is a multivariable, nonlinear, redundant, dynamically interactive system and the control task is a challenging one.

The plant can be seen as three subsystems connected in cascade (figure 1): the muscle control system **MCS**, the biomechanical system **BM** and the external system **E**. Each is discussed further below. In brief, the **MCS** transforms a 700-dimensional vector of efferent drives $\boldsymbol{\alpha}_{\mathbf{m}}$ into a 700-dimensional vector of tensions $\mathbf{t}$. The **BM** system transforms the tensions generated by 700 functional muscles into a 220-dimensional vector $\mathbf{q}$ comprised of 110 elemental movements $\boldsymbol{\theta}$ and 110 co-contractions $\mathbf{c}$ about those movements. These 220 elemental movements and co-contractions operate on the external world **E** and produce a $p$-dimensional vector of responses $\mathbf{r}$, with $p$ determined by the particular system **E**.

Responses $\mathbf{r}$ are detected by exteroceptors and/or interoceptors (e.g., vision, pressure). Tensions $\mathbf{t}$ and movements and co-contractions $\mathbf{q}$ are detected by interoceptors distributed in muscles, tendons, joints and skin. Together these provide kinesthetic sensation that also includes a 'sense of effort' derived from feedback of outgoing motor commands, the so-called efference copy (McCloskey 1981, Gandevia 1987). From a central view it is the relationships between these

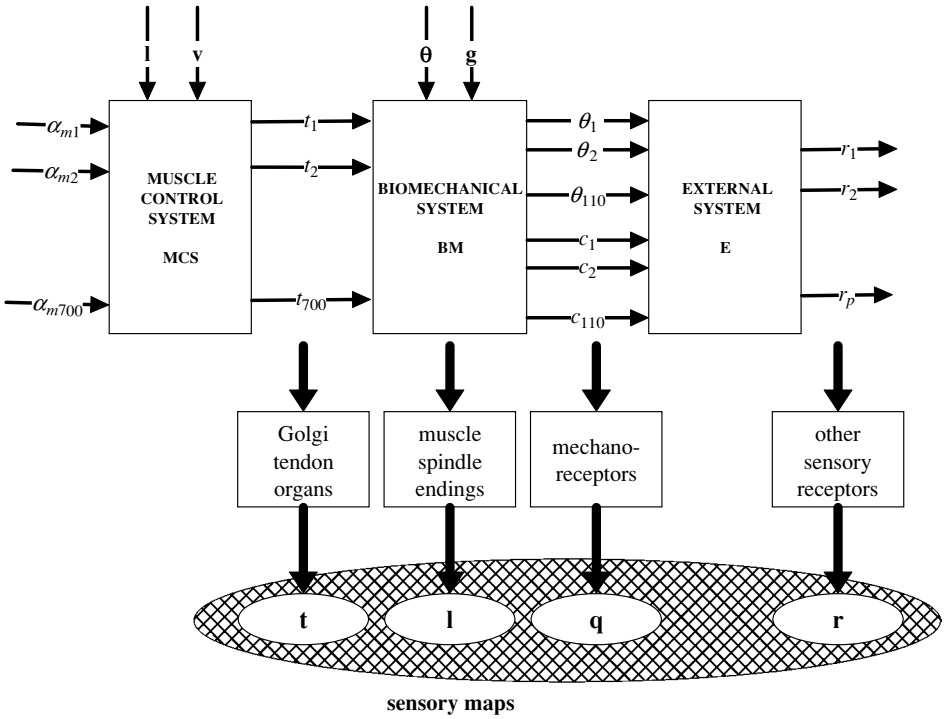



**Figure 1.** A schematic diagram illustrating how the plant can be identified within the central nervous system as three subsystems connected in cascade; muscle control system **MCS**, biomechanical system **BM** and external system **E**. 700-dimensional vector $\boldsymbol{\alpha}_{\mathbf{m}} = [\alpha_{m1}, \ldots, \alpha_{m700}]^{\mathrm{T}}$ of descending motor commands at the input to **MCS** is detected, either directly or indirectly, via efference copy. Output from **MCS** is a 700-dimensional vector $\mathbf{t} = [t_1, \ldots, t_{700}]^{\mathrm{T}}$ of tensions generated by 700 functional muscles detected by Golgi tendon organs and represented in a cortical sensory map. The 700 tensions are inputs to **BM**. Output from **BM** is a 220-dimensional vector $\mathbf{q} = [q_1, \ldots, q_{220}]^{\mathrm{T}}$ comprised of 110 elemental movements $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_{110}]^{\mathrm{T}}$ and 110 co-contractions $\mathbf{c} = [c_1, \ldots, c_{110}]^{\mathrm{T}}$. Elemental movements and co-contractions **q** are detected by mechanoreceptors in joints and skin and are represented in cortical sensory maps. Vector $\mathbf{l} = [l_1, \ldots, l_{700}]^{\mathrm{T}}$ of the lengths of 700 functional muscles is detected by muscle spindle endings and represented in a cortical sensory map. 220 elemental movements and co-contractions **q** are inputs to **E** that responds with a vector of reafferent response outputs $\mathbf{r} = [r_1, \ldots, r_p]^{\mathrm{T}}$. The reafferent responses are detected by a variety of sensory receptors and represented in cortical sensory maps. Thus the nervous system has a central representation of the inputs and outputs from each subsystem. The inputs **l** and **v** to **MCS** represent 700-dimensional vectors of muscle lengths and rates of change of muscle length respectively that interact in a nonlinear fashion with the $\boldsymbol{\alpha}_{\mathbf{m}}$ to **t** relationship. The inputs $\boldsymbol{\theta}$ and **g** to **BM** represent body posture and gravitational influences that interact in a nonlinear fashion with the **t** to **q** relationships.

sensory signals that define the subsystems **MCS**, **BM** and **E**. But before examining each of these separately we comment briefly on reflexes.

### *2.1. Reflex system*

It is well known that the **MCS** and **BM** systems are embedded within multiple reflex loops and that control of these reflex systems is mediated by descending modulation of the gains of interneurons in reflex pathways as well as by both alpha and gamma motor neurons (Neilson and Lance 1978, Neilson 1993a). Reflex systems are important in movement control. They provide feedback around **MCS** and **BM** systems that can modify closed-loop dynamics as well as oppose internally and externally generated perturbations and uncertainties. They play a role in stabilizing unstable biomechanical configurations and they are involved in tuning the 'shock absorber' (stiffness and viscosity) characteristics of elemental movements. Reflex systems are an integral part of AMT (see Neilson *et al* (1997)). They do not feature in this paper because the focus here is on feedforward control and on the external perceptual/motor feedback loop that predominates the multitude of internal minor loops.

Ignoring reflexes in this context does not create a logical problem because either (i) the feedback loops are contained within the **MCS** or (ii) **MCS** and **BM** systems are embedded in the feedback loops. As indicated above, **MCS**, **BM** and **E** are well defined by input/output relationships between sensory signals and thus can be identified centrally regardless of whether they correspond to closed-loop or open-loop relationships. Certainly monosynaptic and polysynaptic stretch reflex circuits influence relationships between descending drives to $\alpha$ motor neurons and muscle tensions. But even this does not disrupt the arguments to come about feedforward control. In fact, reflexes themselves raise questions about the nature of feedforward control. We assume in AMT that stretch reflexes function according to the well-accepted *servo-assist* model (Matthews 1972) where $\alpha$ and $\gamma$ drives controlling each muscle are functionally linked so that stretch reflexes are only brought into play when actual muscle length trajectories deviate from intended trajectories. To achieve this $\alpha$–$\gamma$ linkage the drive to $\alpha$ motor neurons must be such that the resulting muscle tensions operate on the **BM**

system to produce movements and muscle length changes that exactly unload the $\gamma$-generated mechanical deformations of muscle spindle endings. How an appropriate $\alpha$–$\gamma$ linkage is generated centrally is exactly the issue of feedforward control addressed in this paper.

### 2.2. Muscle control system *MCS*

The multivariable nonlinear dynamical relationships between the vector of outgoing motor commands to the $\alpha$ motor neurons $\boldsymbol{\alpha}_{\mathbf{m}} = [\alpha_{m1}, \ldots, \alpha_{m700}]^{\mathrm{T}}$ and the vector of tension signals $\mathbf{t} = [t_1, \ldots, t_{700}]^{\mathrm{T}}$ define the muscle control system (figure 1). We designate the **MCS** as a system because, in addition to some 700 functional muscles, it also includes multiple cross-coupled Golgi tendon organ tension feedback loops and Renshaw cell feedback loops around $\alpha$ motor neurons. Nevertheless, despite the presence of these internal loops **MCS** can be defined and identified within the CNS by the relationship between its inputs $\boldsymbol{\alpha}_{\mathbf{m}}$ and its outputs, the tensions **t** detected by Golgi tendon organs at musculotendonous junctions throughout the body. Because muscles have nonlinear tension–length–velocity characteristics it follows that the input/output relationships defining **MCS** include nonlinear interactions with muscle lengths $\mathbf{l} = [l_1, \ldots, l_{700}]^{\mathrm{T}}$ and muscle velocities $\mathbf{v} = [v_1, \ldots, v_{700}]^{\mathrm{T}}$. These interactions are shown in figure 1 by the inputs to **MCS** labeled **l** and **v**.

### 2.3. Biomechanical system *BM*

The vector of tensions **t** at the output of **MCS** pulls on bones and/or tissues and exerts turning and/or translation forces about the 110 elemental movements $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_{110}]^{\mathrm{T}}$ shown at the output of the biomechanical system **BM** (figure 1). The net turning or translation force about each elemental movement is determined by the difference in tensions between a set of agonist muscles and a set of antagonist muscles. This arrangement of agonist and antagonist muscles implies that muscles can co-contract about each elemental movement. While co-contraction does not change the net turning or displacement force about the joint, it is important in controlling the mechanical stiffness and viscosity of elemental movements to external force perturbations. Thus we also show 110 co-contractions $\mathbf{c} = [c_1, \ldots, c_{110}]^{\mathrm{T}}$ at the output of **BM**. Rotations of joints and translations of body parts (i.e., elemental movements $\boldsymbol{\theta}$) are detected by mechanoreceptors in the joint capsules, connective tissues and skin. Since co-contraction of muscles changes the stress patterns within connective tissues and joint capsules, the joint receptors are also well situated to detect co-contractions **c** about each elemental movement. Of course movements are the result of changes in length of functional muscles. Muscle lengths **l** are detected by muscle spindle endings and the relationship between $\boldsymbol{\theta}$ and **l** is almost certainly a static nonlinear (e.g., polynomial) one.

The biomechanical system **BM** is defined by the nonlinear dynamical relationships between the vector of tensions **t** at its input and the vector of elemental movements and co-contractions **q** at its output. All factors that influence the relationships between muscle tensions and body movements are included within **BM**. If one puts on heavy boots, hefts a weight or turns a spring-loaded steering wheel, the biomechanical system is changed. In other words, **BM** incorporates all the complex changing and uncertain mechanical interactions between the body and its environment. The multivariable, nonlinear, dynamical relationships between muscle tensions and body movements are highly interactive and are influenced by inertial, viscous, centrifugal, coriolis and gravitational forces. Clearly the relationship between the change of tension in a muscle and the resulting movement of the body depends on the posture of the body and on its orientation in the gravitational field. To indicate these nonlinear interactions within the biomechanical system figure 1 shows vectors $\boldsymbol{\theta}$ and **g** as inputs to **BM**, $\boldsymbol{\theta}$ being the vector of all elemental movements (that equivalently specifies body posture) and **g** being a vector of measures specifying orientation in the gravitational field.

### 2.4. External system *E*

Elemental movements[1] and co-contractions **q** at the output of **BM** operate on the external system **E** to produce responses $\mathbf{r} = [r_1, \ldots, r_p]^{\mathrm{T}}$ (figure 1). These are detected by sensory receptors (visual, auditory, labryrinthine, olfactory, thermal, etc) according to the nature of **E**. Controlling an external system is part of everyday experience. Consider steering an automobile. The elemental arm movements that turn the steering wheel are the inputs to a system **E** and are detected kinesthetically by joint and cutaneous receptors. When the car turns, the visual image translates across the retinae and is processed by the visual system. These and other sensory signals, such as those from inertial and pressure changes, are the outputs from **E**. The dynamical relationships between these sensory inputs and outputs define the external system **E**, in this case the steering characteristics of the automobile.

Speech production provides a less obvious example. Here muscular contractions control the stiffness and viscosity of the articulators and produce displacements detected by joint and cutaneous receptors. Similar detections occur for the accompanying respiratory and laryngeal activity. All these provide knowledge of the inputs to an external system **E**. In this case **E** is essentially a musical instrument comprised of a thoracic (respiratory) component, a laryngeal (phonatory) component and a vocal tract (articulatory) component as well as a component determined by the internal and external acoustic environment. The output of **E** is an acoustic speech signal detected by the auditory system via bone/tissue and air conduction. The nonlinear dynamical relationships between (a) respiratory, phonatory and articulatory movements and forces and (b) the resulting auditory feature signals (e.g., intensity, pitch, formant frequencies, damping of formants, etc) define an external system **E**.

### 2.5. Three control problems

There are three problems associated with controlling the plant of figure 1 that must be addressed by any theory of movement

[1] In an isometric task an elemental movement is replaced by the equivalent net turning or displacement force about the elemental movement.

control. First there is the well-known *problem of redundancy* in the musculoskeletal system where 700 functional muscles control 220 elemental movements and co-contractions. In other words, the musculoskeletal system is a non-square system with more inputs than outputs. This means that an infinite number of different patterns of muscle activation can produce the same movement and pattern of co-contractions. This long-known difficulty in human movement control is addressed in section 3. Second is a problem less frequently discussed in human movement literature, the *problem of limited central resources*. By definition each of the 220 elemental movements and co-contractions can be controlled independently one at a time. But simultaneous independent control of these components is another matter, relating directly to the availability of central processing resources. This issue is explored in section 4. Third is the *problem of nonlinear dynamical interactions*. The control task for a plant with interactive dynamics is to render these unobservable in the output measures, allowing the outputs to accurately track the inputs presented to the controller. In human movement, this is precisely the requirement for effective skilled performance where outputs must emulate intended multivariable movement trajectories. The 'case of the disappearing dynamics' in movement control is presented in section 5.

## 3. Solving the problem of redundancy

### *3.1. Wired-in synergy generator*

Bernstein (1967) has long been recognized for his insight that the problem of redundancy in the musculoskeletal system can be overcome by the formation of synergies. A synergy occurs when a number of functional muscles are controlled as an entity. One way this could be achieved is by means of a fixed 'wired-in' synergy generator in the descending pathways between the cortex and the muscles. Descending pathways form during the early stages of motor development and thereby create motor maps in the cortex. Much has been learned experimentally about these motor maps. For example, it is known that any one muscle can be activated by stimulating differing sites in the cortex, and conversely, stimulation of any one cortical site usually activates multiple muscles (Krakauer and Ghez 2000). Thus descending pathway connectivity can be thought of as a matrix, resembling a neural network rather than a private-line system for each muscle. Also, it is known from transcranial magnetic stimulation studies in awake humans that the gains of the descending pathways are modulated by central influences during movement and even thinking about making a movement is sufficient to modulate the gains of certain descending pathways (Stinear and Byblow 2003, 2004, Sohn *et al* 2003). Thus it is not unreasonable to postulate that descending pathways serve as a wired-in synergy generator with modulable gains.

Figure 2 depicts such a wired-in synergy generator (**WSG**) consisting of a matrix of connectivity between 220 central commands $\mathbf{m}_\alpha = [m_{\alpha 1}, \ldots, m_{\alpha 220}]^{\mathrm{T}}$ and the 700 $\alpha$ motor neuron commands $\boldsymbol{\alpha}_\mathbf{m}$ that serve as input to the plant in figure 1. The addition of the **WSG** allows an input of 220 central commands $\mathbf{m}_\alpha$ to generate an output of 220 elemental movements and co-contractions $\mathbf{q}$ from the **BM** subsystem. The musculoskeletal system, consisting of **WSG**, **MCS** and **BM** in cascade, is now square and the redundancy problem appears solved provided a 700 $\times$ 220 **WSG** is established. Or is it?

### *3.2. Minimizing demand for metabolic energy*

When we ask how might the CNS adaptively form wired-in muscle synergies, the problem of redundancy reappears. There are an infinite number of ways of establishing a 700 $\times$ 220 matrix of connectivity between the central commands $\mathbf{m}_\alpha$ and the peripheral commands $\boldsymbol{\alpha}_\mathbf{m}$. In other words, there are an infinite number of possible synergies that can compensate for redundancy in the musculoskeletal system. Which synergy does the CNS use and why? Selecting a particular connectivity is a problem of optimization, there is no argument about that. Optimization is the only mathematical tool available for finding a unique solution to a system of redundant equations. The interesting thing about optimization theory is that whatever solution we might choose from the infinite number possible, some unknown cost function will be optimized. The theory does not tell us what that cost function is. The only way to apply optimization theory is to specify a cost function, determine the optimal solution and then test experimentally whether the solution applies. Hence the debate begins when it comes to deciding what is optimized.

Here we explore the proposal that motor maps form automatically to minimize the demand by the muscles for metabolic energy, thereby creating a minimum-energy **WSG**. This determines the optimal activation of agonist and antagonist muscles required to generate pre-specified turning forces and co-contractions about elemental movements. It does not determine what those forces and co-contractions are. It has no influence on performance criteria such as for example, maximum velocity, maximum joint stiffness, maximum force or maximum ability to perform a series of tasks. Nor does it determine kinematic patterns of movement. The **WSG** simply assures that the multiple agonist (and antagonist) muscles at each elemental movement are activated in the appropriate proportions to generate the specified turning forces and co-contractions with minimum demand for muscular energy, whatever those specifications may be.

Existence of a minimum-energy **WSG** does not cap the energy expended but ensures that it is used to best advantage. A leopard, for example, expends a large amount of energy to capture prey. But because of its heavy, powerful construction, it is unable to sustain long periods of high-speed running. If its muscles are activated efficiently so as to minimize demand for energy during a chase, the leopard can run for a longer time and so is more likely to catch its prey. Any organism that can capture prey, escape predators and achieve goals with a minimum demand on metabolic energy, has an evolutionary advantage. The proposal that motor maps function to provide a minimum-energy **WSG** therefore makes good sense.

Using optimization methods (Lagrange multiplier and right pseudo-inverse), we have shown (Neilson and Neilson

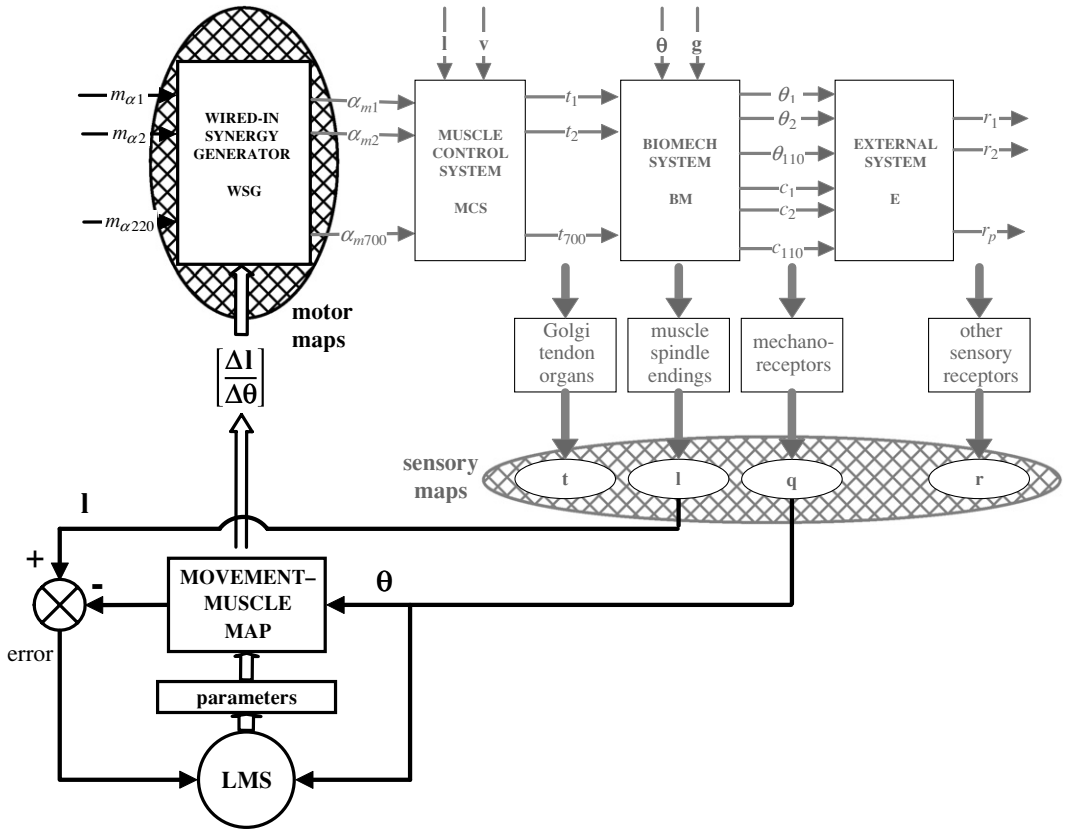



**Figure 2.** A schematic diagram illustrating adaptive formation of motor maps: the gray part of the diagram is derived from figure 1 and is not described again here. The block labeled WIRED-IN SYNERGY GENERATOR **WSG** corresponds to motor maps that connect 220 central motor commands $\mathbf{m}_\alpha = [m_{\alpha 1}, \ldots, m_{\alpha 220}]^{\mathrm{T}}$ to 700 alpha drives $\boldsymbol{\alpha}_\mathbf{m} = [\alpha_{m1}, \ldots, \alpha_{m700}]^{\mathrm{T}}$ at the input to the muscle control system **MCS**. The **WSG** represents connectivity in descending motor pathways. The matrix of signals $\boldsymbol{\Delta l}/\boldsymbol{\Delta\theta}$ modulates the gains of the descending pathways within the **WSG**. The matrix of signals $\boldsymbol{\Delta l}/\boldsymbol{\Delta\theta}$ is derived from an adaptive MOVEMENT-MUSCLE MAP shown at bottom left. This adaptively models the nonlinear (polynomial) relationships between the flexion and extension components of elemental movements $\boldsymbol{\theta}$ and the lengths of functional muscles **l** derived from cortical sensory maps. Only elemental movement components $\boldsymbol{\theta}$ are extracted from the sensory map labeled $\mathbf{q} = [\theta_1, \ldots, \theta_{110}, c_1, \ldots, c_{110}]^{\mathrm{T}}$. The circle labeled LMS represents least mean square adaptive filter circuitry described in section 7. The LMS adaptively tunes the parameters of the MOVEMENT-MUSCLE MAP to reduce the error signal at the output of the comparator to a minimum.

1999b, 2005) that to generate a required pattern of turning forces and co-contractions about elemental movements with minimum demand for metabolic energy, *each muscle must be activated in proportion to its moment arm about that elemental movement*. Any one elemental movement is usually operated on by more than one functional muscle and any one functional muscle usually operates across more than one elemental movement. Consequently, any one muscle participates in multiple elemental-movement synergies. In general, movement involves the overlapping of many elemental movement synergies. Minimal energy demand is achieved if each muscle responds to the sum of the optimal activation patterns for each of the elemental movement synergies it participates in.

### *3.3. Adaptive formation of a minimum-energy wired-in synergy generator*

The initial growth in the fetus of anatomical connections between cortex and peripheral motor neurons is no doubt under genetic control. However, the very first movements of the fetus suggest that these early connections are diffuse and poorly focused. The question becomes, how does the CNS automatically tune the gains of the descending pathways to activate muscles about each elemental movement in proportion to their moment arms about that movement? We propose a two-stage process. The first depends on formation of sensory maps by correlational-based mechanisms of synaptic plasticity (e.g., Hebbian learning) based on the patterns of sensory input received. These processes are well studied experimentally and occur slowly; e.g., over 12 weeks in the case of ocular dominance columns in the kitten (Kandel *et al* 2000). In somatosensory cortex the sensory maps arise from movement-based reafference (Merzenich and Jenkins 1993). We see this process as evolving adaptively, bootstrapped by spontaneous fetal movement. Only those reafferent signals that remain correlated over many weeks across the full repertoire of changing movement patterns end up wired together in the somatosensory mappings. We propose that these maps define elemental movements $\boldsymbol{\theta}$ (based on signals from joint and skin

mechanoreceptors) and functional muscles in terms of lengths **l** (from muscle spindle endings) and tensions **t** (from Golgi tendon organs).

The second stage of establishing a minimum-energy **WSG** is to form an internal model of the static nonlinear (polynomial) relationships between the flexion/extension components of elemental movements $\boldsymbol{\theta}$ and muscle lengths **l**. This is shown in figure 2 as the MOVEMENT-MUSCLE MAP. We propose that the modeling of signals $\boldsymbol{\theta}$ and **l** from the sensory maps is achieved by means of least mean square (LMS) neural processing described in section 7. The nonlinear relationships between $\boldsymbol{\theta}$ and **l** are algebraic and are not changed by factors such as orientation in the gravitational field or variation in mechanical loads. They do alter with long-term change such as growth or surgical reconstruction and the model will retune over that time frame. Just as for the sensory maps this tuning process must be slow in order to break up short-term correlations between elemental movements. The nonlinearity of the relationship between $\boldsymbol{\theta}$ and **l** means that its slope changes with operating point, i.e., with body posture. This slope is specified at each operating point by $\boldsymbol{\Delta l/\Delta\theta}$ that can be derived from the model[2]. It is well known in the field of biomechanics that the derivative $\Delta l_i/\Delta\theta_j$ provides a measure of the moment arm of the *i*th muscle about the *j*th movement. Thus modulating the gains of descending pathways in the **WSG** by the approximate Jacobian matrix $\boldsymbol{\Delta l/\Delta\theta}$ derived from the model, as in figure 2, focuses descending connectivity so muscles are activated in an appropriate synergy to achieve required turning forces and co-contractions with a minimum demand by the muscles for metabolic energy regardless of posture.

This approximate Jacobian matrix is exactly what is required to modulate the **WSG** so that it implements a minimum energy synergy. Changes of descending pathway gains with movement and with fatigue have been shown experimentally using transcranial magnetic stimulation (see Gandevia (2001), Taylor and Gandevia (2001)) and some forms of so-called central muscle fatigue are attributed to reduction in gains of descending pathways. We suggest that modulation of descending pathway gains for posture (as above) and modulation for fatigue may be additive processes, the latter due to a separate mechanism that detects and compensates for changes in the relationships between $\alpha$ drives and the resulting muscle tensions.

## 4. Solving the problem of limited central resources

While the problem of redundancy in the musculoskeletal system can be solved by self-organizing connectivity within the descending motor pathways, the problem of limited central resources remains. By definition, each of the 110 elemental movements, and the 110 co-contractions about those movements, can be controlled independently one at a time, but it is not possible to achieve independent control of all 220 movements and co-contractions simultaneously. We have limited central processing resources (Marois and Ivanoff 2005, Wickens 1984) and common experience tells us that our central control capability allows only a limited number of degrees of freedom of movement to be controlled at the same time.

### *4.1. Control degrees of freedom*

It is sometimes stated that in movement control we 'coordinate' the activity of many degrees of freedom DFs) in order to perform a task. This only makes sense if 'degrees of freedom' is interpreted as meaning the *biomechanical* DFs in a movement. Consider the wired skeleton found in an anatomy classroom. For such a skeleton one can find maybe 150 independent joint rotations and translation movements. Accordingly, we could say that the skeleton has 150 biomechanical DFs. To illustrate, moving the arm of the skeleton involves three movements of the scapula, three rotations at the shoulder, one rotation at the elbow, one rotation of the forearm, two rotations at the wrist, and maybe 15 rotations involving fingers and thumb, a total of say 25 biomechanical DFs. In other words, 25 independent movements can be simultaneously imposed on the forelimb of the skeleton.

Now imagine an intact alive human in a motion analysis laboratory equipped to record these 25 biomechanical DFs of arm movement. Suppose the subject performs a task involving this full movement range and the 25 components are displayed as time signals. We can now ask how many DFs are contained in this set of 25 signals. To emphasize the distinction about to be made, take an extreme example where, after amplitude scaling, all 25 signals have exactly the same waveform. In this case all signals are perfectly correlated and thus have only one DF. While all 25 biomechanical DFs of arm movement are involved, the resultant kinematic pattern involves only one DF of movement from the point of view of control. We will refer to this as a *control* degree of freedom of movement.

Consider now the sets of 25 signals measured from a subject seated at a table and performing a reach and grasp task to objects of varying size at varying distances and directions. Many such experiments have been done, although never with records of all 25 biomechanical DFs but the evidence from diverse studies (see Jeannerod (1999), Krakauer and Ghez (2000), Smeets and Brenner (1999)) points clearly to there being only two independent control DFs in this particular behavior. Such objective findings clearly endorse the fact that what is commonly described as 'coordination' in purposeful movement often arises from strong interrelationships between the underlying elemental movements. The hallmark of a purposeful, coordinated movement is the fact that the number of control DFs is much less than the underlying number of biomechanical DFs. Indeed, if we imagine a forelimb movement in which all 25 biomechanical DFs vary independently and simultaneously, it would appear uncoordinated and purposeless. We now argue that such a controlled movement is impossible, even if we had good reason to generate it, and therein lies the problem of limited central resources.

[2] We postulate that a slave copy of the internal model of the nonlinear static relationships between $\boldsymbol{\theta}$ and **l** signals is used in descending control of gamma motor neurons as part of the reflex control system. The ratio of signals $\boldsymbol{\Delta l/\Delta\theta}$ can be derived from this slave model making modulation of the gains of descending pathways independent of feedback signals.

The fact that the forelimb has 25 biomechanical DFs of movement that can be controlled independently, one at a time, leads to the mistaken impression that the nervous system has available an independent dedicated controller for each such elemental movement. If this were the case, a coordinated forelimb movement with only, say, two control DFs would require the same amount of central control resources (if not more) than that required for 25 control DFs (i.e., simultaneous, independent control of all 25 elemental movements). Thinking back to the classroom skeleton, this means we need 25 students, each assigned to a joint, to produce any movement, coordinated or uncoordinated. Indeed, producing a coordinated movement may require extra team members to impose synchrony. Clearly this is not the way it works. It is self-evident that it is less demanding to produce a movement that has few control DFs than one that has many. Indeed, simultaneous independent control of all 25 biomechanical DFs of the forelimb is just not possible, let alone of all 220 elemental movements and co-contractions of the entire musculoskeletal system. We simply do not have a supply of independent central controllers to match the biomechanical DFs available. This problem of limited central resources forces us to focus on the control DFs of the system that are far fewer in number.

Just how many DFs can be controlled simultaneously is uncertain. While this remains to be determined experimentally it is known that simultaneous control of at least two DFs is definitely achievable. As indicated earlier, reach and grasp behavior is well established experimentally as having at least two underlying control DFs. Also, subjects can perform a two DF visual pursuit tracking task with no greater a central time delay than required to perform a one DF task (Navon *et al* 1984, Oytam *et al* 1998, 2005). This is an important finding. It shows that central mechanisms can plan and execute at least two independent responses simultaneously without interference. If the CNS can plan two dimensions of response in parallel, perhaps it can plan more than two.

Let us say that the CNS has sufficient dedicated neural circuitry to plan as many as $N_m$ independent dimensions of response simultaneously without interference. But how many is $N_m$? Intuitively, if we think of a scenario such as carrying on a conversation while driving an automobile in a busy street, we might conclude that $N_m$ is somewhere between 5 and 20 control DFs. One thing for certain, it is not possible to control all 220 elemental movements and co-contractions independently, simultaneously. What the CNS can do, we argue, is link large numbers of elemental movements and co-contractions together to form functional multijoint movements or *task-dependent synergies* that have only a small number of control DFs. In other words, implementing the planning and control of movement in a low-dimensional task space can solve the problem of limited resources.

In the following sections we show how this can be achieved in the framework of AMT. We propose that during each and every action, real-time networks of adaptive nonlinear dynamical filters continuously extract a small set of task-dependent orthogonalized sensory feature signals from the many interrelated reafferent signals produced during that action (i.e., from efference copy, from tension feedback, from elemental movement and co-contraction feedback, and from feedback of responses via the external perceptual-motor loop). We see this on-line self-organizing process of feature extraction as an intrinsic part of real-time sensory analysis. Likewise we propose that to produce a particular action, real-time networks of neural adaptive filters continuously convert a small set of orthogonal central motor command signals into the many intercorrelated peripheral motor signals required to produce the action (that in turn will give rise to the intended reafference). Thus extraction of orthogonal feature signals from efference copy and generation of synergetic motor command signals are complementary. The adaptive filter parameters that define the sensory features of the efference that produced an action are precisely the same parameters that will generate the correct motor synergy for that action from a small set of orthogonal central commands. The deorthogonalizing process is thus the basis for *task-dependent synergy generation*.

The dimension of the orthogonal feature signals for any particular task equals the number of control DFs in the task and hence defines the dimensionality of the task space. This means that the planning of response trajectories to achieve perceptual goals and the transformation of those trajectories into appropriate central motor commands can be implemented using low-dimensional orthogonal feature signals **M**, **T**, **Q** and **R** rather than the corresponding high-dimensional signals $\mathbf{m}_\alpha$, **t**, **q** and **r** (we use the convention of upper case for orthogonalized signals and lower case for the unorthogonalized equivalents throughout the text and figures). This large reduction in dimensionality means that processing demand is minimized thus solving the central resources problem. In the remainder of this section we show schematically how this can be achieved by means of the adaptive neural filters of AMT. These are inherently self-organizing, nonlinear and dynamic and are described further in section 7.

### 4.2. Orthogonalizing networks

An adaptive filter model of the relationship between an input and an output signal enables the output to be partitioned into two components, one related to the input, the other unrelated (i.e., orthogonal). The related component is the filter output and the orthogonal component is the difference between the output and the filter output signals. Based on this it is possible to construct a network of adaptive filters that can tune on-line to extract a set of orthogonal signals from a set of interrelated (i.e., nonorthogonal) input signals (Neilson 1993b). The structure of such a network is illustrated in figure 3(a) where inputs $u_1(k)$, $u_2(k)$, $u_3(k)$, $u_4(k)$ pass through nonlinear dynamic adaptive filters $r_{21}$, $r_{31}$, $r_{41}$, $r_{32}$, $r_{42}$, $r_{43}$ to produce orthogonal signals $U_1(k)$, $U_2(k)$, $U_3(k)$, $U_4(k)$. Similar structures apply for any number of inputs.

To clarify how the network in figure 3(a) functions, consider the adaptive filter $r_{21}$. The input signal $u_1(k)$ is transformed through $r_{21}$ to give an output signal that is subtracted from input signal $u_2(k)$. This gives the signal $U_2(k)$.

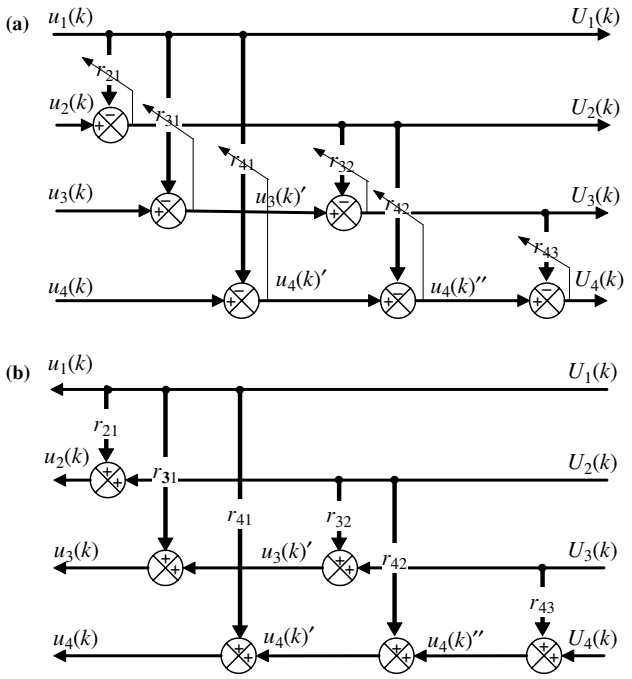



**Figure 3.** (a) Orthogonalizing network: schematic illustration of a network of nonlinear dynamic adaptive filters $r_{21}$, $r_{31}$, $r_{41}$, $r_{32}$, $r_{42}$, $r_{43}$ that tune on-line to transform a set of interrelated input signals $u_1(k)$, $u_2(k)$, $u_3(k)$, $u_4(k)$ into a set of orthogonal output signals $U_1(k)$, $U_2(k)$, $U_3(k)$, $U_4(k)$. Diagonal arrows indicate the error signal used to tune each adaptive filter. The structure of the network is based on the Gram-Schmidt orthogonalizing algorithm as explained in the text. (b) Deorthogonalizing network: schematic diagram illustrating how the same network shown in (a) using slave copies of the nonlinear dynamic adaptive filters $r_{21}$, $r_{31}$, $r_{41}$, $r_{32}$, $r_{42}$, $r_{43}$ and with the direction of flow reversed and subtractions changed to additions can perform the reverse operation and transform the set of orthogonal signals $U_1(k)$, $U_2(k)$, $U_3(k)$, $U_4(k)$ back into the set of interrelated signals $u_1(k)$, $u_2(k)$, $u_3(k)$, $u_4(k)$.

If $r_{21}$ does not accurately specify the relationship between inputs $u_1(k)$ and $u_2(k)$, the signal $U_2(k)$ functions as an error signal. Adaptive tuning based on a least mean square algorithm adjusts the parameters of $r_{21}$ to minimize the variance of $U_2(k)$. When this tuning is accurate, signal $U_2(k)$ specifies the component of $u_2(k)$ that is not related to $u_1(k) = U_1(k)$. In other words, signals $U_1(k)$ and $U_2(k)$ are orthogonal. Similar tuning of adaptive filters $r_{31}$, $r_{41}$, $r_{32}$, $r_{42}$, $r_{43}$ results in the set of orthogonalized output signals $U_1(k)$, $U_2(k)$, $U_3(k)$, $U_4(k)$.

Although the number of outputs from such a network always equals the number of inputs, many of the outputs will be negligible if the inputs are substantially interrelated. In other words, a large number of interrelated inputs can flow through an orthogonalizing network to give a small number of non-zero orthogonal outputs. For example if $u_1(k)$, $u_2(k)$, $u_3(k)$, $u_4(k)$ in figure 3(a) were all perfectly related, the network would extract only one non-zero orthogonal signal $U_1(k)$. The other outputs would converge to zero. In this case the orthogonalizing network requires only the adaptive filters labeled $r_{21}$, $r_{31}$ and $r_{41}$ and the other adaptive filters can be discarded. Thus an orthogonalizing network with a large number of input signals can be greatly simplified if these inputs are strongly interrelated with only a few control DFs.

The network of figure 3(a) offers a general way of orthogonalizing signals in real time. The meaningful information contained in a set of interrelated signals will always be captured in a reduced set of orthogonalized signals (and the remaining noise signals discarded) provided the adaptive filters $r_{ij}$ are able to adequately model the relationships between the inputs. In the simplest case of a purely algebraic relationship between the inputs, the network provides an on-line implementation of what is known as Gram–Schmidt orthogonalization, or Q–R factorization. In this circumstance each of the $r_{ij}$ reduces to essentially a single parameter, namely a regression coefficient, but if these change over time, the adaptive filter circuitry will incorporate the changes as they happen. Likewise if the inputs are known to have a linear dynamic relationship the filters need only implement an adaptive least mean squares algorithm for each parameter in the filter. The methodology is well understood (e.g., Widrow and Stearns (1985)). More challenging is the case of nonlinear dynamic relationships between signals that are likely to be non-Gaussian. In section 7 we present a newly-developed solution for self-tuning adaptive filter modeling of such relationships. What results is a type of on-line nonlinear dynamic independent components analysis (Cichocki and Amari 2002). This gives a means of implementing an orthogonalizing network using adaptive filters $r_{ij}$ that can self-tune to nonlinear dynamic relationships like those between sensory and motor signals.

The orthogonalizing network (figure 3(a)) has a crucial property. It can operate in the reverse direction (figure 3(b)). Simply by changing the direction of signal flow and replacing subtractors with summers, the nonlinear dynamic adaptive filters $r_{21}$, $r_{31}$, $r_{41}$, $r_{32}$, $r_{42}$, $r_{43}$ of figure 3(a) transform orthogonal signals $U_1(k)$, $U_2(k)$, $U_3(k)$, $U_4(k)$ back into interrelated signals $u_1(k)$, $u_2(k)$, $u_3(k)$, $u_4(k)$ in figure 3(b). If some of the $U_1(k)$, $U_2(k)$, $U_3(k)$, $U_4(k)$ signals are zero, the reverse transformation still works. Thus a reverse deorthogonalizing network can reconstruct the original large set of signals from a small set of orthogonal signals.

### *4.3. Feature extraction and synergy generation*

The orthogonalizing network of figure 3(a) gives a means by which the nervous system could implement a real-time process to convert a large set of interrelated sensory signals from the periphery into a small set of orthogonal sensory feature signals represented centrally. Likewise, the deorthogonalizing network of figure 3(b) gives a means by which the nervous system could implement a real-time process to convert a small set of orthogonal central motor command signals into a large set of appropriately interrelated peripheral motor commands. In each case the conversion preserves all the meaningful information in the original set of signals. In AMT we propose this as the mechanism of the continuous real-time task-dependent processes of sensory feature extraction and motor synergy generation.

Figure 4 illustrates how these processes underlie the performance of a task that has $N$ control DFs. It depicts both processes in the context of the signals and systems established

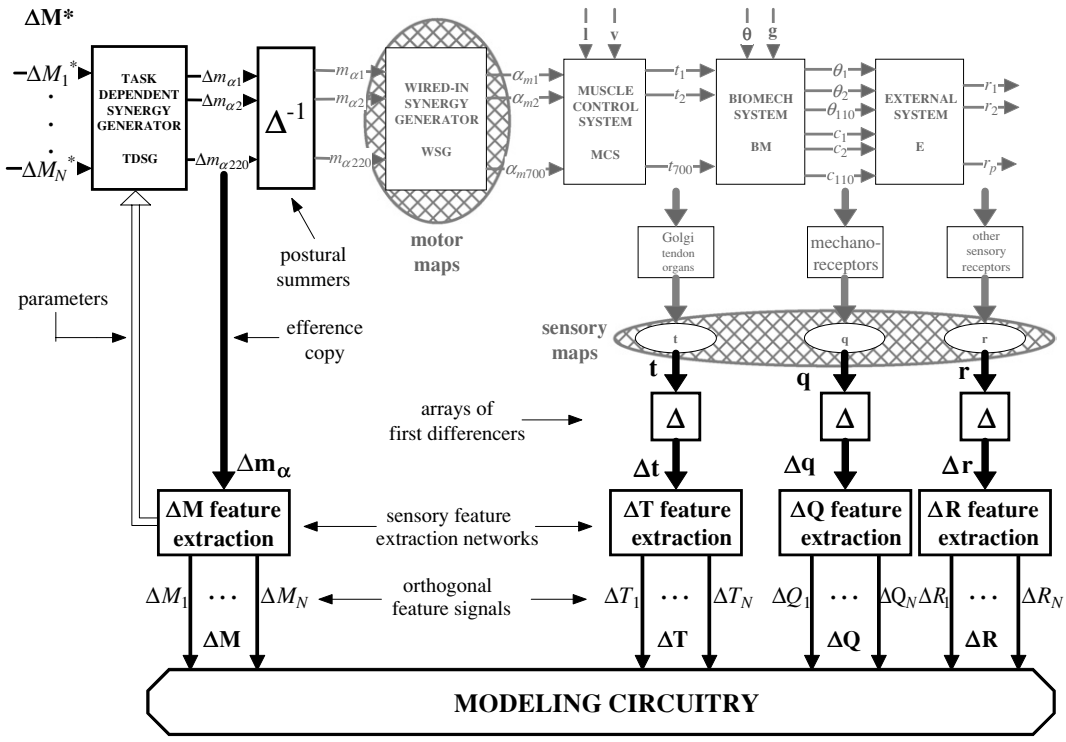



**Figure 4.** A schematic diagram illustrating extraction of orthogonal feature signals and adaptive formation of a TASK DEPENDENT SYNERGY GENERATOR (**TDSG**): The gray part of the figure derives from figures 1 and 2 and will not be described again here. Vectors of tension signals $\mathbf{t} = [t_1, \ldots, t_{700}]^{\mathrm{T}}$, elemental movements and co-contractions $\mathbf{q} = [q_1, \ldots, q_{220}]^{\mathrm{T}}$, and reafferent responses $\mathbf{r} = [r_1, \ldots, r_p]^{\mathrm{T}}$ derived from sensory maps in the cortex are transformed through arrays of first-difference filters in the blocks labeled $\boldsymbol{\Delta}$ to give vectors of first-differenced signals $\boldsymbol{\Delta t} = [\Delta t_1, \ldots, \Delta t_{700}]^{\mathrm{T}}$, $\boldsymbol{\Delta q} = [\Delta q_1, \ldots, \Delta q_{220}]^{\mathrm{T}}$, and $\boldsymbol{\Delta r} = [\Delta r_1, \ldots, \Delta r_p]^{\mathrm{T}}$. A vector of first-differenced central motor commands $\boldsymbol{\Delta m}_\alpha = [\Delta m_{\alpha 1}, \ldots, \Delta m_{\alpha 220}]^{\mathrm{T}}$ is assumed to be available via efference copy as discussed in the text. The vectors of differenced signals $\boldsymbol{\Delta m}_\alpha$, $\boldsymbol{\Delta t}$, $\boldsymbol{\Delta q}$, $\boldsymbol{\Delta r}$ are transformed by networks of nonlinear dynamic adaptive filters labeled $\boldsymbol{\Delta M}$ feature extraction, $\boldsymbol{\Delta T}$ feature extraction, $\boldsymbol{\Delta Q}$ feature extraction, $\boldsymbol{\Delta R}$ feature extraction that tune on-line to obtain vectors of $N$-dimensional orthogonal feature signals $\boldsymbol{\Delta M} = [\Delta M_1, \ldots, \Delta M_N]^{\mathrm{T}}$, $\boldsymbol{\Delta T} = [\Delta T_1, \ldots, \Delta T_N]^{\mathrm{T}}$, $\boldsymbol{\Delta Q} = [\Delta Q_1, \ldots, \Delta Q_N]^{\mathrm{T}}$, $\boldsymbol{\Delta R} = [\Delta R_1, \ldots, \Delta R_N]^{\mathrm{T}}$. The number $N$ is task dependent and represents the number of control degrees of freedom in the task. The adaptive filter networks in the blocks labeled feature extraction are based on the orthogonalizing network in figure 3(a). The adaptive filter network in the block labeled TASK DEPENDENT SYNERGY GENERATOR (**TDSG**) is based on the deorthogonalizing network in figure 3(b). Adaptive filter parameters in the **TDSG** are copies of the adaptive filter parameters tuned on-line in the $\boldsymbol{\Delta M}$ feature extraction network. The **TDSG** transforms an N-vector of centrally generated orthogonal first-differenced required motor commands $\boldsymbol{\Delta M}^* = [\Delta M_1{}^*, \ldots, \Delta M_N{}^*]^{\mathrm{T}}$ into a 220 dimensional vector of interrelated first-differenced central motor commands $\boldsymbol{\Delta m}_\alpha = [\Delta m_{\alpha 1}, \ldots, \Delta m_{\alpha 220}]^{\mathrm{T}}$. The required orthogonal first-differenced feature signals $\boldsymbol{\Delta R}^*$, $\boldsymbol{\Delta T}^*$, $\boldsymbol{\Delta Q}^*$ mentioned in the text occur at levels above $\boldsymbol{\Delta M}^*$ and are not shown in this diagram. Forward and inverse nonlinear dynamical relationships between the vectors of orthogonal first-differenced feature signals $\boldsymbol{\Delta M}$ and $\boldsymbol{\Delta T}$, $\boldsymbol{\Delta T}$ and $\boldsymbol{\Delta Q}$, $\boldsymbol{\Delta Q}$ and $\boldsymbol{\Delta R}$ are modeled on-line by adaptive filters in the block labeled MODELING CIRCUITRY.

in figures 1 and 2. In overview, the input $\boldsymbol{\Delta M}^*$ (top left) represents a $N$-vector of first-differenced orthogonal motor commands that have been generated centrally by a feedforward controller (see sections 5 and 6) to produce a first-differenced orthogonal response vector $\boldsymbol{\Delta R}$ (bottom right). We use the superscript $*$ to indicate *required* signals such as $\boldsymbol{\Delta M}^*$, $\boldsymbol{\Delta T}^*$, $\boldsymbol{\Delta Q}^*$, $\boldsymbol{\Delta R}^*$ generated centrally to achieve an intended response $\mathbf{R}^*$. Such signals are sometimes referred to in the literature as *intended*, *desired* or *reference* signals. If all the intervening transformations are correct the vector $\boldsymbol{\Delta M}^*$ will supply exactly the changes required to implement a required response $\mathbf{R}^*$ in the $N$-space of the task. The transformations include a TASK-DEPENDENT SYNERGY GENERATOR (**TDSG**) that generates synergetic motor command changes $\boldsymbol{\Delta m}_\alpha = [\Delta m_{\alpha 1}, \ldots, \Delta m_{\alpha 220}]^{\mathrm{T}}$. These are summed by postural summers $\boldsymbol{\Delta}^{-1}$ and passed to the wired-in synergy generator **WSG** that provides the minimum energy peripheral motor commands $\boldsymbol{\alpha}_{\mathbf{m}} = [\alpha_{m1}, \ldots, \alpha_{m700}]^{\mathrm{T}}$ as in figure 2, resulting in the required response changes $\boldsymbol{\Delta R} = [\Delta R_1, \ldots, \Delta R_{\mathrm{N}}]^{\mathrm{T}}$.

The processes shown in gray have been detailed earlier, so we now focus on the new components of the figure, beginning with the summers $\boldsymbol{\Delta}^{-1}$ and first differencers $\boldsymbol{\Delta}$. Postural summers $\boldsymbol{\Delta}^{-1}$ hold motor commands $\mathbf{m}_\alpha$ on line, so only changes $\boldsymbol{\Delta m}_\alpha$ need to be processed centrally. This gives a further way of reducing demand on resources since it allows unchanged $\alpha$ drives to be supplied to all those muscles that maintain fixed aspects of posture in a task. But it also brings implications for establishing suitable internal models for use in the feedforward controller just mentioned. We show in section 5 that formation of this controller requires models of both the forward and inverse dynamics of the plant. For internal models of nonlinear dynamical systems to perform accurately their adaptive tuning must be based on signals with statistical properties similar to those that they will ultimately

transform (Widrow and Walach 1996). Inclusion of postural summers $\mathbf{\Delta}^{-1}$ implies that the models will transform first-differenced signals. Consequently, it is important that their adaptive tuning is based on first-differenced signals. Thus the sensory feedback signals **t**, **q** and **r** pass through arrays of first-differencers $\mathbf{\Delta}$ to give signals $\mathbf{\Delta t}$, $\mathbf{\Delta q}$ and $\mathbf{\Delta r}$. The vector of central motor commands $\mathbf{\Delta m}_\alpha$ is already differenced and, for reasons given below, we postulate that this supplies efference copy. It is therefore processed the same way as the other sensory inputs, as follows.

Each of the differenced signals $\mathbf{\Delta m}_\alpha$, $\mathbf{\Delta t}$, $\mathbf{\Delta q}$ and $\mathbf{\Delta r}$ flows in real time through its own orthogonalizing network, labeled respectively $\mathbf{\Delta M}$-, $\mathbf{\Delta T}$-, $\mathbf{\Delta Q}$- and $\mathbf{\Delta R}$-feature extraction (figure 4). The feature extraction networks remove redundancy from each set of highly interrelated inputs $\mathbf{\Delta m}_\alpha = [\Delta m_{\alpha 1}, \ldots, \Delta m_{\alpha 220}]^{\mathrm{T}}$, $\mathbf{\Delta t} = [\Delta t_1, \ldots, \Delta t_{700}]^{\mathrm{T}}$, $\mathbf{\Delta q} = [\Delta q_1, \ldots, \Delta q_{220}]^{\mathrm{T}}$ and $\mathbf{\Delta r} = [\Delta r_1, \ldots, \Delta r_p]^{\mathrm{T}}$ by extracting the orthogonal, independently varying feature signals $\mathbf{\Delta M} = [\Delta M_1, \ldots, \Delta M_N]^{\mathrm{T}}$, $\mathbf{\Delta T} = [\Delta T_1, \ldots, \Delta T_N]^{\mathrm{T}}$, $\mathbf{\Delta Q} = [\Delta Q_1, \ldots, \Delta Q_N]^{\mathrm{T}}$ and $\mathbf{\Delta R} = [\Delta R_1, \ldots, \Delta R_N]^{\mathrm{T}}$. Depending on the number of control DFs in the task that gives rise to the signals, $N$ maybe as low as 1 or as high as $N_{\mathrm{m}}$ which we previously supposed to be possibly 20. As in figure 3(a) the adaptive parameters in the filters of each network tune to the nonlinear dynamic relationships between the signals. These parameters define the task-dependent models that underlie the on-line extraction of orthogonal sensory feature signals.

The process of extracting a set of orthogonal features $\mathbf{\Delta M}$ for a task leads directly to establishing the task-dependent synergy generator **TDSG** based on the equivalent deorthogonalizing network (figure 3(b)). As shown in figure 4, copies of the parameters from the adaptive filters in the $\mathbf{\Delta M}$ feature extraction network tune the adaptive filters in the **TDSG** so that the adaptive formation of the $\mathbf{\Delta M}$ feature extraction network and the **TDSG** are one and the same process. Once a **TDSG** is established for a task with $N$ control DFs it provides the means to transform a small number $N$ of central commands $\mathbf{\Delta M}^*$ into a larger number (220) of interrelated commands $\mathbf{\Delta m}_\alpha$ that in turn drive the **WSG** and subsequent systems.

In figure 4 the vector of first-differenced central motor commands $\mathbf{\Delta m}_\alpha$ is labeled as efference copy. Why? Although the notion of efference copy is long established (von Holst and Mittelstaedt 1950/1973) as is experimental evidence that it provides sensory information (McCloskey 1981), its source remains unclear. Some recent work focuses on central origins in primary motor cortex (Gross *et al* 2002), or perhaps upstream in premotor cortex (Chronicle and Glover 2003) while other investigations concentrate on efference copy as a 'cancellation' signal at brainstem level (Cullen and Roy 2004). From the point of view of extracting orthogonal feature signals $\mathbf{\Delta M}$, the system in figure 4 would work equally well with efference copy derived from the vector of peripheral motor commands $\boldsymbol{\alpha}_{\mathbf{m}}$ or from central commands $\mathbf{m}_\alpha$. The orthogonal feature signals $\mathbf{\Delta M}$ would have the same dimension, but the parameters to generate them would differ. However, the $\mathbf{\Delta M}$ feature extraction network does more than extract the signals $\mathbf{\Delta M}$. As shown by the open arrow at the left of the figure, the parameters of the nonlinear adaptive filters of the $\mathbf{\Delta M}$ network also tune the **TDSG**. For this to work it is essential that efference copy is obtained centrally, above the **WSG**. Feature extraction based on the 700-dimensional vector $\boldsymbol{\alpha}_{\mathbf{m}}$ would involve a different network structure to that required for the **TDSG** and the proposal would not work. Also, note that if $\mathbf{m}_\alpha$ rather than $\mathbf{\Delta m}_\alpha$ were taken as the input for efference copy feature extraction, an additional array of first differencers $\mathbf{\Delta}$ would be required, as for signals **t**, **q** and **r**. Since $\mathbf{\Delta m}_\alpha$ signals are already available at the inputs to the postural summers $\mathbf{\Delta}^{-1}$ it is parsimonious to designate these as supplying efference copy rather than $\mathbf{m}_\alpha$.

By definition of elemental movements (section 2), the structure of the **WSG**, **MCS** and **BM** systems is such that each elemental movement and co-contraction can be controlled independently by the vector $\mathbf{\Delta m}_\alpha$ of 220 central command changes. Providing nonlinear dynamics are taken into account during feature extraction, it follows that the constraints introduced by the **TDSG** into the relationships between these 220 command changes are similarly introduced into the relationships between the 700 $\alpha$ motor neuron command changes $\mathbf{\Delta}\boldsymbol{\alpha}_{\mathbf{m}}$, the 700 muscle tension changes $\mathbf{\Delta t}$, and the resulting changes in the 220 elemental movements and co-contractions $\mathbf{\Delta q}$. In other words, the number of control DFs in the vectors $\mathbf{\Delta}\boldsymbol{\alpha}_{\mathbf{m}}$, $\mathbf{\Delta t}$ and $\mathbf{\Delta q}$ equals the number of control DFs in $\mathbf{\Delta m}_\alpha$ (Neilson and Neilson 2004). But the control DFs in $\mathbf{\Delta m}_\alpha$ are equal to the control DFs in $\mathbf{\Delta M}$ and $\mathbf{\Delta M}^*$ that in turn equal (by definition) the $N$ control DFs in the response vector $\mathbf{\Delta R}$. Thus all the vectors of orthogonal feature signals $\mathbf{\Delta M}$, $\mathbf{\Delta T}$, $\mathbf{\Delta Q}$ and $\mathbf{\Delta R}$ are $N$-dimensional.

### *4.4. Forward and inverse internal models of **MCS**, **BM** and **E** systems*

As it turns out, the process of extracting orthogonal feature signals and forming task-dependent synergies does more than address the problem of limited central resources. It also solves two extra difficulties that would arise if signals $\mathbf{\Delta m}_\alpha$, $\mathbf{\Delta t}$, $\mathbf{\Delta q}$ and $\mathbf{\Delta r}$ were to be used directly in modeling the **MCS**, **BM** and **E** subsystems. From system identification and adaptive filter theory (Cichocki and Amari 2002, Haykin 1986, Ljung 1987, Widrow and Stearns 1985) it is known that (i) it is not possible to model the relationship between multiple inputs and multiple outputs if the inputs are highly interrelated and (ii) it is not possible to model the inverse of a multiple input/multiple output system if the system is redundant; i.e., more inputs than outputs. Both these conditions apply to **MCS**, **BM** and **E**.

A simplified version of the problem of interrelated inputs will be familiar from multiple regression analysis (Kenny 1979) where the solution involves partialling out the relationships between the inputs. This is exactly what is done here but in a setting of nonlinear dynamical relationships. In AMT we solve the interrelationship problem and the redundancy problem simultaneously by virtue of the fact that the $\mathbf{\Delta M}$, $\mathbf{\Delta T}$, $\mathbf{\Delta Q}$ and $\mathbf{\Delta R}$ feature signals in figure 4 are not only orthogonal but also of equal dimension $N$. They have exactly the properties necessary for obtaining accurate

and well-conditioned models of multivariable systems; hence they are shown connecting to the block labeled MODELING CIRCUITRY in figure 4. This consists of adaptive filter circuits (section 7) that form task-dependent forward and inverse models of the nonlinear dynamical relationships between each set of feature signals. The relationship between $\mathbf{\Delta M}$ and $\mathbf{\Delta T}$ provides models of **MCS** and its inverse $\mathbf{MCS}^{-1}$, that between $\mathbf{\Delta T}$ and $\mathbf{\Delta Q}$, models of **BM** and $\mathbf{BM}^{-1}$, and that between $\mathbf{\Delta Q}$ and $\mathbf{\Delta R}$ models of **E** and $\mathbf{E}^{-1}$. We designate these models respectively $\mathbf{MCS}_{\mathrm{m}}$ and $\mathbf{MCS}_{\mathrm{m}}{}^{-1}$, $\mathbf{BM}_{\mathrm{m}}$ and $\mathbf{BM}_{\mathrm{m}}{}^{-1}$, and $\mathbf{E}_{\mathrm{m}}$ and $\mathbf{E}_{\mathrm{m}}{}^{-1}$, emphasizing that they represent the nonlinear dynamical relationships between orthogonalized first-differenced feature signals $\mathbf{\Delta M}$, $\mathbf{\Delta T}$, $\mathbf{\Delta Q}$ and $\mathbf{\Delta R}$ (rather than between the fully-dimensioned undifferenced equivalents $\mathbf{m}_{\alpha}$, **t**, **q** and **r** that define the actual **MCS**, **BM** and **E** systems). These are the models required for the feedforward/feedback movement controller discussed in later sections.

To summarize, introduction of sensory feature extraction and task-dependent synergy generation addresses the problem of limited central resources by minimizing the dimensionality of response planning and control. For a task with $N$ control DFs all control processes at levels above the **TDSG** involve the $N$-dimensional vectors $\mathbf{\Delta M}$, $\mathbf{\Delta T}$, $\mathbf{\Delta Q}$ and $\mathbf{\Delta R}$ or their desired equivalents $\mathbf{\Delta M}^*$, $\mathbf{\Delta T}^*$, $\mathbf{\Delta Q}^*$ and $\mathbf{\Delta R}^*$ in the sensory feature space of the task. These processes never have to involve the full dimensionality of the controlled systems.

## 5. Solving the problem of nonlinear interactions

The central controller has the task of producing a vector of motor commands $\mathbf{\Delta M}^*$ to drive the system in figure 4 to produce a response **R** that matches as accurately as possible an intended response $\mathbf{R}^*$. In other words, the relationship between **R** and $\mathbf{R}^*$ must approach an identity operator. To achieve this the controller has to generate an appropriate $\mathbf{\Delta M}^*$ to compensate for the complex nonlinear dynamical interactions in the relationships that characterize the **MCS**, **BM** and **E** systems. The **BM** system, for example, is highly interactive. A turning force applied about one joint of the skeleton causes rotations about many other joints. Inertial, viscous, centrifugal, coriolis and gravitational forces all contribute to this mechanical coupling between joints. The strength of each reaction force varies as a function of posture and orientation in the gravitational field and with interactions between body and environment. We pick up, put down, push, pull, rotate and squeeze all sorts of objects in the environment and we stand, lean, sit and lie in an infinite variety of ways on all sorts of surfaces. All such interactions with the environment profoundly change the nonlinear dynamics of the **BM** system. Even movement at a single joint, such as abduction of the arm about the shoulder joint, requires activation of muscles throughout the body, not just at the shoulder, to maintain balance and to compensate for interactive reaction forces.

Despite the complexity and uncertainty of these changing nonlinear dynamics, the CNS is able to control rotations of individual joints and generate pre-specified coordinated movements, such as those of a ballerina or gymnast, with relative ease. To illustrate this phenomenon imagine performing the following task. An experimenter puts a finger in the air and moves it about slowly in three dimensions. Your task is to place a fingertip close to that of the experimenter's without actually touching and to track the movement as accurately as possible. Providing the experimenter moves slowly this task is not difficult. The position of your fingertip in 3D space will closely match that of the experimenter. Indeed, as shown repeatedly in visual tracking experiments (McRuer and Krendel 1974, Neilson *et al* 1995), the relationship between the positions of the two fingers can be described by a relatively simple linear dynamical transfer function consisting of a gain, a time delay, and a first or second order linear filter. But if we attempt to analyze the dynamics involved in the finger task, which equates to controlling some 58 functional muscles acting on some 25 elemental movements to track a 3D position under influence of a gravitational field, we are faced with an immensely complex system of multivariable, unstable, nonlinear differential equations. Yet all these complex dynamics are unobservable in the target–response relationship. As with the finger task, it is common experience that the movements we make closely match the ones we intend to make. So what has happened to the complex interactive dynamics between central commands and body movements? This is the case of the disappearing dynamics.

From experimental work on reaching and grasping tasks, Jeannerod (1981, 1999) was aware of the remarkable absence of interaction in the task components and observed that the CNS seems to behave as if it has separate parallel visuomotor channels for each control DF of movement (see Krakauer and Ghez (2000)). Jeannerod argued for an anatomical basis for such channels. In contrast, we suggest that it is the ability of the CNS to form adaptive internal models that allows it to cancel all the interactive nonlinear dynamics while simultaneously providing independent control over an appropriate reduced set of control DFs as in section 4. Developments in the theory of control of nonlinear dynamical systems (Isidori 1995, Jurdjevic 1997, Khalil 1996, Nijmeijer and van der Shaft 1996, Slotine and Li 1991) provide insight into how this might be achieved.

### *5.1. Noninteractive feedback control of a nonlinear dynamical plant*

We have recently used nonlinear control theory to explore how the CNS might provide noninteractive control in human movement (Neilson and Neilson 2001, 2004). We show mathematically how it is possible to design a state feedback controller for a square multiple input/multiple output nonlinear dynamical plant that renders all of the plant dynamics (other than time delays) unobservable in the output signals and, at the same time, decouples the closed-loop input/output relationships so that the closed-loop system behaves like an array of parallel single input/single output channels. Design of this controller is based on the so-called *zero dynamics algorithm* (see Isidori (1995)).

In essence, the feedback controller is designed so that state feedback partitions the state space into a 'foliation' of output contour submanifolds. Outputs are constant for all

states within such an output contour submanifold. If the inputs to the controller are held constant, the state of the system is constrained by state feedback to remain within an output contour submanifold. Consequently, any change of state within an output contour submanifold leaves the outputs unchanged and so the state changes are unobservable in the output signals. On the other hand, changes of the input signals to the controller cause the state of the plant to move from one output contour submanifold to the next. This causes the output signals to change but the state feedback is so organized that all of the dynamics of the plant, except for time delays, are unobservable in the input/output relationships and, at the same time, the input/output relationships are rendered noninteractive. Thus input 1 controls output 1 via a time delay, input 2 controls output 2 via another time delay, and so on, with no cross coupling between channels. Just as observed in movement control, the state feedback controller compensates for all the nonlinear dynamical interactions within the plant and effectively creates an array of parallel, independent channels of control. This is illustrated schematically in figure 5.

### 5.2. *Noninteractive feedforward control*

While it is possible to design a state *feedback* controller to cancel the dynamics and decouple the nonlinear plant, animals are able to maintain control even when deafferented (Taub and Berman 1968). Thus the question arises, is it possible to design a *feedforward* controller to cancel the dynamics and decouple the input/output relationships? The feedback controller in figure 5(a) requires on-line feedback of the state $\mathbf{X}(k)$ of the plant. However, as commonly employed in control engineering, it is possible to estimate the state of the plant from input/output measurements using an internal model of the forward dynamics (e.g., Kalman filter). As we have described (Neilson and Neilson 2004), such a state estimator operating open-loop (i.e., no feedback from the output of the plant) in combination with the controller of figure 5(a) corresponds exactly to a feedforward controller able to decouple and render the dynamics of the plant unobservable, as illustrated in figure 5(b) for the state feedback controller. Of course, the absence of feedback in the state estimator means the resulting feedforward controller is unable to compensate for uncertainties in the system such as noise and external disturbances. Nevertheless, the argument indicates that a feedforward controller able to cancel dynamics and decouple control is feasible. We argue (Neilson and Neilson 2004) that the relationship between $\mathbf{\Delta R}^*$ and $\mathbf{\Delta M}^*$ (figure 5(a)) for the equivalent feedforward controller provides a definition for what is meant by the term 'inverse dynamics' of a square multiple input/multiple output nonlinear dynamical system. Except for time delays, the feedforward controller cancels the nonlinear dynamics of the plant (equivalent to multiplication by an inverse) and creates an identity relationship between $\mathbf{\Delta R}^*$ and $\mathbf{\Delta R}$. We now show that a feedforward controller can be formed adaptively, without worrying about state estimation, by using slave copies of the inverse models $\mathbf{E}_m^{-1}$, $\mathbf{BM}_m^{-1}$, $\mathbf{MCS}_m^{-1}$.

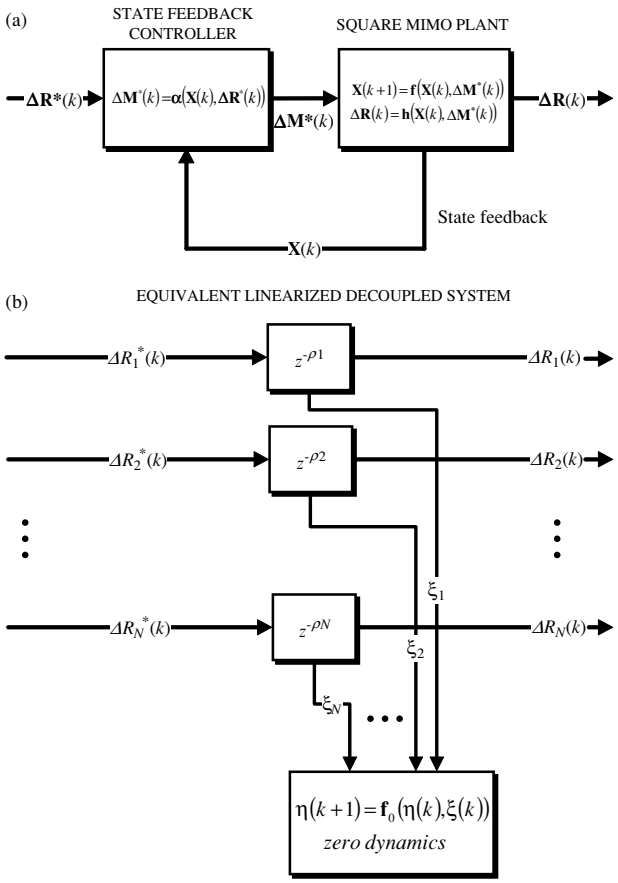



**Figure 5.** (a) Block diagram illustrating a unique state feedback controller for a square, multi-input/multi-output, interactive, nonlinear, dynamical plant. The input/output relationship for the plant can be taken to be the relationship between the orthogonal signals $\mathbf{\Delta M}^*(k)$ and $\mathbf{\Delta R}(k)$ in figure 4. The task for the controller is to transform a vector of orthogonal first-differenced required response signals $\mathbf{\Delta R}^*(k)$ into an appropriate vector of first-differenced motor commands $\mathbf{\Delta M}^*(k)$ to drive the plant and generate a response vector $\mathbf{\Delta R}(k)$ that matches the required response vector $\mathbf{\Delta R}^*(k)$ as accurately as possible. The state vector of the plant represented by $\mathbf{X}(k)$ is detected and fed back to the controller. (b) Block diagram illustrating the equivalent closed-loop relationships between the required response vector $\mathbf{\Delta R}^*(k)$ and the actual response vector $\mathbf{\Delta R}(k)$ created by the unique state feedback controller shown in (a). The controller renders the closed-loop system non-interactive, equivalent to a set of N independent parallel channels of control. The controller renders all the dynamics, except for time delays, unobservable in the closed-loop input/output relationships. The symbols $z^{-\rho i}$ represent time delays of $\rho_i$ sample intervals in the $i$th parallel control channel; that is, a train of $\rho_i$ delayers $z^{-1}$ in each channel. The symbol $\boldsymbol{\xi_i}$ represents the state vectors of the $\rho_i$ delayers $z^{-1}$ in the $i$th channel. The state vectors $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_N$ are observable in the closed-loop input/output relationships. Unobserved state variables are represented by the symbol $\boldsymbol{\eta}(k)$ and the unobserved dynamics are represented by the zero dynamics block $\boldsymbol{\eta}(k+1) = \mathbf{f}_0(\boldsymbol{\eta}(k), \boldsymbol{\xi}(k))$ at the bottom of the diagram.

### 5.3. *Adaptive formation of a noninteractive feedforward controller*

As outlined in section 4.4, modeling circuitry within sensory systems forms internal models of the forward and inverse

relationships between $\mathbf{\Delta M}$ and $\mathbf{\Delta T}$, $\mathbf{\Delta T}$ and $\mathbf{\Delta Q}$, $\mathbf{\Delta Q}$ and $\mathbf{\Delta R}$. A cascade connection of copies of the corresponding inverse models $\mathbf{MCS}_m^{-1}, \mathbf{BM}_m^{-1}, \mathbf{E}_m^{-1}$, provides exactly the neural circuitry required to transform preplanned vectors of desired response $\mathbf{\Delta R}^*$ into required movements and co-contractions $\mathbf{\Delta Q}^*$, then into required muscle tensions $\mathbf{\Delta T}^*$ to produce those movements and co-contractions, and finally into required motor commands $\mathbf{\Delta M}^*$ to generate the tensions. The transformations all involve low-dimensional orthogonal first-differenced feature signals. The inverse models $\mathbf{E}_m^{-1}, \mathbf{BM}_m^{-1}, \mathbf{MCS}_m^{-1}$ connect in cascade to form exactly the required feedforward controller described above. However, in section 4.4 we provided no detail about the internal structure of the MODELING CIRCUITRY block of figure 4. In this section (and in section 7) we address that internal structure. Here we describe the structure for forward and inverse modeling of each multiple input/multiple output system (**MCS**, **BM**, **E**), then in section 7 we describe the circuitry of the individual adaptive filters that make up each multiple input/multiple output model.

While **MCS**, **BM** and **E** systems are modeled independently and in parallel, the same modeling structure applies to each and so, for simplicity, we describe the structure only once. To maintain continuity of the symbols for input and output signals across the various figures we have chosen to describe in figure 6 modeling of the overall plant corresponding to the relationship between $\mathbf{\Delta M}^*$ and $\mathbf{\Delta R}$ in figure 4. The plant is comprised of **MCS**, **BM** and **E** connected in cascade but it should be kept in mind that the modeling described in figure 6 for the total plant is in fact done three times in parallel for each of the subsystems **MCS**, **BM**, **E**.

The input $\mathbf{\Delta M}^*(k)$ and the output $\mathbf{\Delta R}(k)$ of the plant in figure 6 correspond to orthogonal first-differenced feature signals and, as mentioned above, the plant is taken to be a square, interactive, multiple input/multiple output, nonlinear, dynamical system corresponding to the overall input/output relationship between $\mathbf{\Delta M}^*$ and $\mathbf{\Delta R}$ in figure 4. Details of the modeling process based on the adaptive LMS algorithm indicated in figure 6 are given in section 7. A copy of the vector of orthogonal input signals $\mathbf{\Delta M}^*(k)$ connects to the forward model and to the LMS-based modeling process. Outputs $\mathbf{\Delta R}(k)$ are compared with the outputs from the forward model to create a vector of error signals $\mathbf{e}_1(k)$. Error signals $\mathbf{e}_1(k)$ and input signals $\mathbf{\Delta M}^*(k)$ drive the adaptive LMS algorithm and tune the parameters of the forward model to minimize the mean square values of the error signals $\mathbf{e}_1(k)$. Nonlinear interactions with posture $\boldsymbol{\theta}$ and orientation in the gravitational field $\mathbf{g}$ are included in the internal model, but not shown in the figure. The forward model enables the afferent signals $\mathbf{\Delta R}(k)$ to be separated into reafferent components (model outputs) and exafferent components (error signals). When the forward model is accurate, the error signals $\mathbf{e}_1(k)$ provide an estimate of the first-differenced exafferent (i.e., disturbance) signals $\mathbf{\Delta D}(k)$ at the output of the plant. We show in section 6 that such estimates of exafference can be used to provide feedback compensation.

Ability to separate afference into reafference and exafference makes it possible to form accurate models of

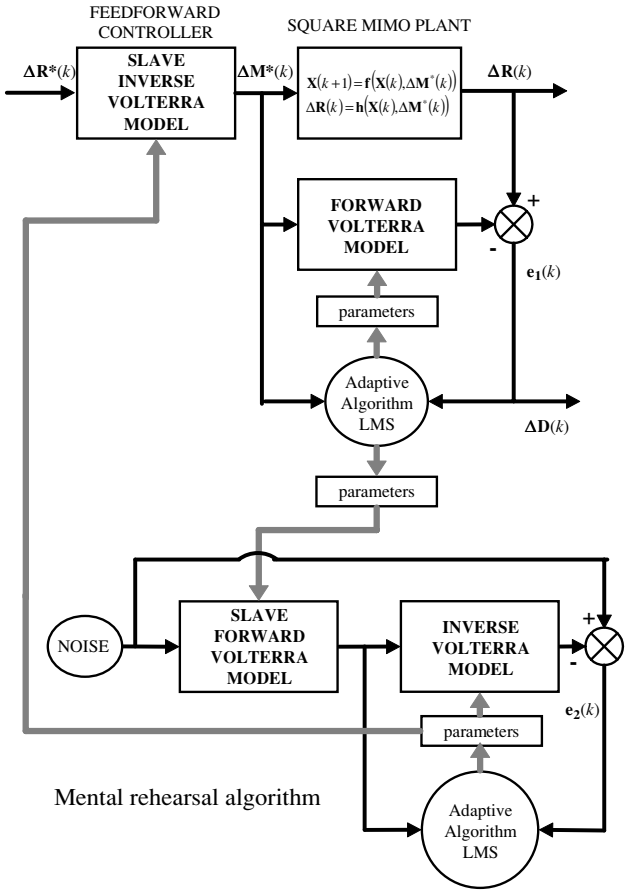



**Figure 6.** Schematic diagram illustrating adaptive formation of both forward and inverse models of the square multiple input/multiple output, interactive, nonlinear, dynamical plant in figure 5. Bottom part of diagram illustrates mental rehearsal algorithm used to form an inverse model of a slave copy of the forward model. A slave copy of the inverse model forms the adaptive feedforward controller. While the figure illustrates modeling of the overall plant defined by the relationship between $\mathbf{\Delta M}^*(k)$ and $\mathbf{\Delta R}(k)$ in figure 5, it is important to appreciate that the modeling procedure works for any arbitrary square, nonlinear, dynamical plant. The modeling procedure is applied simultaneously and in parallel to form forward and inverse models of the **MCS**, **BM** and **E** systems. The controller is comprised of the inverse models of these three subsystems connected in cascade. For simplicity the three subsystems have been combined into a single system defined by the overall relationship between $\mathbf{\Delta M}^*(k)$ and $\mathbf{\Delta R}(k)$ and its inverse. $\mathbf{\Delta M}^*(k)$ = vector of required orthogonal first-differenced motor commands, $\mathbf{\Delta M}(k)$ = vector of orthogonal first-differenced motor commands, $\mathbf{\Delta R}(k)$ = vector of orthogonal first-differenced response signals, $\mathbf{\Delta R}^*(k)$ = vector of centrally-generated required orthogonal first-differenced response signals, $\mathbf{\Delta D}(k)$ = estimate of vector of orthogonal first-differenced disturbance signals.

inverse dynamics even in the presence of a large amount of disturbance variation $\mathbf{\Delta D}(k)$ added to the output signals. One way an inverse model can be formed is illustrated in figure 6. We refer to this as the *mental rehearsal algorithm* because the forward model can 'teach' the inverse model independently of actually performing the task, for example, while sitting in an armchair thinking about performing the task. A vector of noise signals is generated and connected to the inputs of a 'slave' copy of the forward model as in

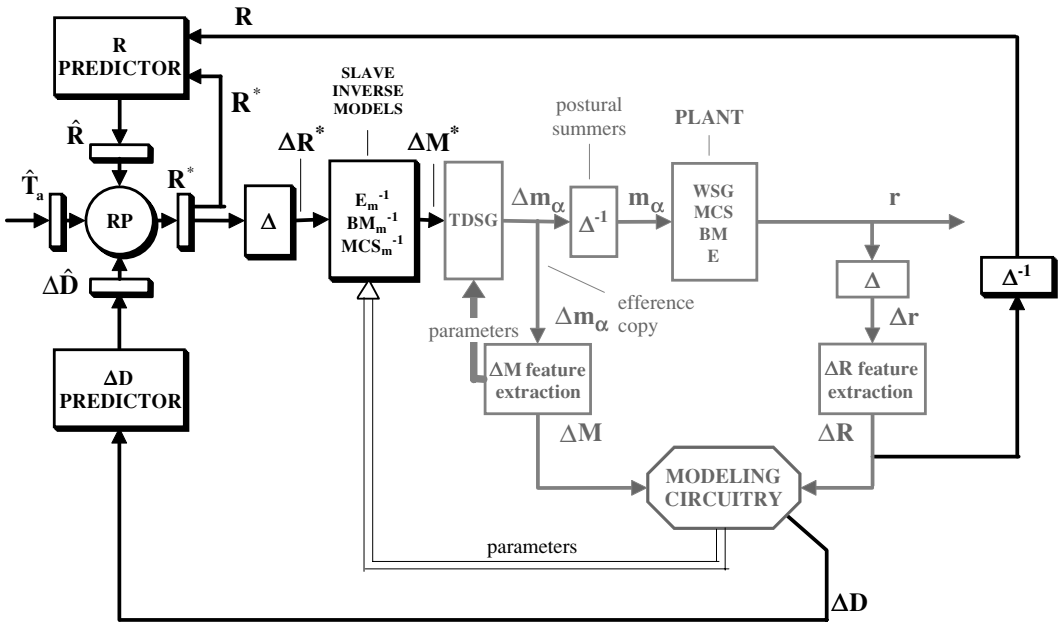



**Figure 7.** Schematic diagram of the adaptive feedforward/feedback control system proposed in AMT: The gray section of the diagram corresponds to figure 4 with the wired-in synergy generator **WSG**, muscle control system **MCS**, biomechanical system **BM** and external system **E** collapsed into the single block PLANT. In AMT these subsystems and their inverses are modeled separately and in parallel but for simplicity we illustrate only modeling and control of the combined plant. An adaptive feedforward controller consisting of SLAVE INVERSE MODELS and a task-dependent synergy generator **TDSG** control the plant. The input to this controller is the vector of pre-planned required orthogonal first-differenced response trajectories $\mathbf{\Delta R}^*$. The response planning system is shown at the left of the diagram. The input signals for response planning are the predicted future values of the orthogonalized target feature signals $\hat{\mathbf{T}}_\mathbf{a}$, predicted orthogonalized response feature signals $\hat{\mathbf{R}}$, and predicted orthogonalized first-differenced disturbance signals $\mathbf{\Delta}\hat{\mathbf{D}}$. The latter two signals are provided by feedback of the actual orthogonal response **R** and disturbance changes $\mathbf{\Delta D}$. (The vector of disturbance signals $\mathbf{\Delta D}$ is estimated during the modeling process.) The response planner **RP** requires a fixed interval of time (100–200 ms) to accomplish three tasks: (a) read the predicted future values of sensory feature signals $\hat{\mathbf{T}}_\mathbf{a}$, $\hat{\mathbf{R}}$, and $\mathbf{\Delta}\hat{\mathbf{D}}$ from working memory buffers, indicated by small rectangular blocks, (b) generate a minimum acceleration required response trajectory $\mathbf{R}^*$ to move the response **R** from its current state into alignment with the predicted target state $\hat{\mathbf{T}}_\mathbf{a}$, and (c) write $\mathbf{R}^*$ into a working memory buffer ready for execution. While $\mathbf{R}^*$ is read from working memory and executed in real-time, the **RP** system iterates its operation. **RP** operates intermittently at reaction time rates creating a concatenated sequence of error-correction submovements. As in figure 4 adaptive filters in the block labeled MODELING CIRCUITRY form both forward and inverse adaptive filter models of the multi-input/multi-output nonlinear dynamical relationships between $\mathbf{\Delta M}$ and $\mathbf{\Delta T}$, $\mathbf{\Delta T}$ and $\mathbf{\Delta Q}$, and $\mathbf{\Delta Q}$ and $\mathbf{\Delta R}$. For simplicity only the overall relationship between $\mathbf{\Delta M}$ and $\mathbf{\Delta R}$ is illustrated. Copies of the adaptive filter parameters derived from the modeling circuitry modulate the SLAVE INVERSE MODELS to give the required transformation $\mathbf{\Delta R}^*$ to $\mathbf{\Delta M}^*$. The **TDSG** (adaptively tuned as in figure 4) generates required changes $\mathbf{\Delta m}_\alpha$ to increment or decrement the motor commands $\mathbf{m}_\alpha$ held on-line by postural summers $\mathbf{\Delta}^{-1}$. This gives a new 220-vector of interrelated central motor commands $\mathbf{m}_\alpha$ to drive the plant.

figure 6. For modeling nonlinear dynamical systems it is important that the spectral density functions (PSDF) and the amplitude probability density functions APDF) of the noise signals should match those of the actual input signals $\mathbf{\Delta M}^*(k)$. For on-line modeling the noise generator can be replaced by the actual inputs $\mathbf{\Delta M}^*(k)$. Outputs from the 'slave' forward model connect to the inputs of an adaptive inverse model. Outputs from the adaptive inverse model are compared with the noise signals to create error signals $\mathbf{e}_2(k)$ and an adaptive algorithm adjusts the parameters of the inverse model to minimize the mean square value of the errors $\mathbf{e}_2(k)$. The inverse model is thus tuned to provide the inverse of the input/output transformation of the forward model. As shown in figure 6, a 'slave' copy of this inverse is used as the inverse model in the feedforward controller. Keep in mind however that for an actual simulation of an adaptive feedforward movement controller this modeling structure is implemented three times in parallel for each of the **MCS**, **BM** and **E** systems and the controller comprises a cascade connection of inverse models $\mathbf{E}_\mathrm{m}^{-1}$, $\mathbf{BM}_\mathrm{m}^{-1}$ and $\mathbf{MCS}_\mathrm{m}^{-1}$.

## 6. Bringing the solutions together

### *6.1. The AMT movement control system*

The computational solutions to the problems of redundancy, resource and nonlinear interactions provide us with the core of a complete feedback/feedforward movement control system. Figure 7 provides a simplified illustration of the computational structure of this system. The feedforward controller consists of the slave inverse model and the task-dependent synergy generator **TDSG**. The slave inverse comprises inverse models $\mathbf{E}_\mathrm{m}^{-1}$, $\mathbf{BM}_\mathrm{m}^{-1}$ and $\mathbf{MCS}_\mathrm{m}^{-1}$ in cascade and transforms the required response changes $\mathbf{\Delta R}^*$ into the appropriate required motor changes $\mathbf{\Delta M}^*$. The intervening change signals $\mathbf{\Delta T}^*$ and $\mathbf{\Delta Q}^*$ are omitted for simplicity. The

gray section of the diagram concerns the transformation of $\Delta\mathbf{M}^*$ to $\Delta\mathbf{R}$, the actual response change signals. This part of the figure is exactly the square MIMO system of figure 4 but redrawn with the wired-in synergy generator **WSG** and subsystems **MCS**, **BM** and **E** collapsed into the single block labeled PLANT to simplify the extended diagram. The feedback of signals $\Delta\mathbf{T}$ and $\Delta\mathbf{Q}$ in figure 4 is still inherently there, but these and their related modeling is now included within the plant block. The only feedback shown is that from the resultant external response signals **R**. In other words, the diagram shows only the main perceptual-motor feedback loop.

Taking them in reverse order, the solutions to the problems of redundancy, resources, and nonlinear interactions are incorporated in figure 7 as follows:

(1) Required response changes $\Delta\mathbf{R}^*$ are input to a feedforward controller that not only cancels all the complex nonlinear dynamics of the plant, but also decouples control so that each of the $N$ orthogonal feature signals comprising $\Delta\mathbf{R}^*$ and $\Delta\mathbf{R}$ is controlled independently, in parallel. In other words, the troublesome interactive dynamics of the plant disappear and the control problem becomes simply that of an aggregate of $N$ single input/single output channels equal in number to the control DFs of the task (section 5).

(2) Sensory processing networks ($\Delta\mathbf{M}$- and $\Delta\mathbf{R}$-feature extraction) tune themselves adaptively to remove redundancy from efference copy and sensory signals ($\Delta\mathbf{m}_\alpha$ and $\Delta\mathbf{r}$) by extracting orthogonal feature signals ($\Delta\mathbf{M}$ and $\Delta\mathbf{R}$). Parameters from the $\Delta\mathbf{M}$ feature extraction network simultaneously tune the task-dependent synergy generator **TDSG**. For a task with $N$ control DFs the orthogonal signals $\Delta\mathbf{M}$ and $\Delta\mathbf{R}$ and their required equivalents $\Delta\mathbf{M}^*$ and $\Delta\mathbf{R}^*$ are all $N$-dimensional. This allows planning and control to be implemented economically in the reduced space of the task rather than with the full dimensionality of the plant, thus minimizing computational demand on limited central processing resources (section 4).

(3) The **TDSG** appropriately transforms the $N$ orthogonal motor command changes $\Delta\mathbf{M}^*$ required for an $N$ DF task into 220 interrelated command changes $\Delta\mathbf{m}_\alpha$ that are added to unchanged postural commands and passed to the **WSG**. This two-step process of synergy generation ensures that synergies are implemented with minimum energy, providing a unique solution to the redundancy problem (section 3).

The solutions thus come neatly together in figure 7, but there are still some loose ends. We alluded above that orthogonalization of signals provides economies in planning as well as in control, but we have not yet considered the process of planning a required response $\mathbf{R}^*$ or equivalently, the change signal $\Delta\mathbf{R}^*$ that serves as input to the controller. Nor have we closed the perceptual-motor loop that provides the feedback aspect of the controller and we have also left in mid-air the vector of disturbance signals $\Delta\mathbf{D}$ corresponding to exafferent signals derived during the modeling process (figure 6). These aspects of the movement control system are shown in the dark part of figure 7 and without them we cannot construct the complete computational process of generating a response **R** to a given target signal $\mathbf{T_a}$.

### *6.2. Response planning and the perceptual-motor loop*

Unlike the continuous process of sensory analysis (i.e., extracting orthogonal sensory feature signals and modeling their interrelationships) and the continuous process of response execution (i.e., generating synergetic motor commands from planned required responses and transforming these through the inverse model of the plant), the process of response planning proposed in AMT is intermittent (Neilson *et al* 1992). The response planning module **RP** (figure 10 mid left) implements a process of trajectory generation that occurs during a finite planning interval. During this time the processor is fully occupied and no further trajectory can be generated. This introduces both central time delay and intermittency into the perceptual-motor loop. The **RP** module receives information about target signals $\mathbf{T_a}$, current response signals **R** and first-differenced disturbance signals $\Delta\mathbf{D}$. Target signals $\mathbf{T_a}$ can come directly from the environment as in visual tracking tasks but in self-paced tasks they have to come from higher (cognitive) levels of response planning not addressed here. Sensory signals are necessarily affected by peripheral and/or central time delay, and possibly by delays in the external system, so we postulate that the CNS is able to generate predictions of their future values. Predicted signals $\hat{\mathbf{T}}_\mathbf{a}$, $\hat{\mathbf{R}}$ and $\Delta\hat{\mathbf{D}}$ are therefore shown in the figure as direct inputs to **RP**, along with an **R** predictor and $\Delta\mathbf{D}$ predictor (the $\mathbf{T_a}$ predictor is off stage).

How might these predictions be achieved? We know from adaptive filter theory (Box and Jenkins 1976, Goodwin and Sin 1984, Widrow and Stearns 1985, Haykin 1986) that the filter networks for forming internal models of relationships between signals are also precisely what is required to generate predictions of future values of stochastic signals. Thus the neural adaptive filters of AMT (section 7) can be used for predicting as well as for modeling (Neilson *et al* 1995). Providing the stochastic properties of a signal do not change too rapidly, an adaptive filter will tune continuously on-line to provide an optimal prediction of future values of the signal. Of course, the further ahead in time the prediction, the greater the prediction error. Also, the closer the characteristics of the stochastic signal to pure randomness (i.e., the broader the bandwidth), the greater the prediction error (Neilson *et al* 1988b).

According to AMT the **RP** module operates in a fixed interval of time to accomplish three tasks: (a) read in from working memory the predicted future values of sensory feature signals, $\hat{\mathbf{T}}_\mathbf{a}$, $\hat{\mathbf{R}}$ and $\Delta\hat{\mathbf{D}}$, (b) generate a minimum acceleration required-response trajectory $\mathbf{R}^*$ to move the response **R** from its current state into alignment with the predicted target state $\hat{\mathbf{T}}_\mathbf{a}$, and (c) write $\mathbf{R}^*$ into working memory ready for execution. While the current required response $\mathbf{R}^*$ is being executed in real-time, the next is being generated by the **RP** system in accelerated time. This can occur because the **RP** trajectory generator uses parallel array processing so that within the fixed

planning interval it can generate a vector of required response trajectories that extend far longer into the future than the next planning time. The computational structure of this generator involves adaptive filters, as for modeling and prediction, and is given in Neilson *et al* (1995). Accelerated response planning has the advantage that execution or disturbance errors in an ongoing trajectory can be corrected by a new response long before the incorrect response concludes. This gives rise to a control strategy known as *receding horizon predictive control*. In our simulations of **RP** it is possible to vary the horizon, so the strategy applied in AMT might better be called *variable horizon predictive control*. Increasing the prediction horizon increases the duration of the S-shaped minimum acceleration trajectories and effectively reduces the speed of the required response $\mathbf{R}^*$. As a result the gain, phase and bandwidth of the perceptual-motor loop vary with the prediction horizon, relating to observable characteristics of accuracy–energy or speed–accuracy behavior (Neilson *et al* 1995).

In summary, dedicated **RP** circuitry allows a small number $N$ of required response trajectories comprising the vector $\mathbf{R}^*$ to be generated simultaneously without interference. Dimension $N$ of the $\mathbf{R}^*$ vector equals the number of control DFs for the task. Thus planning as well as control computations are carried out using only low-dimensional vectors of orthogonal feature signals reducing the central processing load. Nevertheless, an inherent central processing limitation remains, associated with intermittency and refractoriness. The adaptive inverse-model feedforward controller in figure 7 is embedded within an intermittently operating, perceptual-motor feedback loop that is subject to 'reaction time' delay mitigated by prediction. We discuss these time constraints when we examine behavioral evidence on planning and control of movement in section 8.

## 7. Do the solutions work?

In the preceding sections we have built up a cohesive scheme for feedforward/feedback control of movement that simultaneously addresses three major problems for any movement control theory, computational or otherwise. Figure 7 gives the design for a simulable system for the planning and control of movement in terms of an $N$-vector of responses in an $N$-dimensional task space. We now consider how such a simulation might be implemented in a biologically realistic way. It is not enough to propose a mathematical description that accounts for what we observe in movement control if there is no viable neural basis to underpin the proposal. In this section we briefly review the basic neural filters that were the genesis of AMT (Neilson *et al* 1985) in order to extend this concept to our newly developed nonlinear filter network that appears to be equally justifiable in terms of the structure and resources of the CNS. We then document the performance of this network in simulations that test its robustness to extremes that may well exceed those presented by realistic neural signals. The adaptive filter processes set out and tested below serve as 'plug-ins' to the movement control solutions presented above. If these processes successfully model nonlinear dynamical relationships between signals that are not necessarily Gaussian, the solutions should work.

### *7.1. The basic neural adaptive filter of AMT*

The fundamental neural adaptive processing of AMT arose in terms of cortico-cerebellar-cortical loops. The neuroanatomical connections of these loops are well known (Barlow 2002, Ito 1984) and the connectivity is illustrated schematically in figure 8(a). Except for the shaded blocks at the top of the diagram, all the structure and connectivity shown in figure 8(a) is available in textbook descriptions of the cerebellum (e.g., Ghez and Thach (2000)). The shaded blocks represent hypothetical components, one a comparator and the other a multiplier plus working memory. This working memory communicates with long-term memory (LTM). If present within the nervous system, these transform the circuit into a neural adaptive filter module that implements the least mean square (LMS) adaptive filter algorithm (Widrow 1970, Widrow and Stearns 1985). The connections are shown again in a more circuit-like diagram in figure 8(b). Via descending modulation of the inferior olive and climbing fibers, the circuit has the ability to tune the synaptic gains of the input activity of a small cluster of Purkinje cells to their parallel fibers. The existence of this circuit is a fundamental hypothesis of AMT. If true, it follows from known cerebellar structure that this module is repeated hundreds of thousands of times, each module operating more or less independently in parallel with the others. The combined action of these many modules is precisely what is required to establish an internal model between sensory and motor signals. We propose that this is accomplished as follows.

The shaded block at top right of figure 8(a) represents a cortical column comparator. This comparator compares a sensory feedback signal from the periphery, in this case the feedback of a muscle tension signal $t$ from Golgi tendon organs in a functional muscle, with feedback via the thalamus of a signal $t_m$ from the dentate nucleus in the cerebellum. The resulting error signal $t - t_m$ is communicated to another cortical column represented by the shaded block at top left in figure 8(a). This cortical column contains a multiplier, shown schematically in the center of figure 8(b). It multiplies the error signal with another selected signal, in this case the motor command $m_\alpha$. These signals can be seen entering the multiplier from right and left respectively in both parts of the figure. A small fraction $\mu$ of the product increments or decrements the level of neural activity held on-line in a working memory module. In figure 8(a) this module is included with its multiplier input in the shaded cortical column at top left. The neural activity held on-line in the working memory module then modulates the gains of Purkinje cells in cerebellar cortex via the inferior olive and climbing fibers. This modulation is indicated as a multiplier in figure 8(b). As shown in figure 8(a) the signal $m_\alpha$ also projects to the same Purkinje cells via corticobulbar pathways, pons, mossy fibers, granule cells and parallel fibers. The Purkinje cells influence the activity of a cluster of cells in the dentate nucleus that, in turn, closes the loop by projecting back to the cerebral cortex via the thalamus.

The simple circuit of figure 8 is exactly what is required to implement Widrow's adaptive filter LMS algorithm. The behavior of the circuit is wondrous. It updates at each

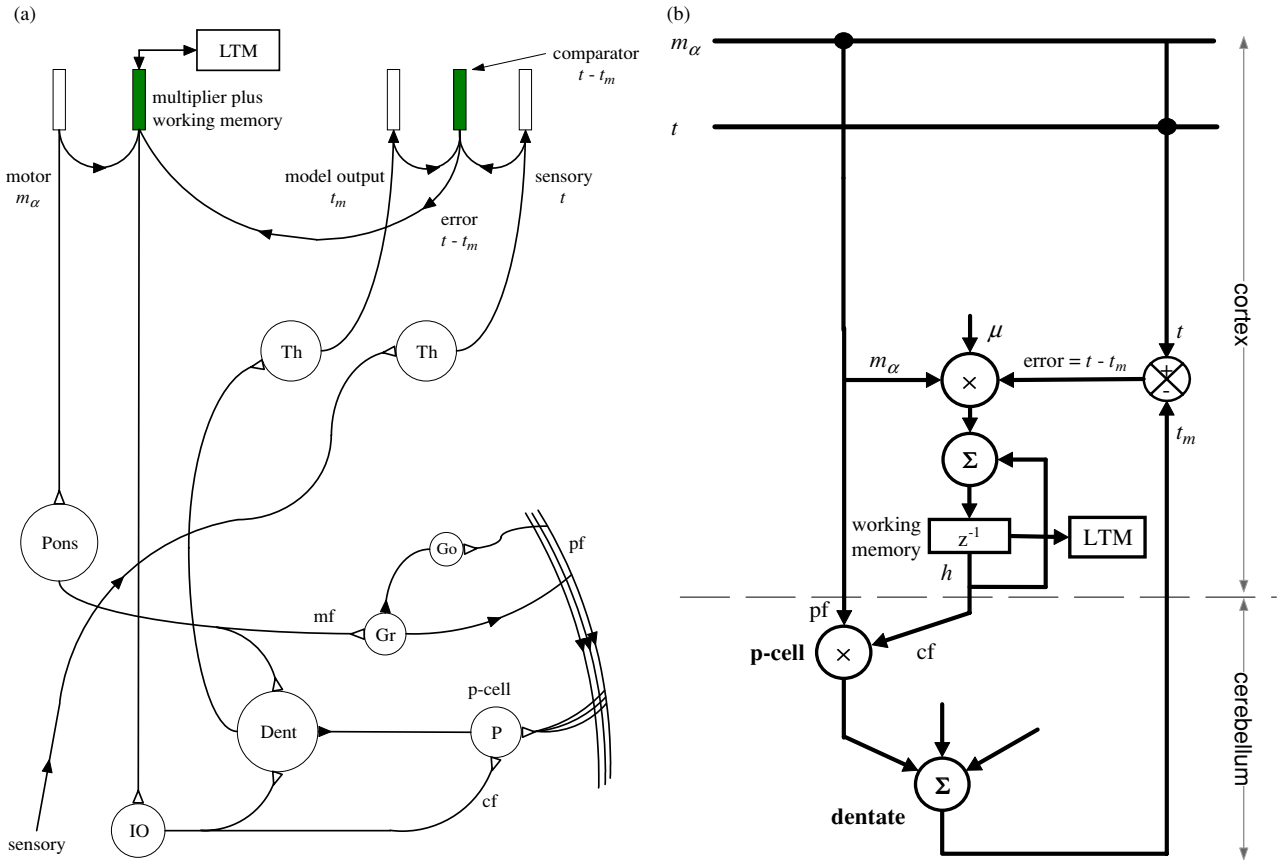



**Figure 8.** (a) Schematic diagram of known neuroanatomical connections in cortico-cerebellar-cortical pathways. The shaded blocks are hypothetical and represent a multiplier plus working memory and a comparator. Working memory communicates with long-term memory LTM. With the hypothetical components added, the neuroanatomical connections resemble the circuitry of a least mean squares LMS adaptive filter module shown in (b). $m_\alpha$ = motor command signal, $t$ = fed back muscle tension signal, $t_m$ = feedback of tension signal from cerebellum model, Pons = pontine nuclei, Gr = granule cells, Go = Golgi cells, P-cells = Purkinje cells, Dent = dentate nucleus, Th = thalamic nuclei, IO = inferior olive, mf = mossy fibers, cf = climbing fibers, pf = parallel fibers. (b) Circuit for an adaptive filter least mean square LMS module that tunes the adaptive parameter $h$ to model the linear relationship $t = h_t\, m_\alpha$ between the input signal $m_\alpha$ and the output signal $t$ where $h_t$ is the true value of the parameter. The same notation used in (a) applies in (b). The adaptive parameter is represented by the signal $h$ at the output of the working memory. Working memory is shown as a sample-and-hold block labeled $z^{-1}$ with a feedback loop around it forming a discrete-time summer. The parameter $h$ is held on-line by the feedback loop and is updated at each sample interval by the signal $\mu m_\alpha(t - t_m)$ at the output of the multiplier. The symbol $\mu$ represents a small constant (typically <0.01) known as the adaptive gain of the algorithm. It controls the speed of convergence. It is assumed that the parameter $h$ can be stored into and retrieved from long-term memory (LTM). The parameter signal $h$ and the input signal $m_\alpha$ connect to the multiplier labeled p-cell where they are multiplied together. The output $t_m = h\, m_\alpha$ from the multiplier is compared with the tension signal $t = h_t\, m_\alpha$ to give the error signal $t - t_m$. The LMS module tunes the parameter $h$ to reduce the variance of the error signal to a minimum so that $h$ converges to $h_t$. The flow of signals in the LMS module exactly matches the flow of signals in the cortico-cerebellar-cortical pathways shown in (a).

sample using the current error value squared to estimate the error variance but this 'noisy' estimate is smoothed by the low-pass filter characteristics of the loop. The amount of smoothing and the speed of convergence are controlled by the setting of the adaptive gain $\mu$. In terms of the cerebellum the signal from working memory is precisely the adaptive parameter required to adjust the gains of the Purkinje cells to reduce the mean square value of the error signal $t - t_m$ to a minimum. In other words, the circuit adjusts the adaptive filter parameter to form the best possible estimate of the regression coefficient between the signals $m_\alpha$ and $t$.

How realistic is this proposal? The existence of a neural comparator does not seem problematic because the various synaptic inputs to a neuron have an additive or subtractive influence on membrane potential depending on whether the synapses are excitatory or inhibitory. However, many engineers are aware of the complex circuitry required to implement a multiplier in floating-point hardware for a digital computer. It is not surprising therefore that they might query how a multiplier can be implemented within the nervous system. It may not be that difficult. It is now well established that many second messenger G-protein synapses have a multiplicative action on the membrane potential responses to inputs from other synapses. Depending on the synapse, these modulatory influences can last from a few milliseconds to days or even longer (Siegelbaum *et al* 2000).

### 7.2. Neural internal models

The adaptive neural circuit described above has the ability to establish one parameter in an adaptive filter model of a relationship between two signals. If the relationship is non-dynamic and linear, this circuit is all that is required. If the relationship is non-dynamic and $n$th order polynomial, $n$ modules working independently in parallel are needed to determine the model. Likewise a linear dynamic relationship requires $m$ modules to establish a model using $m$ lags, and so on. The method is general and multiple repetitions of this module working independently in parallel give the means to establish a complete set of model parameters, even for nonlinear dynamical relationships.

The signal $\mu$ in the LMS algorithm controls the speed of convergence of the adaptive parameters. Typically it is set to a small constant value less than 0.01. Such a constant value can easily be wired-into the neural circuit as a synaptic gain controlling the proportion of the product of the input and error signals that increments or decrements the adaptive parameter held on-line in working memory. On the other hand the value of $\mu$ might be varied by the nervous system. If $\mu$ is set high the adaptive parameter converges quickly but it is noisy and inaccurate. If $\mu$ is set low the parameter converges slowly but it is less noisy and more accurate. While yet to be explored, it is possible that the nervous system might take advantage of this. When a person is unskilled at a task it could increase the value of $\mu$ so the adaptive parameters converge quickly towards their true values. As the internal model improves with practice the value of $\mu$ could be decreased thereby decreasing the noise in the parameter estimates, improving model accuracy and protecting the parameters against perturbations.

How these parameters might be maintained in the CNS gives rise to another key aspect of AMT. We propose that adaptive filter parameters held on-line in working memory can be transferred into and retrieved from long-term memory, forming the basis for a system of procedural/motor memory. We see the storage, retrieval and continuing adaptation of these parameters as enabling the CNS to maintain a repertoire of internal models that can be selected and switched according to the task. This is discussed further in section 8. Clearly the extent of that repertoire will depend both on the extent of available long-term memory and on the number of parameters necessary to define each internal model. We now turn to this issue.

When interrelationships between sensory and motor signals are approximately linear, the number of parameters required to define the models between them is modest. The known extent of cortico-cerebellar-cortical modularity and cortical motor memory would seem to be sufficient to establish and maintain a reasonable range of human performance. The stumbling block comes when sensory–motor relationships are nonlinear, and clearly they frequently are. Whether or not the CNS can realistically accommodate the parameter requirements for a large number of nonlinear internal models depends on how many parameters are needed to define a nonlinear relationship. This is addressed in the following section.

### 7.3. Structure of a network for nonlinear dynamical modeling

Theoretically speaking, an adaptive filter model of a Volterra series description of the input/output relationship of a nonlinear dynamical system (Marmarelis 2004, Westwick and Kearney 2003) can be constructed using a parallel repetition of the adaptive filter module illustrated in figure 8. However, there are a number of practical problems. Firstly, it is well known that if a signal with a Gaussian amplitude probability density function APDF) is transformed through a nonlinear dynamical system, the APDF of the output signal will be non-Gaussian (Bendat 1990, 1998). For a Volterra model to emulate changes in the skewness (asymmetry) and kurtosis (peakiness) of the output signal APDF, it needs to be of third order at the least. But to implement the linear, bilinear and trilinear components of that model for even a single input/single output system requires hundreds of thousands of adaptive filter parameters. Thus practically, we need to find an adaptive filter structure that can implement a third-order Volterra model using only a small number of adaptive filter parameters.

Secondly, if one considers the $\mathbf{\Delta m}_\alpha$, $\mathbf{\Delta t}$, $\mathbf{\Delta q}$ and $\mathbf{\Delta r}$ signals from typical motor tasks, it becomes apparent that the need to form adaptive filter models of nonlinear dynamical relationships between signals with non-Gaussian APDFs and non-white power spectral density functions (PSDFs) cannot be avoided. Furthermore, the APDFs and PSDFs of these signals will change from task to task and cannot be predicted in advance. As a result, the CNS must be able to form adaptive filter models independently of the APDFs and PSDFs of the input signals. This is a major threat to theories proposing adaptive formation of internal models for motor control. As yet, no theoretical solution has been proposed, although a number of researchers are currently working on the problem (see Marmarelis (2004), Westwick and Kearney (2003)).

Building on the ideas of these authors, we have developed an adaptive filter (figure 9) that requires only 60 adaptive parameters to form a third-order Volterra model of the nonlinear forward or inverse relationship between input/output signals that can tune independently of APDFs and PSDFs of the input. The main features of the filter are as follows. The signal $u(k)$ applied to the input of the system to be modeled and the output signal $y(k)$ are sensed and connected to the adaptive filter. Both the APDF and PSDF of $u(k)$ can vary and the modeled system can have arbitrary nonlinear dynamical input/output characteristics. In the adaptive filter the input signal $u(k)$ passes through a parallel array of ten linear filters ($\mathrm{SVF}_1, \ldots, \mathrm{SVF}_{10}$). The orthonormal impulse response functions of these filters correspond to the first ten singular vectors from a singular value decomposition of an arbitrarily generated symmetrical $30 \times 30$ matrix of low-pass filtered (bandwidth <6.0 Hz) random numbers. This configuration is predicated on the fact that any matrix of low-pass filtered random numbers can be accurately reconstructed by an appropriately weighted combination of the first ten singular vectors (see Noble and Daniel (1988)). Providing the input and output signals of an arbitrarily generated third-order Volterra system have spectral bandwidths less than 6.0 Hz, it is

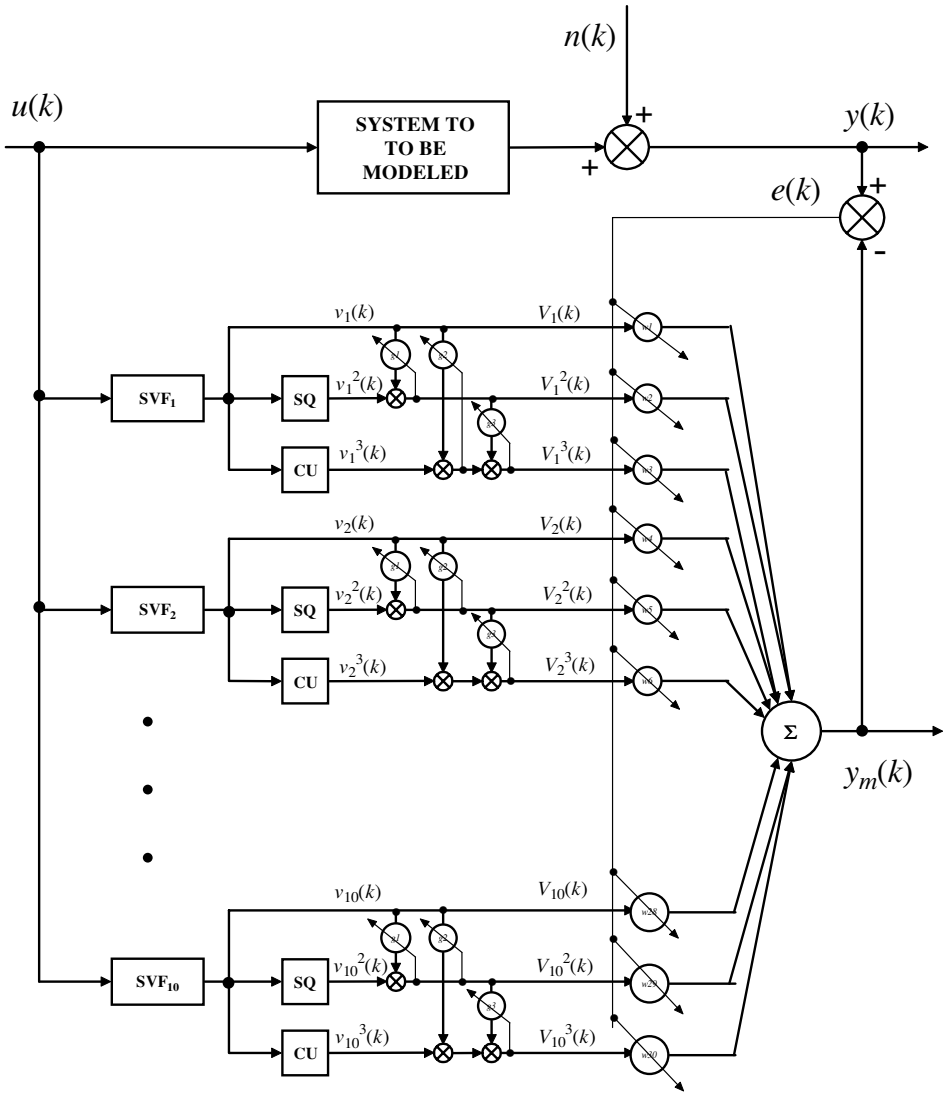



**Figure 9.** Schematic diagram of an adaptive filter able to model the nonlinear dynamical input/output relationship of a SYSTEM TO BE MODELED independently of the power spectral density function and/or the amplitude probability density function of the input signal $u(k)$. The output signal $y(k)$ from the system to be modeled includes additive noise $n(k)$. The input signal $u(k)$ is passed through 10 singular vector filters $\mathrm{SVF}_1, \ldots, \mathrm{SVF}_{10}$ to give 10 output signals $v_1(k), \ldots, v_{10}(k)$. Each of these signals is transformed by a squarer SQ and a cuber CU to give the signals ${v_1}^2(k), \ldots, {v_{10}}^2(k)$ and ${v_1}^3(k), \ldots, {v_{10}}^3(k)$, respectively. The three signals $v_i(k)$, ${v_i}^2(k)$, ${v_i}^3(k)$ at the output of the $i$th singular value filter $\mathrm{SVF}_i$, $1 \leqslant i \leqslant 10$, are transformed through a network of three LMS adaptive filter modules with adaptive parameters $g_{i1}$, $g_{i2}$, $g_{i3}$. Each network is arranged to perform an on-line Gram-Schmidt orthogonalization of its three input signals $v_i(k)$, ${v_i}^2(k)$, ${v_i}^3(k)$ to give three orthogonal output signals $V_i(k)$, ${V_i}^2(k)$, ${V_i}^3(k)$. A linear combination $y_m(k)$ of the 30 orthogonalized signals $V_i(k)$, ${V_i}^2(k)$, ${V_i}^3(k)$, $1 \leqslant i \leqslant 10$, is formed by an array of 30 LMS adaptive filter modules with adaptive parameters $w_1, \ldots, w_{30}$ to give the output signal $y_m(k)$ from the adaptive filter. The output $y_m(k)$ is compared with the output $y(k)$ from the system being modeled to produce an error signal $e(k)$. The error signal connects to all 30 of the $w$-parameter LMS modules. Each $w$-parameter tunes adaptively to reduce the error signal $e(k)$ to a minimum.

possible to reconstruct the first-order Volterra kernel from an appropriate combination of the impulse response functions of the filters $\mathrm{SVF}_1, \ldots, \mathrm{SVF}_{10}$. The output signal from each SVF is then squared and cubed to give output signals $v_i(k)$, ${v_i}^2(k)$, ${v_i}^3(k)$ $(i = 1, \ldots, 10)$. This is based on the demonstration by Bendat (1990, 1998) that a second-order Volterra kernel can be constructed by a cascade combination of a linear filter followed by a squarer and similarly, a third-order Volterra kernel can be constructed as a cascade combination of a linear filter followed by a cuber. By tuning parameters $w_1, \ldots, w_{30}$ (figure 9 right) to obtain an appropriate linear combination $y_m(k)$ of the 30 linear, squared and cubed signals at the outputs of the SVFs, the output $y(k)$ from the modeled system can be reconstructed.

The problem associated with changing the APDF and/or the PSDF of the input signal $u(k)$ can be attributed to the fact that the linear, square and cube signals, $v_i(k)$, ${v_i}^2(k)$, ${v_i}^3(k)$ at the outputs of the singular vector filters are strongly correlated with each other and the sizes of the regression coefficients

**Table 1.** Performance of the nonlinear adaptive filter of figure 9 according to the statistical characteristics of the input signal and of the system being modeled.

| Input | | | | | System | | | | | Performance | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| APDF | Bandwidth (Hz) | Scaling | Skew-ness | Kurtosis | Gain | Resonant freq (Hz) | Linear weight | Bilinear weight | Trilinear weight | %ve | %mse |
| nor | 1 | $u = 1.5s$ | −0.07 | 3.01 | 2 | 0.5 | 1 | 0 | 0 | 0.02 | 1.23 |
| nor | 1 | $u = 1.5s$ | −0.07 | 3.01 | 2 | 0.5 | 1 | 1 | 1 | 9.52 | 27.18 |
| exp | 1 | $u = 2s^2$ | 6.87 | 113.71 | 2 | 0.5 | 1 | 0 | 0 | 4.8 | 15.09 |
| exp | 1 | $u = 2s^2$ | 6.87 | 113.71 | 2 | 0.5 | 1 | 1 | 1 | 11.18 | 31.37 |
| bga.05 | 1 | $u = 6s$ | −0.08 | 7.85 | 2 | 0.5 | 1 | 0 | 0 | 0.02 | 1.43 |
| bga.05 | 1 | $u = 6s$ | −0.08 | 7.85 | 2 | 0.5 | 1 | 1 | 1 | 0.09 | 29.95 |
| nor | 1 | $u = 1.5s$ | −0.07 | 3.00 | 2 | 2.0 | 1 | 0 | 0 | 0.78 | 8.62 |
| nor | 1 | $u = 1.5s$ | −0.07 | 3.00 | 2 | 2.0 | 1 | 1 | 1 | 11.86 | 31.42 |
| exp | 1 | $u = 2s^2$ | 7.64 | 147.42 | 1 | 2.0 | 1 | 0 | 0 | 3.05 | 13.95 |
| exp | 1 | $u = 2s^2$ | 7.64 | 147.42 | 1 | 2.0 | 1 | 1 | 1 | 2.9 | 15.36 |
| bga.05 | 1 | $u = 3s$ | −0.03 | 6.79 | 2 | 2.0 | 1 | 0 | 0 | 4.7 | 21.34 |
| bga.05 | 1 | $u = 3s$ | −0.03 | 6.79 | 2 | 2.0 | 1 | 1 | 1 | 9.1 | 28.90 |
| nor | 4 | $u = s$ | 0.002 | 2.96 | 1 | 0.5 | 1 | 0 | 0 | 0.0 | 0.33 |
| nor | 4 | $u = s$ | 0.002 | 2.96 | 2 | 0.5 | 1 | 1 | 1 | 6.6 | 23.12 |
| exp | 4 | $u = 0.5s^2$ | 6.90 | 95.92 | 1 | 0.5 | 1 | 0 | 0 | 5.4 | 9.78 |
| exp | 4 | $u = 0.5s^2$ | 6.90 | 95.92 | 1 | 0.5 | 1 | 1 | 1 | 3.68 | 12.70 |
| bga.05 | 4 | $u = 3s$ | 0.23 | 16.29 | 1 | 0.5 | 1 | 0 | 0 | 0.18 | 4.24 |
| bga.05 | 4 | $u = 3s$ | 0.23 | 16.29 | 1 | 0.5 | 1 | 1 | 1 | 3.79 | 19.06 |
| nor | 4 | $u = s$ | 0.002 | 2.96 | 1 | 2.0 | 1 | 0 | 0 | 0.09 | 2.95 |
| nor | 4 | $u = s$ | 0.002 | 2.96 | 1 | 2.0 | 1 | 1 | 1 | 11.78 | 33.17 |
| exp | 4 | $u = 0.5s^2$ | 5.53 | 55.00 | 1 | 2.0 | 1 | 0 | 0 | 4.8 | 18.16 |
| exp | 4 | $u = 0.5s^2$ | 5.53 | 55.00 | 1 | 2.0 | 1 | 1 | 1 | 2.13 | 13.32 |
| bga.05 | 4 | $u = 3s$ | 0.04 | 15.83 | 1 | 2.0 | 1 | 0 | 0 | 1.5 | 12.36 |
| bga.05 | 4 | $u = 3s$ | 0.04 | 15.83 | 1 | 2.0 | 1 | 1 | 1 | 4.2 | 20.62 |

Notes: APDF = amplitude probability density function of random number input, $s$ = output from eighth-order Butterworth digital filter, $u$ = stochastic input to adaptive filter, %ve = variance of error as a percentage of variance of system output $y(k)$, %mse = mean square of error as a percentage of mean square of system output $y(k)$. The simulation runs for the shaded rows are illustrated in figure 9.

between them vary with changes in the APDF and/or PSDF of $u(k)$. These correlations interfere with the tuning of the 30 $w$-parameters. This problem is overcome by the set of adaptive filter parameters $g_1, \ldots, g_{30}$ (mid figure 9). The 30 $g$-parameters (three $g$-parameters for each of the 10 SVFs) perform an adaptive Gram–Schmidt orthogonalization of the $v_i(k)$, ${v_i}^2(k)$ and ${v_i}^3(k)$ signals at the output of each singular vector filter. Consequently, if the APDF or PSDF of the input signal changes, the $g$-parameters adaptively retune to maintain an orthogonal relationship between the signals $V_i(k)$, ${V_i}^2(k)$ and ${V_i}^3(k)$ derived from the output of each singular vector filter. It is then simple for the 30 $w$-parameters to tune adaptively to minimize the mean square value of the error signal $e(k)$.

### *7.4. Testing the nonlinear adaptive filter*

Using MATLAB 6.1/Simulink 5.0 we conducted a series of 24 simulation tests of the discrete-time adaptive filter in figure 9. Sample rate was set to 20 Hz. Performance of the adaptive filter for a variety of input signals and modeled systems is presented in table 1.

*7.4.1. Stochastic input signals u(k).* Stochastic input signals $u(k)$ were generated using command *rpiid.m* from the MATLAB Higher Order Spectral Analysis toolbox. This command generates a sequence of identically independently distributed random numbers as samples of a random variable drawn from one of five different pre-selected distributions. For each of the 24 simulation runs corresponding to the rows in table 1 we generated a time series of 12 000 random numbers using either (i) a normal Gaussian distribution (nor), (ii) an exponential distribution (exp) or (iii) a Bernoulli–Gaussian distribution with spike probability of 0.05 (bga.05). This is indicated in the column labeled APDF. The random number series was then transformed through a low-pass eighth-order Butterworth digital filter with bandwidth set to either 1.0 Hz or 4.0 Hz as shown in the column labeled bandwidth. For some simulation runs the stochastic signal $s(k)$ at the output of the Butterworth filter was squared to produce an extreme deviation of the APDF from Gaussianity. The signal $s(k)$ or $s^2(k)$ was multiplied by a scaling factor (0.5, 1, 1.5, 2, 3 or 6 as shown in the column labeled scaling) in order to keep the average amplitude of input $u(k)$ in the range 0–1. Despite this, for some highly skewed inputs the maximum peak amplitude reached as high as 10.8, 50 times the mean value of the signal, providing a very difficult test for the filter.

The skewness of the 12 000 samples in each stochastic input signal $u(k)$ was determined using the MATLAB command *skewness.m*. This computes the third central moment divided by the cube of the standard deviation. It does not adjust for bias. These measures are given in the table 1 column labeled skewness. Likewise the kurtosis of each input signal was determined using the command *kurtosis.m*. This

computes the fourth central moment divided by the fourth power of the standard deviation. It does not adjust for bias. These measures are given in the column labeled kurtosis. As seen in the table, kurtosis for a Gaussian distribution is close to 3.0 and increases as the distribution becomes more peaky (i.e. leptokurtic).

*7.4.2. System to be modeled.* For the system to be modeled we used a discrete-time second-order linear filter for which dc gain, damping ratio and natural resonant frequency could be preset prior to each simulation run. For the 24 simulation runs presented in table 1 the dc gain was set to either 1 or 2 as in the gain column, the damping ratio was set to 0.5, and the natural resonant frequency was set to either 0.5 Hz or 2.0 Hz as in resonant freq column. The linear output from the discrete-time second-order filter was both squared and cubed to give bilinear (squared) and trilinear (cubed) outputs as well as the original linear output. The output signal $y(k)$ was then constructed as a linear combination of the linear, bilinear and trilinear outputs. As specified in table 1 the linear, bilinear and trilinear weights were preset prior to each simulation run to either 1, 0, 0 respectively, giving a linear system, or to 1, 1, 1, giving a nonlinear third-order Volterra system with equal weights on each component of the output.

*7.4.3. Performance.* Table 1 gives information for 24 simulation runs with differing combinations of input signal and modeled system characteristics. Each run began by initializing to zero the 30 $g$-parameters in the orthogonalizing networks and the 30 $w$-parameters (figure 9). During each run a separate LMS adaptive filter module (the same as the cerebellum LMS module in figure 8(b)) tuned each parameter. These modules are shown in figure 9 by the diagonal arrows drawn across each parameter and connected to the appropriate error signal.

Preliminary testing showed that performance of the adaptive filter was very poor when $u(k)$ was either highly skewed or highly leptokurtic. Such stochastic input signals contain occasional very large amplitude spikes, 10 to 50 times larger than the average amplitude. These spikes perturbed the LMS modules causing the adaptive $g$- and $w$-parameters to jump in a step fashion away from convergence. In most simulation runs these jumps created an instability in the system from which the adaptive filter could not recover. The filter output $y_m(k)$ diverged to very large amplitude ($>10^{17}$ units) oscillations. To overcome this problem we installed sigmoid-shaped saturation filters into the LMS modules to limit the amplitudes of the input and error signals to $\pm 2$ units from zero just prior to multiplication (equivalent to filtering signals $m_\alpha$ and $t - t_m$ on the inputs to the multiplier in figure 8(b)). By locating the sigmoid filters thus, the large amplitude spikes of the input were still transformed through the adaptive filter but the increments in each $h$-parameter (figure 8(b)) were limited by the saturation filtering. This protected the $g$- and $w$-parameters against discontinuous jumps associated with occasional large amplitude spikes in either the input or error signal. With the addition of this saturation filtering the adaptive filter performed excellently. The filter no longer became unstable but, as can be seen in figures 10(b)–(d), because of the saturation filtering the filter output $y_m(k)$ tended to undershoot the large amplitude spikes in the system output $y(k)$.

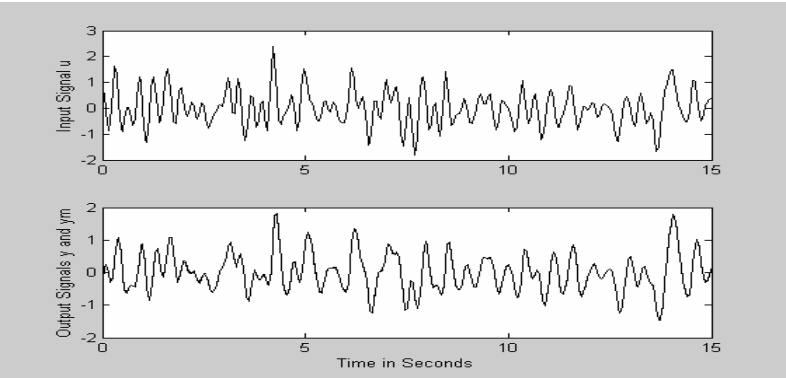
Each simulation run lasted 10 min corresponding to 12 000 samples in $u(k)$ at 20 samples/s. After each simulation a 15 s time-window (i.e., 300 samples at 20 samples/s) of (i) input $u(k)$, (ii) system output $y(k)$ and (iii) adaptive filter output $y_m(k)$ was taken arbitrarily from the ninth minute of the run. From this an error signal $e(k) = y(k) - y_m(k)$ was computed to give two performance measures, %ve, the variance of error as a percentage of the variance of $y(k)$, and %mse, the mean square value of the error signal as a percentage of the variance of $y(k)$. These are given in table 1 for each of the 24 simulation runs. The tendency mentioned above for $y_m(k)$ to undershoot large amplitude spikes in $y(k)$ introduces a non-zero mean into the error signals and accounts for the increase in the %mse scores relative to the %ve scores.

To provide an intuitive picture of how well the adaptive filter models a linear system and a nonlinear third-order Volterra system independently of the stochastic characteristics of the input, we show in figure 10 the waveforms from the 15 s time-window taken from four of the simulation runs. These correspond to the four shaded rows in table 1. The top trace in each part of the figure shows the input signal $u(k)$ while the bottom traces show the adaptive filter output signal $y_m(k)$ (dotted line) superimposed on the system output signal $y(k)$ (solid line).

In (a) input signal $u(k)$ is Gaussian with a bandwidth of 4.0 Hz. The system being modeled is a second-order linear system with a resonant frequency 2.0 Hz and damping ratio 0.5 (table 1, row 19). The output $y_m(k)$ from the adaptive filter exactly matches the output $y(k)$, illustrating the excellent performance of the adaptive filter when modeling a linear dynamical system with a Gaussian input. In (b) the input is the same as in (a) but the modeled system is a nonlinear third-order Volterra system with the same resonant frequency and damping ratio (table 1, row 20). Comparing (b) with (a) reveals that the output signal $y(k)$ in (b) displays a strong nonlinear relationship to the input $u(k)$. Despite this nonlinearity the output $y_m(k)$ from the adaptive filter closely matches $y(k)$ apart from some undershooting on large amplitude swings in $y(k)$. This is due to the previously mentioned influence of the stablizing saturation filtering required for other inputs. With this removed, undershooting errors did not occur.

Figure 10(c) is based on the same Volterra system as in (b) but now the input signal derives from an exponential APDF and is strongly non-Gaussian with skewness 7.64 and kurtosis 147.42 (table 1, row 10). Despite this, the signal $y_m(k)$ at the output of the adaptive filter closely matches $y(k)$. This demonstrates the ability of the adaptive filter to model accurately a nonlinear third-order Volterra system independently of the APDF and/or PSDF of the input signal. Figure 10(d) is based on the same Volterra system as (b) and (c) but with input derived from a Bernoulli-Gaussian APDF with low pulse probability 0.05 (table 1, row 24). This input is non-skew but strongly leptokurtic (kurtosis = 15.83) as can be seen in (d) where it spends much of the time close to zero with an occasional

(a) Gaussian (4 Hz bandwidth). Linear System




(b) Gaussian (4 Hz bandwidth). Nonlinear System

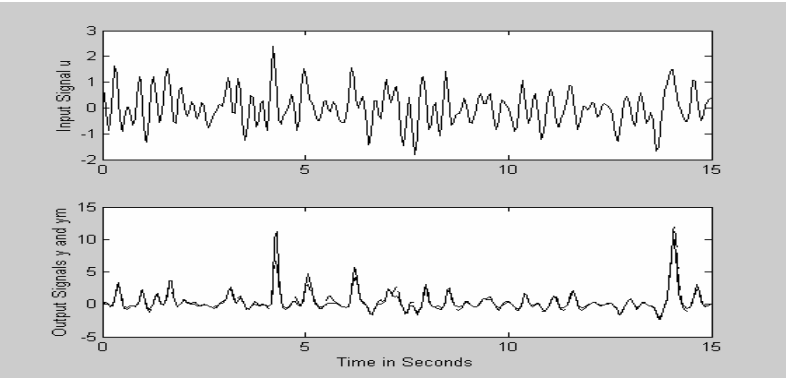


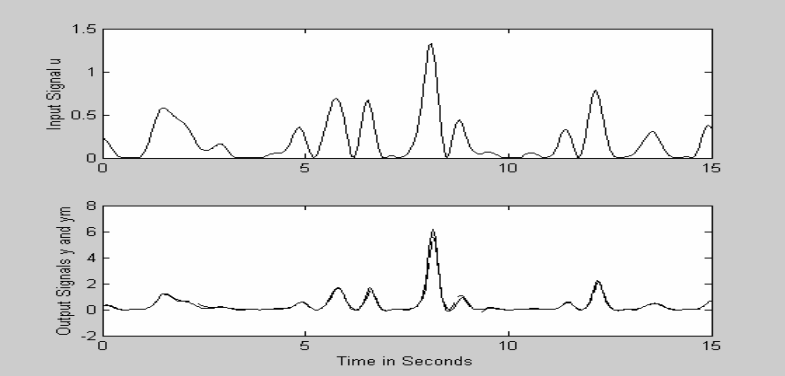

**Figure 10.** Waveforms from four (shaded rows in table 1) of the 24 simulation runs used to assess the performance of the nonlinear adaptive filter shown in figure 9. In each figure (a), (b), (c), (d) the top window shows a 15 s time sample of the stochastic input signal $u(k)$ sampled from the ninth minute of a ten minute simulation run. The bottom window shows the corresponding output signal $y_m(k)$ from the adaptive filter superimposed on the signal $y(k)$ from the output of the modeled system. The match between $y_m(k)$ and $y(k)$ indicates the performance of the adaptive filter in modeling either linear or nonlinear third-order Volterra systems with differing input APDFs and PSDFs. Details are given in section 7.4. (a) Gaussian input signal $u(k)$ with a bandwidth of 4 Hz driving a linear second-order modeled system with a resonant frequency of 2 Hz. The adaptive filter output $y_m(k)$ accurately matches the system output $y(k)$. (b) Gaussian input signal $u(k)$ (same as in (a)) driving a nonlinear third-order Volterra system. The adaptive filter output $y_m(k)$ matches the system output $y(k)$ except for an undershooting error on large amplitude peaks. (c) Input signal $u(k)$ with a bandwidth of 1 Hz and an extreme deviation from Gaussianity with skewness = 7.64 and kurtosis = 147.42 driving a nonlinear third-order Volterra system. Despite the extreme deviation from Gaussianity of the input, the adaptive filter output $y_m(k)$ closely matches the output $y(k)$ from the modeled nonlinear system. (d) Input signal $u(k)$ with a bandwidth of 4 Hz and with a strongly leptokurtic APDF derived by filtering random numbers with a Bernoulli–Gaussian APDF and a low spike probability of 0.05 through an eighth-order Butterworth digital filter with a bandwidth of 4 Hz. Despite the deviation of the APDF of the input signal from Gaussianity the output $y_m(k)$ from the adaptive filter closely matches the output $y(k)$ from the modeled nonlinear system.

random pulse. Again the simulation shows that the output from the adaptive filter $y_m(k)$ closely matches the output $y(k)$ from the nonlinear third-order Volterra system.

When saturation filters were removed from the four simulations in figure 10 the output signal $y_m(k)$ for runs with the exponential and Bernoulli–Gaussian input signals became unstable and performance measures approached infinity. For the nonlinear system in figure 10(b) with a Gaussian input the %ve score increased from 11.8% to 58.8% due to occasional bursts of instability in $y_m(k)$ from which the system recovered. Performance of the linear system in (a) with Gaussian input was not changed. Clearly, for satisfactory performance the LMS modules must be protected against the occasional large amplitude spikes in the input and/or error signals associated with non-Gaussian inputs and/or outputs.

In summary, these simulations demonstrate the ability of the adaptive filter circuit in figure 9 to tune its 30 $g$- and 30 $w$-parameters so that its input/output relationship accurately matches that of an arbitrary linear system or an arbitrary nonlinear third-order Volterra system independently of changes in the APDF and/or PSDF of the input signal. The filter has been tested with inputs of bandwidth up to 4.0 Hz (i.e., a bandwidth appropriate for human movement control) and with saturation filters for extremely ill-behaved inputs. It functions remarkably well independently of the amplitude and spectral characteristics of the signals. The method appears to be extremely promising. Because of its economy of parameters, and because it can be implemented by parallel operation of identical modules like those proposed to exist in corticocerebellar circuits (figure 8), we suggest that this method gives a viable basis for adaptive modeling of nonlinear dynamical relationships in the nervous system.

(c) Exponential and Square (1 Hz bandwidth). Nonlinear System




(d) Bernoulli-Gaussian (4 Hz bandwidth). Nonlinear System

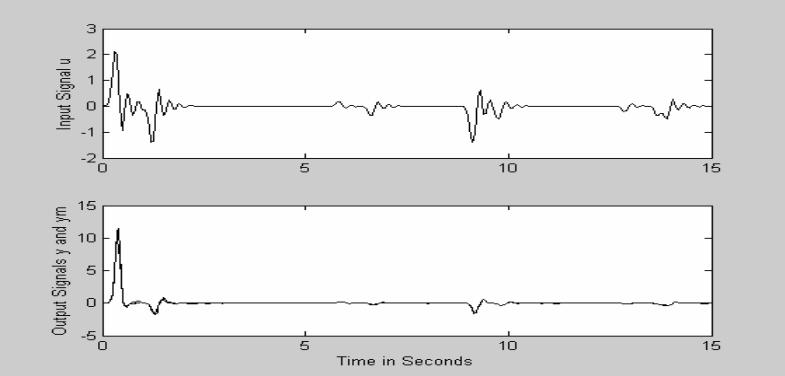



**Figure 10.** (Continued.)

## 8. Discussion

We have set out computational solutions to the movement control problems of redundancy, resources, and nonlinear interactions (sections 3, 4, 5) and have brought these together in a feedforward/feedback structure that introduces intermittency into the perceptual-motor loop (section 6, figure 7). Except for the plant, adaptive multivariable nonlinear internal models and the corresponding modeling processes are distributed throughout this structure. Thus in section 7 we considered how these might be implemented in a biologically realistic way with adaptive neural filter networks. The parallel processing network of figure 9 is economical with respect to parameters and successfully performs the required computations independent of the probability distributions of the signals processed. Plugged into the figure 7 blueprint, it offers a viable basis for modeling and transforming the systems and signals of movement control (section 2) in a way that is neurally feasible. Inasmuch as the filter networks accomplish the necessary modeling, the solutions we propose will work. We believe this provides a comprehensive means for simulating the planning and control of movement in a way not previously possible.

### *8.1. Synergies, degrees of freedom and modeling*

There are alternative ways of viewing the problems, in particular in regard to what is controlled. Take the wiring of fixed synergies for instance. Patterns of muscular output activated by connectivity in the spinal cord are described by some as *motor primitives* (Bizzi and Mussa-Ivaldi 2000). It has been argued (Mussa-Ivaldi 1999) that motor commands do not control individual muscles or individual joint torques but, via interneurons, modulate the viscoelastic force fields produced by sets of muscles. Yet the combination of adaptive task-dependent multijoint synergies (section 4) controlling wired-in patterns of muscles operating across elemental movements (section 3) has exactly the same effect; i.e., generation of viscoelastic force fields that influence broad regions of the body's state space. Mussa-Ivaldi and Giszter (1992) have shown that in the frog, muscles operating across multiple joints can be activated by stimulation in the spinal cord and, moreover, the force fields add vectorially during stimulation of multiple sites. On the other hand, it is simple to demonstrate the elemental movements that, by definition, humans control independently, one at a time. Whether multijoint movements are controlled by a combination of wired-in elemental movement synergies as suggested here or whether individual elemental movements are controlled by a combination of multijoint synergies, as implied by the frog data, seems a minor difference. One is simply a coordinate transformation of the other accomplished by a matrix multiplication of the connectivity in the descending pathways that, according to each theory, implement the wired-in synergies or motor primitives. The experimental evidence can be seen as supporting both views.

When differing theoretical formulations are being considered the matter of degrees of freedom can easily mislead. As presented in section 4 the concept of *control* DFs differs from that of *biomechanical* DFs. It also differs from that of *dynamical* DFs used by those who conceive skilled movement as the free motion of a nonlinear dynamical system with a limit cycle attractor (Mitra *et al* 1998, Newell *et al* 2003). The number of dynamical DFs is defined to be the minimum number of first-order differential equations required to capture the system's time-evolving behavior (i.e., the dimension of the state vector). In AMT we are concerned with control of a nonlinear dynamical system rather than with its free motion. We are concerned not with the dimension of the state vector but with the minimum number of input/output equations required to capture the system's characteristics (i.e., the dimension of the control vector). This dimension specifies the number of control DFs. As shown in section 5 the control vector (i.e., the vector $\mathbf{\Delta M}^*$ with dimension corresponding to the control DFs) operates on the dynamical system and renders the state variables (corresponding to the dynamical DFs) unobservable in the outputs. While consideration of the free motion of nonlinear systems is interesting and attractor dynamics must be taken into account during formation of internal models, the state variables do not provide a coordinate system that informs about control.

In contrast many theorists do concern themselves with control DFs. The question frequently focuses on which coordinate system the CNS uses to reduce dimensionality, end-point of effector, joint positions, joint torques, and center of mass being among the many possibilities (see Scholz and Schöner (1999)). Despite considerable experimental work there is no universal answer, and there appear to be different likely candidates for different tasks. A recent approach is to consider the role of dimensions that are not controlled (Scholz and Schöner 1999, Todorov and Jordan 2002). These findings are not surprising in the context of AMT where a reduced task space is determined by orthogonalization of reafferent signals associated with task responses. Signals that are not relevant to the task will not play a role in the modeling of its feature space as in the concept of an uncontrolled manifold.

For AMT experimental work on reducing the DFs in a task has more been on the process of developing task-dependent synergies (i.e., control or coordinative structures) rather than on what structure is developed. Recently, we obtained experimental evidence providing compelling support for the proposal that the CNS can learn the dynamic relationships between unaccustomed bimanual movements and can use this information to form novel **TDSG**s (Neilson and Neilson 2002). There are also experimental observations concerning the accuracy of internal models (Davidson *et al* 2000, Ghous and Neilson 2002) that support the inclusion of postural summers $\mathbf{\Delta}^{-1}$ and the associated use of first-difference signals for model formation (figure 4). If participants cannot see the response cursor in a visual pursuit tracking task, performance deteriorates at low frequencies (<0.5 Hz) but remains accurate at high frequencies up to 2.0 Hz. This fits the intuition that feedforward control using an accurate internal model is important for accuracy of high-speed movement (e.g., saccades and fast reaches) while feedback error correction is important for accuracy of low-speed movement. Reaching behavior, for example, shows a fast initial movement followed by a sequence of error corrections to acquire the end-point (e.g., Meyer *et al* (1990)).

We indicated in section 4.4 that without orthogonalization it is not possible to form forward and inverse models of multivariable systems with highly interrelated inputs and with unequal numbers of inputs and outputs. Even with orthogonalization of the signals, there remains a potential problem for forming an accurate inverse model. While the modeled system may be stable, or at least can be stabilized by feedback, there is no guarantee the inverse system will be stable. Indeed in laboratory simulations using zero-order hold (ZOH) equivalents of continuous-time systems, we have found that the inverse of the ZOH equivalent system is very often unstable. An unstable inverse model cannot be used as a feedforward controller. Nevertheless, as shown mathematically by Isidori (1995), it is possible to stabilize the inverse model, but only at the expense of detuning it. It can no longer be an accurate inverse and therefore it cannot both cancel the dynamics and decouple the system as can an accurate inverse. Either cancellation can be favored at the expense of decoupling or vice versa.

The consequence of using such a detuned model as a feedforward controller is interesting. The actual response trajectory will no longer accurately match the intended trajectory. What happens if such circumstances are encountered in human movement control? We are currently investigating this using a two input/two output cross-coupled tracking system. Participants perform a visual pursuit tracking task where the system has either a stable inverse or an unstable inverse. Preliminary results show that the latter task can be performed without responses becoming unstable but compared to tracking with a stable inverse, accuracy is considerably poorer. Responses show decreased amplitude and longer time lags, but cross-talk (interaction) between channels increases only marginally. This suggests that the CNS favors decoupling interacting responses at the expense of maintaining their accuracy.

We have yet to explore how the mental rehearsal process (figure 6) might detune the inverse model to achieve decoupling with stability. Meanwhile mental rehearsal is supported experimentally as having a role in internal model acquisition (Bhushan and Shadmehr 1999). These authors advocate a system of forward and inverse modeling with mental rehearsal as the basis of learning and consolidating novel tasks. Unlike in early days of AMT, the concept of forward and inverse models in the CNS, particularly in the cerebellum, is now widely endorsed both theoretically and experimentally (for review see Desmurget and Grafton (2000), Kawato (1999), Wolpert *et al* (1998)) but the configurations in which they are used vary. Davidson *et al* (2002) draw attention to the difficulty of obtaining an inverse model online for a nonlinear system given that the analytical method that holds for a linear system no longer works. The same is true for MIMO systems. However, we have shown experimentally (Ghous and Neilson 2002) that the CNS can form internal

models of a nonlinear MIMO tracking system independently of the APDF of the signals. We also know that a model of a nonlinear MIMO can be successfully tuned online by the method of figure 6.

### *8.2. Intermittency, submovements, trajectories and speed-accuracy*

In section 6 we combined the computational solutions of previous sections into a feedforward/feedback structure that includes a fixed-time planning process introducing refractoriness and intermittency into control of continuous movement (figure 7). To make meaningful comparisons between performance of a simulator and performance of humans in an experimental task such as visual tracking, it is necessary to have an estimate of the probable time constraints in the nervous system. As reviewed by van Beers *et al* (2004), estimates of time delay for proprioceptive feedback vary between 100 and 200 ms. Schmidt and Lee (1999) indicates a similar range for visual feedback processing in motor tasks. Consonant with this, Stark (1968) gives the average refractory period for saccadic eye movements as 200 ms. The intermittent character of saccadic movement is well known, but intermittency pervades even apparently continuous movement and has a long history dating back to Craik (1947, 1948). Consistent with the 100–200 ms timing above, intermittency at a frequency around 10 Hz has been observed in electromyograms and velocity profiles of slow finger movements (Vallbo and Wessberg 1993). More recently, discontinuities at 6–9 Hz in finger movements have been shown to be strongly coherent with synchronized activity in cerebello-thalamo-cortical loops (Gross *et al* 2002). These authors in fact suggest that coupling of activity in this 6–9 Hz range represents the neural mechanism for the intermittent control of continuous movements.

In earlier studies we have implemented the figure 7 structure in more basic forms to simulate movement planning and control. Setting the value of the **RP** fixed time interval in the 100–200 ms range gives good agreement with results of visual tracking experiments (Neilson *et al* 1988a, 1993, 1995). Regardless of the dynamics of the tracking system, open-loop tracking characteristics always converge with practice to a gain, a time delay and a first-order lag low-pass filter (McRuer and Krendel 1974). This is reproduced in AMT simulations of tracking and it equates to formation of accurate internal models of the musculoskeletal system in interaction with the tracking system (Neilson *et al* 1988a). Accurate models (acquired with practice) render the dynamics of **MCS**, **BM** and **E** systems unobservable in tracking behavior, leaving only the influence of intermittent response planning. We have shown (Neilson *et al* 1988a) that the dynamics of intermittent response planning correspond to the gain, time delay and first-order lag filter observed experimentally by McRuer and Krendel. Intermittent planning in a fixed time interval means that each response runs open loop during the subsequent planning time. A 100–200 ms planning time implies intermittent error corrections at five to ten times a second. We have shown that this limits coherent feedback tracking of stochastic input signals to bandwidths of 2.0 Hz or less and have confirmed with visual tracking data that this is the case (Neilson *et al* 1993).

These simulations are predicated on a basic proposal of AMT that purposive movement is broken into a concatenated sequence of submovements that we call *basic units of motor production* or BUMPs (Neilson *et al* 1992). We emphasize that a BUMP duration of 100–200 ms does not imply that comparable segments of movement will be observed in responses. A number of researchers have created movement-parsing algorithms to detect the boundaries of submovements via discontinuities in velocity and acceleration waveforms (e.g., Meyer *et al* (1988)). With these automatic methods conveniently available in computer programs, kinematic discontinuities have become a sort of (unintended) way of defining 'submovement' boundaries. The resulting reports of submovement durations apparently longer than 200 ms do not invalidate our simulations. When an actual response trajectory **R** matches the intended response trajectory **R**$^*$, the position and velocity at the beginning of one submovement will exactly match that at the end of the previous submovement. The two movements will join together smoothly with no detectable discontinuities in position, velocity or acceleration. Discontinuities appear only when **R** deviates from **R**$^*$ as a result of lack of skill and/or other uncertainties in response execution. It is therefore not surprising that discontinuities in velocity and acceleration waveforms are found on average at more than 200 ms apart.

The behavior of the discrete-time ZOH minimum acceleration trajectories generated in AMT simulations (Neilson *et al* 1995) differs from that of the continuous-time trajectories described by others (Flash and Hogan 1985, Uno *et al* 1989). The difference is most striking for a fast ballistic movement of duration 100 ms. Using a 50 ms sample interval our discrete-time trajectory generator produces a positive 50 ms duration rectangular pulse of acceleration followed by a negative 50 ms duration rectangular pulse of deceleration. The position trace undergoes constant acceleration for 50 ms and then constant deceleration for 50 ms and the velocity profile is triangular. Except for the sharp corners in the velocity and acceleration waveforms created by using a rectangular pulse to approximate a bell-shaped pulse (see appendix A.1), this simulated trajectory closely matches fast ballistic movement trajectories observed in man (Berardelli *et al* 1996). A discrete-time trajectory generator also accounts for the phenomenon of isochrony that is not explained by continuous-time trajectory generation. The 100 ms duration fast ballistic movement in man does not change as the amplitude of the movement is increased. This shows that the maximum speed of a fast movement is limited by the timing of the descending neural commands, as for the discrete-time ZOH trajectory generator, and not by the maximum force-generating capabilities of the muscles.

By configuring the discrete-time trajectory generator to provide variable horizon control (section 6.2) we have obtained insights into speed-accuracy phenomena. Unskilled subjects slow and stiffen their movements. In AMT such behavior corresponds to an optimal control strategy involving increasing

the prediction horizon and detuning the inverse model (Neilson and Neilson 1999a). This reduces the gain, decreases the bandwidth and increases the phase lag of error correction in the perceptual-motor loop. Responses are slowed as a consequence. Reduced gain improves stability margins and so increasing the prediction horizon provides a defense against instability. Slowing $\mathbf{R}^*$ by stretching the prediction horizon has another important consequence. For muscles working against predominantly viscoinertial mechanical loads, a reduction in the speed of movement implies a reduction in the net turning and/or displacement forces that the muscles must generate to produce the movement. Variance of broadband noise in neural drives to muscles increases with the level of neural drive or force (Harris and Wolpert 1998, O'Dwyer and Neilson 1998, Schmidt and Lee 1999) and so slowing responses reduces the variance of neural noise and this in turn improves the accuracy of the movement. Also, a slowed $\mathbf{R}^*$ allows more intermittent error corrections to be made in reaching a fixed or slowly moving target and this improves low-speed movement accuracy.

The combination of noise in the nervous system and intermittent error correction gives rise to a logarithmic speed-accuracy tradeoff as in Fitts' law. With the prediction horizon fixed, it also predicts asymmetrical velocity profiles similar to those described by Plamondon and Alimi (1997). Alternatively, if the prediction horizon is changed from one planning interval to the next as required in reaching tasks with a target fixed in both space and time, the theory predicts symmetrical velocity profiles and linear speed-accuracy tradeoff as observed experimentally (Schmidt and Lee 1999). Olsen (2001) has extensively investigated the phenomenon of noise in the nervous system and speed-accuracy tradeoff by means of AMT simulations and has demonstrated these to be consistent with the behavioral evidence.

### *8.3. Model acquisition, storage and selection*

In the 1980s the concept of feedforward controllers based on adaptive inverse models was not familiar in discussions of motor control and learning. Experimental work was often aimed at demonstrating this as a viable theoretical approach. Now that the existence of internal models is virtually a given in the computational approach to human movement, the extent and variety of experimentation has escalated. A current focus is the multiplicity of models that underlie skilled behavior: on whether they encapsulate kinematic and dynamic information separately or together (Krakauer *et al* 1999, Tong *et al* 2002, Wolpert *et al* 2001), on how they are represented and stored (Mussa-Ivaldi 1999, Wolpert *et al* 2001), on how they consolidate or interfere (Baraduc *et al* 2004, Davidson and Wolpert 2003, Shadmehr and Brashers-Krug 1997, Wigmore *et al* 2002), on how they are appropriately selected (Haruno *et al* 2001, Jacobs *et al* 1991, Wolpert *et al* 2003, Wolpert and Kawato 1998), and on if and how they generalize, combine or fragment (Flanagan *et al* 1999, Goodbody and Wolpert 1998, Kawato 1999, Karniel and Mussa-Ivaldi 2002, Sainburg *et al* 1999).

In AMT the **TDSG** and the forward and inverse input/output relationships of the **MCS**, **BM** and **E** systems are modeled independently and in parallel. This has been shown experimentally for dynamic and kinematic systems (Krakauer *et al* 1999, Jansen-Osmann *et al* 2005). It cannot happen with feedback error learning (Kawato and Gomi 1992) where the learned model corresponds to the inverse of the entire plant. Existence of separate models for the **TDSG** and the **MCS**, **BM** and **E** subsystems has the advantage that if a change occurs in only one of the systems, it is only this system that requires an updated model. For example, if a skilled automobile driver was to change grip on the steering wheel to a novel grip then he/she may have to learn a new task-dependent synergy but the internal models of the driver's muscle control systems and biomechanics remain the same, as does that for the external system defined by the relation between steering wheel movement and change of automobile heading. Consequently, one can predict a strong transfer of skill between the two driving tasks, as we know intuitively would be the case.

Proposals that depend on feedback error learning to acquire internal models must account for skill transfer differently. Take the MOSAIC model for sensorimotor learning and control (Haruno *et al* 2001, Wolpert and Kawato 1998). According to this proposal a relatively small number of paired forward and inverse model primitives are held on-line and, by modulating the outputs of the inverse models, an enormous repertoire of behaviors can be generated. From this comes the notion that new models can be created through a process of composition and decomposition of old models (Davidson and Wolpert 2004, Flanagan *et al* 1999, Ghahramani and Wolpert 1997). While AMT predicts that new skills can be acquired by forming new combinations of established **TDSG**, $\mathbf{E}_m{}^{-1}$, $\mathbf{BM}_m{}^{-1}$ and $\mathbf{MCS}_m{}^{-1}$ models, we are not convinced that individual models (such as $\mathbf{BM}_m{}^{-1}$) can be composed from a combination of simpler models. In a visual pursuit tracking experiment we have found that practicing the left hand and then practicing the right hand did not help in performing the two-handed task composed of the two one-handed tasks put together. Practicing the two-handed task did improve performance on the two-handed task and also on each of the one-handed tasks (Williams and Neilson 1998). In other words, a two-handed task could be decomposed into two one-handed tasks but two one-handed tasks could not be composed into one two-handed task. This is against the idea that complex models can be constructed from a combination of previously learned simpler models.

How models are stored and retrieved from a repertoire is an area yet to be fully developed in AMT and is a present focus. We are proceeding from the proposal that the parameters for all adaptive filter models are stored into and retrieved from long-term memory under control of response planning processes (Neilson *et al* 1992). We regard skilled motor behavior as a concatenated sequence of task-dependent synergies and refer to these functional units of movement as *synergemes* (Neilson and Neilson 2004). Thus a behavior such as lifting a glass, drinking, then replacing it, is comprised of a concatenated sequence of synergemes; e.g., reach, grasp,

pickup, transport, and so on. Each synergeme represents a class of skilled movement acquired through practice. Each synergeme requires a different **TDSG** and/or inverse model. In the transition from one synergeme to another, the feedback/feedforward controller of figure 7 must be able to switch internal models quickly and smoothly in the correct sequence and with appropriate timing. Since we hypothesize that these seamless transitions occur in the normal course of events, obtaining experimental evidence on model switching is likely to be difficult. Nevertheless we have serendipitously found a circumstance that appears to reveal rapid switching of models during a brief period of interference when subjects return from a highly practiced incompatible task to normal tracking (Neilson *et al* 1998).

We are currently considering the process of model selection as follows. Selection of a controller implies retrieval of a set of adaptive filter parameters from long-term memory. The retrieved set of parameters corresponds to a pattern of neural activity held on line in cortical working memory. This activity modulates the gains of neural adaptive filters. We see working memory as portals to recursively connected neurons that function as long-term memory association networks. Partial activation recalls the entire pattern, as in a Kohonen neural network. We refer to the pattern of activity that recalls the parameters for a particular controller as a *gesture code*. Gesture codes are created within sensory analysis systems in response to temporospatial patterns of reafference. This reafference becomes associated with the spatial pattern of cortical activity for the parameters of the controller that created it. Many different gestures might associate with each controller. In this way, over time, the CNS could establish associations between a large variety of gesture codes (clenching the fist, grasping with finger and thumb, lifting the foot, turning the head, arching the back, etc) and the parameters of the corresponding controllers. Each gesture code is a symbol that points to the adaptive parameters for an appropriate controller, just as a library catalogue points to the accession number that retrieves the right book.

The symbolic nature of gesture codes suggests a role in higher level cognitive processing. Implementation of a sequence of gesture codes will retrieve the parameters for the appropriate controllers to accomplish the task creating a sequence of goal-oriented synergemes. The gesture codes sit at the interface of cognitive and motor processing, much as the symbolic elements of language give rise to speech. This suggests that the parallels with prosodic, syntactic, morphemic and phonemic processing will be informative in understanding the selection and sequencing of adaptive parameters for implementing skilled movement in general. These ideas will hopefully provide fruitful extensions to AMT.

### *8.4. Developmental learning*

Theory concerned with adult skill acquisition should also be relevant to the motor learning that occurs during development. To touch on this briefly, we already put forward in section 3 the proposal that the learning of wired-in synergies begins in utero as slowly-tuning motor maps. Like the slowly-tuning sensory maps that come to define elemental movements and functional muscles, the motor maps are bootstrapped by spontaneous fetal movement (Neilson and Neilson 2005). We propose that the formation of feature extraction networks and of forward and inverse models comes later in development, beginning around two months after birth when there is a well-known rapid acceleration in motor development. We attribute this developmental spurt to the emergence of task-dependent motor learning in the form of the task-dependent synergy generators (section 4). This is in keeping with the time that neural circuitry in cortico-basal ganglia-cortical and cortico-cerebellar-cortical loops is likely to myelinate (Volpe 1987). It is exactly this circuitry that we suggest has the ability to form neural adaptive filter models of nonlinear dynamical relationships between neural signals.

We have argued that the primary functional disability in individuals severely disabled with cerebral palsy is a consequence of neural lesions that disrupt the adaptive modeling of efference copy signals during this latter stage of motor development (Neilson and O'Dwyer 1984, Neilson *et al* 1997). This early prediction from AMT was borne out by positron-emission tomography and ultrasound studies (e.g., Volpe *et al* (1985), Weindling *et al* (1985)) associating the diagnosis of cerebral palsy with white matter lesions known as periventricular leukomalacia. This impairment of cortico-cerebellar connections will, according to AMT, disrupt the availability of efference copy to the cerebellar modeling circuits, thus impairing formation of a repertoire of correct **TDSG** and **MCS** models. This is keeping with the fact that cerebral-palsied individuals move inappropriately, even though they know exactly what they want to do. Moreover, without efference copy the CNS is unable to relate movement feedback with the motor commands that generated it. Inappropriate movements are consequently interpreted as illusions of exafference and subjectively experienced as involuntary (Neilson *et al* 1992, 1997). The increased levels of co-contraction and increased stretch reflex sensitivities seen in cerebral-palsied individuals may in part be attributed to a stiffening reaction to protect against the falsely perceived exafference. Because lack of efference copy has no impact on the development of the **WSG** this proposal is consistent with the fact that cerebral palsy is very difficult to detect behaviorally in the newborn.

### *8.5. Conclusion*

In AMT we seek to provide a mathematical description of the computational processing that must take place within the CNS to account for observed motor behavior. It describes control of the entire musculoskeletal system in interaction with its environment and takes into account the complex, multivariable, changing, uncertain, nonlinear and dynamical nature of the control problem. It is a comprehensive theory with a single set of control principles applying for all movements. In other words, we do not regard balance, posture, eye movements, reaching, speech, and so forth, as if they were unrelated processes.

The solutions proposed here for this difficult control problem are biologically plausible. All the computational

blocks in figure 7, except for the PLANT, are constructed from networks of neural adaptive filters. Even the **RP**-system can be constructed using adaptive filters (Neilson *et al* 1995). But each block is comprised of multiple adaptive filters, each adaptive filter is comprised of multiple LMS modules, and each LMS module corresponds to the neural circuitry in a cortical-subcortical loop. In other words, all the processing in our proposed computational structure is potentially performed in the CNS by the repetition of a simple module repeated over and over very many times in parallel. This is consistent with the observation that the same basic circuitry is repeated everywhere throughout the neocortex (Douglas and Martin 2004) and that cortico-cerebellar-cortical and cortico-basal ganglia-cortical connections are organized into modules that are likewise repeated in discrete circuits or loops (Middleton and Strick 2000).

The CNS has yet to reveal many computational secrets for proposals to be endorsed or otherwise. Whether or not it implements LMS algorithms ubiquitously or at all is presently a moot point. But what we can do is configure the adaptive filter modules of AMT as a multivariable, nonlinear, dynamical control system and use this as a predictor of actual behaviors. Following the philosophy of Gregory (1978), the simulation structure set out in figure 7 is a precise hypothesis about the neurobiological processes involved in human movement control. If it predicts correctly, as have earlier implementations to date, then we can say that we know one way the CNS may do it. If the prediction fails, we know to look elsewhere.

## Acknowledgments

This work was supported by Australian Research Council Large Grant A00106147 '*How babies learn to move: simulation based on a generalised theory of motor control and development*' as well as by previous funding from the Australian Research Council, the National Health and Medical Research Council of Australia, and the Spastic Centre of New South Wales. We thank Zhanwen Li and Asim Ghous for programing contributions in developing the nonlinear adaptive filter network.

## Appendix A.1. Discrete-time signals

Throughout the paper we employ discrete-time signal theory to describe neural signals and relationships between neural signals. There are two reasons for this. The first comes from recognition that transmission and central processing time delays are important in shaping the dynamical behavior of the nervous system. Time delays are easily included in discrete-time models but not so readily in continuous-time models. The second reason concerns the nature of neural signals. The neural filtering processes of AMT involve signals that correspond to temporospatial patterns of action potentials in ensembles of nerve fibers. A temporospatial pattern can encode information in many ways (DeCharms and Zador 2000) but the simplest involves the total number of impulses arriving per unit of time; i.e., by the summation of impulses arriving over a time interval. Such a basis for encoding and decoding information by ensembles of neurons is consistent with the structure of cortical columns (Mountcastle 1978, 1993, 1998) and with the bursting patterns of ensembles of neurons in the cortex (Braitenberg 1986, Shaw and Silverman 1988, von Seelen *et al* 1986, Sardesai *et al* 2001). The burst pattern of cortical columns is estimated by constructing a post-stimulus histogram of the activity of individual neurons in the ensemble. This provides a statistical estimate of the number of active neurons in the ensemble as a function of time. Typically a post-stimulus histogram shows a sequence of bell-shaped bursts at a frequency of 20–40 Hz that decrease in amplitude with time. Also, recent work on the phenomenon of corticomuscular coherence shows that oscillatory activity at around 20 Hz transmits from motor cortex to the spinal motor neurones (see Salenius and Hari (2003)).

The amplitude of neural bursts seems a very likely way for the nervous system to encode and transmit information. While not exact, each bell-shaped burst can be approximated by a rectangular pulse. Consequently, the bursts of temporospatial activity in the nerve fiber bundles that interconnect cortical and subcortical structures with the motor neuron pools can be approximated by a staircase waveform of concatenated rectangular pulses. Such a staircase waveform is well known in discrete-time signal processing as a zero order hold (ZOH) signal. ZOH signals are continuous-time signals but they are easily realized in discrete time by sampling the height of each step in the staircase. Thus the decision to approximate temporospatial patterns of neural signals by ZOH waveforms opens a vast library of existing digital signal processing techniques for use in AMT analyses and simulations. This seems reasonable given the evidence on neural bursts and it might further be argued that the processing of ZOH signals provides a good approximation to the type of processing actually performed in the nervous system.

Following from the above, the signals used throughout this paper should be taken as discrete-time sampled equivalents of ZOH continuous-time signals. The sample interval is not specified but, based on the 20–40 Hz bursting of cortical activity, it is likely to be fixed between 25–50 ms. We use the usual discrete-time notation $u(k)$ to represent these signals but for brevity a signal is often given simply as $u$ with sampling number $k$ implicit. A single ZOH signal is shown unbolded (e.g., $q_1$) while a vector of ZOH signals is given in bold (e.g., $\mathbf{q}$). As in discrete-time signal theory $z^{-1}$ represents a backward shift of one sample interval while $z^{-\rho}$ represents a backward shift of $\rho$ sample intervals. The symbol $\Delta = 1 - z^{-1}$ represents a first-difference operator and its inverse $\Delta^{-1} = 1/(1 - z^{-1})$ represents a summer. Bold $\boldsymbol{\Delta}$ defines an array of first-differencers or equivalently, a diagonal matrix of first-differencers $\Delta$, where the dimension of the matrix is taken to be the dimension of the vector of input signals. Likewise $\boldsymbol{\Delta}^{-1}$ defines an array of summers or equivalently, a diagonal matrix of summers $\Delta^{-1}$ with dimension equal to that of the vector of input signals.

## Appendix A.2. Table of symbols

This table describes the symbols used in the paper and gives the section number where they first appear.

| Symbol | Description | Location |
|---|---|---|
| $\boldsymbol{\alpha}_{\mathbf{m}} = [\alpha_{\mathrm{m1}}, \ldots, \alpha_{\mathrm{m700}}]^{\mathrm{T}}$ | Vector of 700 descending alpha drives to pools of alpha motor neurons | Section 2 |
| APDF | Amplitude probability density function | Section 7.3 |
| **BM** | Biomechanical system | Section 2 |
| $\mathbf{BM}_{\mathrm{m}}$ | Model of biomechanical system based on orthogonalized differenced signals $\boldsymbol{\Delta}\mathbf{T}$ and $\boldsymbol{\Delta}\mathbf{Q}$ | Section 4.4 |
| $\mathbf{BM}^{-1}$ | Inverse of biomechanical system | Section 4.4 |
| $\mathbf{BM}_{\mathrm{m}}{}^{-1}$ | Model of inverse biomechanical system based on orthogonalized differenced signals $\boldsymbol{\Delta}\mathbf{Q}$ and $\boldsymbol{\Delta}\mathbf{T}$ | Section 4.4 |
| $\mathbf{c} = [c_1, \ldots, c_{110}]^{\mathrm{T}}$ | Vector of co-contractions of functional muscles about 110 elemental movements | Section 2 |
| CNS | Central nervous system | Section 1 |
| DF | Degree of freedom | Section 4.1 |
| $\boldsymbol{\Delta}\mathbf{D}$ | Vector of first-differenced disturbance signals | Section 5.3 |
| $\boldsymbol{\Delta}\hat{\mathbf{D}}$ | Vector of first-differenced disturbance signals predicted ahead in time | Section 6.2 |
| $\boldsymbol{\Delta}$ | Diagonal matrix of first-differencers | Section 4.3 |
| $\boldsymbol{\Delta}^{-1}$ | Diagonal matrix of summers | Section 4.3 |
| **E** | External system | Section 2 |
| $\mathbf{E}_{\mathrm{m}}$ | Model of external system based on orthogonalized differenced signals $\boldsymbol{\Delta}\mathbf{Q}$ and $\boldsymbol{\Delta}\mathbf{R}$ | Section 4.4 |
| $\mathbf{E}^{-1}$ | Inverse of external system | Section 4.4 |
| $\mathbf{E}_{\mathrm{m}}{}^{-1}$ | Model of inverse external system based on orthogonalized differenced signals $\boldsymbol{\Delta}\mathbf{R}$ and $\boldsymbol{\Delta}\mathbf{Q}$ | Section 4.4 |
| $e(k)$ | Error signal driving $w$-parameters in nonlinear adaptive filter | Section 7.3 |
| $\mathbf{e}_1(k)$ | Vector of error signals generated by comparator at output of forward model | Section 5.3 |
| $\mathbf{e}_2(k)$ | Vector of error signals generated by comparator at output of inverse model | Section 5.3 |
| $\mathbf{g}$ | Vector of gravitational inputs to biomechanical system | Section 2.3 |
| $g_1, \ldots, g_{30}$ | 30 adaptive $g$-parameters in nonlinear adaptive filter | Section 7.3 |
| $\mathbf{l} = [l_1, \ldots, l_{700}]^{\mathrm{T}}$ | Vector of lengths of 700 functional muscles | Section 2.2 |
| $\boldsymbol{\Delta}\mathbf{l} = [\Delta l_1, \ldots, \Delta l_{700}]^{\mathrm{T}}$ | Vector of first-differenced lengths of 700 functional muscles | Section 3.3 |
| $\Delta l_i / \Delta \theta_j$ | Ratio of first difference of length of $i$th functional muscle to first difference of $j$th elemental movement | Section 3.3 |
| $\boldsymbol{\Delta}\mathbf{l}/\boldsymbol{\Delta}\boldsymbol{\theta}$ | Approximation to Jacobian matrix with $ji$th element given by $\Delta l_i / \Delta \theta_j$ described above | Section 3.3 |
| LMS | Least mean square adaptive filter algorithm | Section 3.3 |
| LTM | Long-term memory | Section 7.1 |
| $m_\alpha$ | Motor signal in cortico-cerebellar-cortical circuit | Section 7.1 |
| $\mathbf{m}_\alpha = [m_{\alpha 1}, \ldots, m_{\alpha 220}]^{\mathrm{T}}$ | Vector of 220 central motor commands driving the task-dependent synergy generator | Section 3.1 |
| $\boldsymbol{\Delta}\mathbf{m}_\alpha = [\Delta m_{\alpha 1}, \ldots, \Delta m_{\alpha 220}]^{\mathrm{T}}$ | Vector of 220 first-differenced central motor commands. Assumed to be efference copy | Section 4.3 |
| **M** | $N$-dimensional vector of orthogonalized motor commands | Section 4.1 |
| $\boldsymbol{\Delta}\mathbf{M} = [\Delta M_1, \ldots, \Delta M_N]^{\mathrm{T}}$ | $N$-dimensional vector of orthogonalized first-differenced motor commands | Section 4.3 |
| $\boldsymbol{\Delta}\mathbf{M}^* = [\Delta M_1{}^*, \ldots, \Delta M_N{}^*]^{\mathrm{T}}$ | $N$-dimensional vector of orthogonalized first-differenced centrally generated required motor commands | Section 4.3 |
| $\boldsymbol{\Delta}\mathbf{M}$ feature extraction | Adaptive orthogonal first-differenced motor command feature extraction network | Section 4.3 |
| **MCS** | Muscle control system | Section 2 |
| $\mathbf{MCS}_{\mathrm{m}}$ | Model of muscle control system based on orthogonalized differenced signals $\boldsymbol{\Delta}\mathbf{M}$ and $\boldsymbol{\Delta}\mathbf{T}$ | Section 4.4 |
| $\mathbf{MCS}^{-1}$ | Inverse of muscle control system | Section 4.4 |
| $\mathbf{MCS}_{\mathrm{m}}{}^{-1}$ | Model of inverse muscle control system based on orthogonalized differenced signals $\boldsymbol{\Delta}\mathbf{T}$ and $\boldsymbol{\Delta}\mathbf{M}$ | Section 4.4 |
| $\mu$ | Adaptive gain constant | Section 7.1 |
| $N$ | Number of control degrees of freedom in a task | Section 4.3 |
| $N_m$ | Maximum number of orthogonal feature signals that can be controlled independently, simultaneously. | Section 4.1 |
| PSDF | Power spectral density function | Section 7.3 |
| $\mathbf{q} = [q_1, \ldots, q_{220}]^{\mathrm{T}}$ | Vector of 220 elemental movements and co-contractions | Section 2 |

| Symbol | Description | Location |
|---|---|---|
| $\mathbf{\Delta q} = [\Delta q_1, \ldots, \Delta q_{220}]^{\mathrm{T}}$ | Vector of 220 first-differenced elemental movements and co-contractions | Section 4.3 |
| $\mathbf{Q}$ | $N$-dimensional vector of orthogonalized elemental movements and co-contractions | Section 4.1 |
| $\mathbf{\Delta Q} = [\Delta Q_1, \ldots, \Delta Q_N]^{\mathrm{T}}$ | $N$-dimensional vector of orthogonalized elemental movement/co-contraction feature signals | Section 4.3 |
| $\mathbf{\Delta Q}^* = [\Delta Q_1{}^*, \ldots, \Delta Q_N{}^*]^{\mathrm{T}}$ | Vector of $N$ orthogonal first-differenced required elemental movement/co-contraction feature signals | Section 4.3 |
| $\mathbf{\Delta Q}$ feature extraction | Adaptive orthogonal first-differenced elemental movement/co-contraction feature extraction network | Section 4.3 |
| $\mathbf{r} = [r_1, \ldots, r_p]^{\mathrm{T}}$ | Vector of $p$ reafferent response signals | Section 2 |
| $\mathbf{\Delta r} = [\Delta r_1, \ldots, \Delta \mathrm{r}_p]^{\mathrm{T}}$ | Vector of $p$ first-differenced reafferent response signals | Section 4.3 |
| $\mathbf{R} = [R_1, \ldots, R_N]^{\mathrm{T}}$ | Vector of $N$ orthogonalized response feature signals | Section 4.1 |
| $\mathbf{\Delta R} = [\Delta R_1, \ldots, \Delta R_N]^{\mathrm{T}}$ | Vector of $N$ orthogonalized first-differenced response feature signals | Section 4.3 |
| $\mathbf{\Delta R}$ feature extraction | Adaptive orthogonal first-differenced response feature extraction network | Section 4.3 |
| $\hat{\mathbf{R}} = \left[\hat{R}_1, \ldots, \hat{R}_N\right]^{\mathrm{T}}$ | Vector of $N$ orthogonalized predicted future values of response feature signals | Section 6.2 |
| $\mathbf{R}^* = [R_1{}^*, \ldots, R_N{}^*]^{\mathrm{T}}$ | Vector of $N$ centrally-generated orthogonal required response feature signals | Section 4.3 |
| $\mathbf{\Delta R}^* = [\Delta R_1{}^*, \ldots, \Delta R_N{}^*]^{\mathrm{T}}$ | Vector of $N$ centrally-generated orthogonal first-differenced required response feature signals | Section 4.3 |
| $r_{21}, r_{31}, r_{41}, r_{32}, r_{34}, r_{43}$ | Adaptive filters in a Gram–Schmidt orthogonalizing network | Section 4.2 |
| **RP** | Response planning | Section 6.2 |
| $\mathrm{SVF}_1, \ldots, \mathrm{SVF}_{10}$ | Ten singular vector filters in nonlinear adaptive filter | Section 7.3 |
| $t$ | Tension signal in cortico-cerebellar-cortical circuit | Section 7.1 |
| $\mathbf{t} = [t_1, \ldots, t_{700}]^{\mathrm{T}}$ | Vector of 700 tension signals generated by 700 functional muscles | Section 2 |
| $\mathbf{\Delta t} = [\Delta t_1, \ldots, \Delta t_{700}]^{\mathrm{T}}$ | Vector of 700 first-differenced tension signals | Section 4.3 |
| $\mathbf{T}$ | $N$-dimensional vector of orthogonalized tension signals | Section 4.1 |
| $\mathbf{\Delta T} = [\Delta T_1, \ldots, \Delta T_N]^{\mathrm{T}}$ | Vector of $N$ orthogonalized first-differenced tension feature signals | Section 4.3 |
| $\mathbf{\Delta T}^* = [\Delta T_1{}^*, \ldots, \Delta T_N{}^*]^{\mathrm{T}}$ | Vector of $N$ orthogonalized first-differenced required tension feature signals | Section 4.3 |
| $\mathbf{\Delta T}$ feature extraction | Adaptive orthogonal first-differenced tension feature extraction network | Section 4.3 |
| $\mathbf{T_a}$ | $N$-vector of orthogonal target signals | Section 6.2 |
| $\hat{\mathbf{T}}_\mathbf{a} = \left[\hat{T}_{a1}, \ldots, \hat{T}_{aN}\right]^{\mathrm{T}}$ | Vector of $N$ orthogonalized predicted future values of target feature signals | Section 6.2 |
| **TDSG** | Task-dependent synergy generator | Section 4.3 |
| $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_{110}]^{\mathrm{T}}$ | Vector of 110 elemental movement signals | Section 2 |
| $\mathbf{\Delta}\boldsymbol{\theta}$ | Vector of first-differenced elemental movement signals | Section 3.3 |
| $t_m$ | Model output tension signal in cortico-cerebellar-cortical circuit | Section 7.1 |
| $u(k)$ | Input signal to system to be modeled and to nonlinear adaptive filter | Section 7.3 |
| $u_1(k), u_2(k), u_3(k), u_4(k)$ | Four interrelated signals | Section 4.2 |
| $U_1(k), U_2(k), U_3(k), U_4(k)$ | Four orthogonalized signals | Section 4.2 |
| $v_i(k), v_i^2(k), v_i^3(k)$ | The output, squared output, and cubed output signals from the $i$th singular vector filter in the nonlinear adaptive filter circuit | Section 7.3 |
| $V_i(k), V_i^2(k), V_i^3(k)$ | The orthogonalized output, squared output and cubed output signals from the $i$th singular vector filter in the nonlinear adaptive filter circuit | Section 7.3 |
| $\mathbf{v} = [v_1, \ldots, v_{700}]^{\mathrm{T}}$ | Vector of velocities of length changes of 700 functional muscles | Section 2.2 |
| $w_1, \ldots, w_{30}$ | 30 adaptive $w$-parameters in nonlinear adaptive filter | Section 7.3 |
| **WSG** | Wired-in synergy generator | Section 3.1 |
| $y(k)$ | Output from system to be modeled in nonlinear adaptive filter | Section 7.3 |
| $y_m(k)$ | Output from nonlinear adaptive filter model | Section 7.3 |
| ZOH | Zero order hold | Section A.1 |
| $z^{-\rho}$ | Time delay of $\rho$ sample intervals | Section A.1 |

## References


Abernethy B and Sparrow W A 1992 The rise and fall of dominant paradigms in motor behaviour research *Approaches to the Study of Motor Control and Learning* ed J J Summers (Amsterdam: Elsevier) pp 3–45

Baraduc P, Lang N, Rothwell J C and Wolpert D M 2004 Consolidation of dynamic motor learning is not disrupted by rTMS of primary motor cortex *Current Biol.* **14** 252–6

Barlow J S 2002 *The Cerebellum and Adaptive Control* (Cambridge: Cambridge University Press)

Bendat J 1990 *Nonlinear System Analysis and Identification from Random Data* (New York: Wiley)

Bendat J 1998 *Nonlinear System Techniques and Applications* (New York: Wiley)

Berardelli A, Hallett M, Rothwell J C, Agostino R, Manfredi M, Thompson P D and Marsden C D 1996 Single-joint rapid arm movements in normal subjects and in patients with motor disorders (review article) *Brain* **119** 661–74

Bernstein N A 1967 *The Co-ordination and Regulation of Movements* (London: Pergamon)

Bhushan N and Shadmehr R 1999 Computational nature of human adaptive control during learning of reaching movements in force fields *Biol. Cybern.* **81** 39–60

Bizzi E and Mussa-Ivaldi F A 2000 Toward a neurobiology of coordinate transformation *The New Cognitive Neurosciences* ed M S Gazzaniga (Cambridge, MA: MIT Press) pp 489–500

Box G E P and Jenkins G M 1976 *Time Series Analysis: Forecasting and Control* (San Francisco, CA: Holden-Day)

Braitenberg V 1986 Two views of the cerebral cortex *Brain Theory* ed G Palm and A Aertsen (Berlin: Springer) pp 81–96

Brashers-Krug T, Shadmehr R and Bizzi E 1996 Consolidation of human motor memory *Nature* **382** 252–5

Chronicle E P and Glover J 2003 A ticklish question: does magnetic stimulation of the primary motor cortex give rise to an 'efference copy'? *Cortex* **39** 105–10

Cichocki A and Amari S-I 2002 *Adaptive Blind Signal and Image Processing* (New York: Wiley)

Craik K J W 1947 Theory of the human operator in control systems: I. The operator as an engineering system *Br. J. Psychol.* **38** 56–61

Craik K J W 1948 Theory of human operator in control systems: II. Man as an element in a control system *Br. J. Psychol.* **38** 142–8

Cullen K E and Roy J E 2004 Signal processing in the vestibular system during active versus passive head movements *J. Neurophysiol.* **91** 1919–33

Davidson P R, Andreae J H and Sirisena H R 2002 Simulating closed- and open-loop voluntary movement: a nonlinear control systems approach *IEEE Trans. Biomed. Eng.* **49** 1242–52

Davidson P R, Jones R D, Sirisena R H and Andreae J H 2000 Detection of adaptive inverse models in the human motor system *Hum. Mov. Sci.* **19** 761–95

Davidson P R and Wolpert D M 2003 Motor learning and prediction in a variable environment *Curr. Opin. Neurobiol.* **13** 1–6

Davidson P R and Wolpert D M 2004 Internal models underlying grasp can be additively combined *Exp. Brain Res.* **155** 334–40

DeCharms R C and Zador A 2000 Neural representation and cortical code *Ann. Rev. Neurosci.* **23** 613–47

Desmurget M and Grafton S 2000 Forward modeling allows feedback control for fast reaching movements *Trends Cogn. Sci.* **4** 423–31

Douglas R J and Martin K A C 2004 Neuronal circuits of the neocortex *Ann. Rev. Neurosci.* **27** 419–51

Doya K, Kimura H and Kawato M 2001 Neural mechanisms of learning and control *IEEE Control Syst. Mag.* **21** 42–54

Flanagan J R, Nakano E, Imamizu H, Osu R, Yoshioka T and Kawato M 1999 Composition and decomposition of internal models in motor learning under altered kinematic and dynamic environments *J. Neurosci.* **19** rapid communication 34, 1–5

Flash T and Hogan N 1985 The coordination of arm movements: an experimentally confirmed mathematical model *J. Neurosci.* **5** 1688–703

Flash T and Sejnowski T 2001 Computational approaches to motor control *Curr. Opin. Neurobiol.* **11** 655–62

Gandevia S C 1987 Roles of perceived voluntary motor commands in motor control *Trends Neurosci.* **10** 81–5

Gandevia S C 2001 Spinal and supraspinal factors in human muscle fatigue *Physiol. Rev.* **81** 1725–89

Ghahramani Z and Wolpert D M 1997 Modular decomposition in motor learning *Nature* **386** 392–5

Ghez C and Thach T 2000 The cerebellum *Principles of Neural Science* 4th edn ed E R Kandel, J H Schwartz and T M Jessell (New York: McGraw-Hill) pp 832–52

Ghous A and Neilson P D 2002 Evidence for internal representation of a static nonlinearity in a visual tracking task *Hum. Mov. Sci.* **21** 847–79

Goodbody S J and Wolpert D M 1998 Temporal and amplitude generalization in motor learning *J. Neurophysiol.* **79** 1825–38

Goodwin G C and Sin K S 1984 *Adaptive Filtering Prediction and Control* (Englewood Cliffs, NJ: Prentice-Hall)

Gregory R L 1978 Eye and brain *The Psychology of Seeing* 3rd edn (New York: McGraw-Hill)

Gross J, Timmermann L, Kujala J, Dirks M, Schmitz F, Salmelin R and Schnitzler A 2002 The neural basis of intermittent motor control in humans *Proc. Natl. Acad. Sci. USA* **4** 2299–302

Harris C M and Wolpert D M 1998 Signal-dependent noise determines motor planning *Nature* **20** 780–4

Haykin S 1986 *Adaptive Filter Theory* (Englewood Cliffs, NJ: Prentice-Hall)

Haruno M, Wolpert D M and Kawato M 2001 Mosaic model for sensorimotor learning and control *Neural Comput.* **13** 2201–20

Isidori A 1995 *Nonlinear Control Systems* 3rd edn (London: Springer)

Ito M 1984 *The Cerebellum and Neural Control* (New York: Raven Press)

Jacobs R A, Jordan M I, Nowlan S J and Hinton G E 1991 Adaptive mixture of local experts *Neural Comput.* **3** 79–87

Jansen-Osmann P, Richter S, Schinauer T, Fuchs P and Kalveram K-T 2005 Adaptation to separate kinematic and dynamic transformations in children and adults *Motor Control* **9** 197–212

Jeannerod M 1981 Intersegmental coordination during reaching at natural visual objects *Attention and Performance IX* ed J Long and A Baddeley (Hillsdale, NJ: Erlbaum) pp 153–69

Jeannerod M 1999 Visuomotor channels. Their integration in goal-directed prehension *Hum. Mov. Sci.* **18** 201–18

Jordan M I and Wolpert D M 1999 Computational motor control *The Cognitive Neurosciences* ed M Gazzaniga (Cambridge, MA: MIT Press)

Jurdjevic V 1997 *Geometric Control Theory* (Cambridge: Cambridge University Press)

Kandel E R, Jessell T M and Sanes J R 2000 Sensory experience and the fine-tuning of synaptic connections *Principles of Neural Science* 4th edn ed E R Kandel, J H Schwartz and T M Jessell (New York: McGraw-Hill) pp 1115–30

Karniel A and Mussa-Ivaldi F A 2002 Does the motor control system use multiple models and context switching to cope with a variable environment? *Exp. Brain Res.* **143** 520–4

Kawato M 1999 Internal models for motor control and trajectory planning *Curr. Opin. Neurobiol.* **9** 718–27

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

Kenny D A 1979 *Correlation and Causality* (New York: Wiley)

Khalil K 1996 *Nonlinear Systems* 2nd edn (Upper Saddle River, NJ: Prentice-Hall)

Krakauer J and Ghez C 2000 Voluntary movement *Principles of Neural Science* 4th edn ed E R Kandel, J H Schwartz and T M Jessell (New York: McGraw-Hill) pp 756–81

Krakauer J W, Ghilardi M-F and Ghez C 1999 Independent learning of internal models for kinematic and dynamic control of reaching *Nature Neurosci.* **2** 1026–31
Ljung L 1987 *System Identification: Theory for the User* (Englewood Cliffs, NJ: Prentice-Hall)
Marmarelis V Z 2004 *Nonlinear Dynamic Modeling of Physiological Systems* (Hoboken, NJ: Wiley)
Marois R and Ivanoff J 2005 Capacity limits of information processing in the brain *Trends Cogn. Sci.* **9** 296–305
Marr D C and Poggio T 1977 From understanding computation to understanding neural circuitry *Neurosci. Res. Prog. Bull.* **15** 470–88
Matthews P B C 1972 *Mammalian Muscle Receptors and Their Central Actions* (London: Arnold)
McCloskey D I 1981 Corollary discharges: motor commands and perception *Handbook of Physiology: Sec. 1. The Nervous System: Vol. II. Motor Control, Part 2* ed V B Brooks (Bethesda, MD: American Physiological Society) pp 1415–47
McRuer D T and Krendel E S 1974 *Mathematical Models of Human Pilot Behavior (AGARDograph No. 188)* (Neuilly sur Seine, France: North Atlantic Treaty Organization, Advisory Group for Aerospace Research and Development)
Merzenich M M and Jenkins W M 1993 Reorganization of cortical representations of the hand following alterations of skin inputs induced by nerve injury, skin island transfers, and experience *J. Hand Therapy* **6** 89–104
Meyer D E, Abrams R A, Kornblum S, Wright C E and Smith J E K 1988 Optimality in human motor performance: ideal control of rapid aimed movements *Psychol. Rev.* **95** 340–70
Meyer D E, Smith J E K, Kornblum S, Abrams R A and Wright C E 1990 Speed-accuracy tradeoffs in aimed movements: toward a theory of rapid voluntary action *Attention and Performance XIII. Motor Representation and Control* ed M Jeannerod (Hillsdale, NJ: Erlbaum) pp 173–226
Middleton F A and Strick P L 2000 Basal ganglia and cerebellar loops: motor and cognitive circuits *Brain Res. Rev.* **31** 236–50
Mitra S, Amazeen P G and Turvey M T 1998 Intermediate motor learning as decreasing active (dynamical) degrees of freedom *Hum. Mov. Sci.* **17** 17–65
Mountcastle V B 1978 An organizing principle of cerebral function: the unit module and the distributed system *The Mindful Brain* ed G M Edelman and V B Mountcastle (Cambridge, MA: MIT Press) pp 1–50
Mountcastle V B 1993 The columnar organization of the neocortex *Brain* **120** 701–22
Mountcastle V B 1998 *Perceptual Neuroscience: The Cerebral Cortex* (Cambridge, MA: Harvard)
Mussa-Ivaldi F A 1999 Modular features of motor control and learning *Curr. Opin. Neurobiol.* **9** 713–7
Mussa-Ivaldi F A and Giszter S F 1992 Vector field approximation: a computational paradigm for motor control and learning *Biol. Cybern.* **67** 491–500
Navon D, Gopher D, Chillag N and Spitz G 1984 On separability of and interference between tracking dimensions in dual-axis tracking *J. Motor Behav.* **16** 364–91
Neilson M D and Neilson P D 1985 Speech motor control and stuttering *Motor Memory and Control: The Otago Symposium (Dunedin, New Zealand, 1982)* ed D G Russell and B Abernethy (Dunedin, NZ: Human Performance Associates) pp 69–83
Neilson M D and Neilson P D 1987 Speech motor control and stuttering: a computational model of adaptive sensory-motor processing *Speech Commun.* **6** 325–33
Neilson P D 1982 Central processes involved in acquisition of motor skill *Neurosci. Lett. (Suppl. 8) Abstracts of the Australian Neuroscience Society, Second Annual Meeting (15–17 Feb., Sydney)*
Neilson P D 1993a Tonic stretch reflex in normal subjects and in cerebral palsy *Science and Practice in Clinical Neurology* ed S C Gandevia, D Burke and M Anthony (Cambridge: Cambridge University Press) pp 169–90
Neilson P D 1993b The problem of redundancy in movement control: the adaptive model theory approach *Psychol. Res.* **55** 99–106
Neilson P D and Lance J W 1978 Reflex transmission characteristics during voluntary activity in normal man and in patients with movement disorders *Cerebral Motor Control In Man: Long Loop Mechanisms. Progress in Clinical Neurophysiology* ed J E Desmedt (Basel: Karger) vol 4 pp 263–99
Neilson P D and Neilson M D 1999a A neuroengineering solution to the optimal tracking problem *Hum. Mov. Sci.* **18** 155–83
Neilson P D and Neilson M D 1999b Self-organizing motor synergies: A neuroengineering approach to structure function relations in voluntary movement *Abstracts of International Conference—Progress in Motor Control II 'Structure-Function Relations in Voluntary Movement' (State College, PA, August 1999)* ed K Deutsch and M Latash (University Park, PA: The Pennsylvania State University) pp 114–5
Neilson P D and Neilson M D 2001 Neural mechanisms for control of multivariable, redundant, nonlinear musculoskeletal systems. *Proc. 40th IEEE Conf. on Decision and Control (Orlando, FL, Dec. 2001)* (Piscataway, NJ: IEEE Control Systems Society) pp 13–4
Neilson P D and Neilson M D 2002 Anisotropic tracking: evidence for automatic synergy formation in a bimanual task *Hum. Mov. Sci.* **21** 723–48
Neilson P D and Neilson M D 2004 A new view on visuomotor channels: the case of the disappearing dynamics *Hum. Mov. Sci.* **23** 257–83
Neilson P D and Neilson M D 2005 Motor maps and synergies *Hum. Mov. Sci.* at press
Neilson P D, Neilson M D and O'Dwyer N J 1985 Acquistion of motor skills in tracking tasks: learning internal models *Motor Memory and Control: The Otago Symposium (Dunedin, New Zealand, 1982)* ed D G Russell and B Abernethy (Dunedin, NZ: Human Performance Associates) pp 25–36
Neilson P D, Neilson M D and O'Dwyer N J 1988a Internal models and intermittency: a theoretical account of human tracking behavior *Biol. Cybern.* **58** 101–12
Neilson P D, Neilson M D and O'Dwyer N J 1992 Adaptive model theory: Application to disorders of motor control *Approaches to the Study of Motor Control and Learning* ed J J Summers (Amsterdam: Elsevier) pp 495–548
Neilson P D, Neilson M D and O'Dwyer N J 1993 What limits high-speed tracking performance? *Hum. Mov. Sci.* **12** 85–109
Neilson P D, Neilson M D and O'Dwyer N J 1995 Adaptive optimal control of human tracking *Motor Control and Sensory-Motor Integration: Issues and Directions* ed D J Glencross and J P Piek (Amsterdam: North-Holland) pp 97–140
Neilson P D, Neilson M D and O'Dwyer N J 1997 Adaptive model theory: central processing in acquisition of skill *Neurophysiology and Neuropsychology of Motor Development (Clinics in Developmental Medicine)* ed K J Connolly and H Forssberg (London: MacKeith Press) pp 346–70
Neilson P D, Neilson M D and O'Dwyer N J 1998 Evidence for rapid switching of sensory-motor models *Motor Control And Human Skill: A Multidisciplinary Approach* ed J P Piek (Champaign, IL: Human Kinetics) pp 105–126
Neilson P D and O'Dwyer N J 1984 Reproducibility and variability of speech muscle activity in athetoid dysarthria of cerebral palsy *J. Speech Hear. Res.* **27** 502–17
Neilson P D, O'Dwyer N J and Neilson M D 1988b Stochastic prediction in pursuit tracking: An experimental test of adaptive model theory *Biol. Cybern.* **58** 113–22
Newell K M, Broderick M P, Deutsch K M and Slifkin A B 2003 Task goals and change in dynamical degrees of freedom with motor learning *J. Exp. Psychol.: Hum. Perception Performance* **29** 379–87

Nijmeijer H and van der Shaft A 1996 *Nonlinear Dynamic Control Systems* (New York: Springer)
Noble B and Daniel J W 1988 *Applied Linear Algebra* 3rd edn (Englewood Cliffs, NJ: Prentice-Hall)
O'Dwyer N J and Neilson P D 1998 Motor output variability: variation with average level and range of force *Aust. J. Psychol.* **50** (Suppl.) 8
Olsen S M 2001 Adaptive model theory: A simulation study of speed-accuracy tradeoff in the execution of reaching movements *Master of Engineering Cybernetics Thesis* Faculty of Electrical Engineering and Telecommunication, Norwegian University of Science and Technology, Norway and School of Electrical Engineering and Telecommunications, University of New South Wales, Australia
Oytam Y, Neilson P D and O'Dwyer N J 1998 Characteristics of coordination and performance in visuo-motor coupling *Aust. J. Psychol.* **50** 9
Oytam Y, Neilson P D and O'Dwyer N J 2005 Degrees of freedom and motor planning in purposive movement *Hum. Mov. Sci.* at press
Plamondon R and Alimi A M 1997 Speed/accuracy tradeoffs in target directed movements *Behav. Brain Sci.* **20** 279–349
Sainburg R L, Ghez C and Kalakainis D 1999 Intersegmental dynamics are controlled by sequential anticipatory, error correction, and postural mechanisms *J. Neurophysiol.* **81** 1045–55
Salenius S and Hari R 2003 Synchronous cortical oscillatory activity during motor action *Curr. Opin. Neurobiol.* **13** 678–84
Sardesai M, Figge C, Bodner M, Crosby M, Hansen J and Quillfeldt J A *et al* 2001 Reliable short-term memory in the trion model: Towards a cortical language and grammar *Biol. Cybern.* **84** 173–82
Schmidt R A and Lee T D 1999 *Motor Control and Learning: A Behavioral Approach* (Champaign, IL: Human Kinetics)
Scholz J P and Schöner G 1999 The uncontrolled manifold concept: identifying control variables for a functional task *Exp. Brain Res.* **126** 289–306
Shadmehr R and Brashers-Krug T 1997 Functional stages in the formation of human long-term motor memory *J. Neurosci.* **17** 409–19
Shaw G L and Silverman D J 1988 Simulations of the trion model and the search for the code of higher cortical processing *Computer Simulation in Brain Science* ed R M J Cotterill (Cambridge: Cambridge University Press) pp 189–209
Siegelbaum S A, Schwartz J A and Kandel E R 2000 Modulation of synaptic transmission: second messengers *Principles of Neural Science* 4th edn ed E R Kandel, J H Schwartz and T M Jessell (New York: McGraw-Hill) pp 229–42
Slotine J-J E and Li W 1991 *Applied Nonlinear Control* (Englewood Cliffs, NJ: Prentice-Hall)
Smeets J B and Brenner E 1999 A new view on grasping *Motor Control* **3** 237–71
Sohn Y H, Dang N and Hallett M 2003 Suppression of corticospinal excitability during negative motor imagery *J. Neurophysiol.* **90** 2303–9
Stark L 1968 *Neurological Control Systems. Studies in Bioengineering* (New York: Plenum)
Stinear C M and Byblow W D 2003 Motor imagery of phasic thumb abduction temporally and spatially modulates corticospinal excitability *Clin. Neurophysiol.* **114** 909–14
Stinear C M and Byblow W D 2004 Modulation of corticospinal excitability and intracortical inhibition during motor imagery is task-dependent *Exp. Brain Res.* **157** 351–8
Taub E and Berman A J 1968 Movement and learning in the absence of feedback *The Neuropsychology of Spatially Oriented Behavior* ed S J Freedman (Homewood, IL: Dorsey) pp 173–92
Taylor J L and Gandevia S C 2001 Transcranial magnetic stimulation and human muscle fatigue *Muscle Nerve* **24** 18–29
Todorov E and Jordan M I 2002 Optimal feedback control as a theory of motor coordination *Nature Neurosci.* **5** 1226–35
Tong C, Wolpert D M and Flanagan J R 2002 Kinematics and dynamics are not represented independently in motor working memory: evidence from an interference study *J. Neurosci.* **22** 1108–13
Uno Y, Kawato M and Suzuki R 1989 Formation and control of optimal trajectory in human multijoint arm movement—minimum torquechange model *Biol. Cybern.* **61** 89–101
Vallbo Å B and Wessberg J 1993 Organization of motor output in slow finger movements in man *J. Physiol.* **469** 673–91
van Beers R J, Haggard P and Wolpert D M 2004 The role of execution noise in movement variability *J. Neurophysiol.* **91** 1050–63
Volpe J J 1987 *Neurology of the Newborn* 2nd edn (Philadelphia: Saunders)
Volpe J J, Herscovitch P, Perlman J M, Kreusser K L and Raichle M E 1985 Positron emission tomography in the asphyxiated term newborn: parasagittal impairment of cerebral blood flow *Ann. Neurol.* **17** 287–96
von Holst E and Mittelstaedt H 1950/1973 The reafference principle. Interaction between the central nervous system and the periphery *Selected Papers of Erich von Holst: The Behavioural Physiology of Animals and Man* vol 1 (London: Methuen) pp 139–73 (Transl. from German by R Martin, originally published 1950)
von Seelen W, Mallot H A, Krone G and Dinse H 1986 On information processing in the cat's visual cortex *Brain Theory* (Berlin: Springer) pp 49–79
Weindling A M, Rochefort M J, Calvert S A, Fok T-F and Wilkinson A 1985 Development of cerebral palsy after ultrasonographic detection of periventricular cysts in the newborn *Dev. Med. Child Neurol.* **27** 800–6
Westwick D T and Kearney R E 2003 *Identification of Nonlinear Physiological Systems* (Piscataway, NJ: IEEE)
Wickens C D 1984 Processing resources in attention *Varieties of Attention* ed R Parasuraman and D R Davies (Orlando, FL: Academic) pp 63–102
Widrow B 1970 Adaptive filters *Aspects of Network System Theory* ed R E Kalman and N De Claris (New York: Holt, Rinehart and Winston) pp 563–87
Widrow B and Stearns S D 1985 *Adaptive Signal Processing* (Englewood Cliffs, NJ: Prentice-Hall)
Widrow B and Walach E 1996 *Adaptive Inverse Control* (Upper Saddle River, NJ: Prentice-Hall)
Wigmore V, Tong C and Flanagan J R 2002 Visuomotor rotations of varying size and direction compete for a single internal model in motor working memory *J. Exp. Psychol.: Hum. Perception Performance* **28** 447–57
Williams K J and Neilson P D 1998 Studies in hand eye coordination: finding proof for synergy formation *Aust. J. Psychol.* **50** 12
Wolpert D M, Doya K and Kawato M 2003 A unifying computational framework for motor control and social interaction *Phil. Trans. R Soc.* B **358** 593–602
Wolpert D M, Ghahramani Z and Flanagan J R 2001 Perspectives and problems in motor learning *Trends Cogn. Sci.* **5** 487–94
Wolpert D M and Kawato M 1998 Multiple paired forward and inverse models for motor control *Neural Netw.* **11** 1317–29
Wolpert D M, Miall R C and Kawato M 1998 Internal models in the cerebellum *Trends Cogn. Sci.* **9** 718–27