Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Physiology - Paris 97 (2003) 601–612
Journal of Physiology Paris
www.elsevier.com/locate/jphysparis

Recurrent neural networks of integrate-and-fire cells simulating short-term memory and wrist movement tasks derived from continuous dynamic networks

Marc A. Maier [a], Larry E. Shupe [b], Eberhard E. Fetz [b,*]

[a] University Paris-VII and INSERM U483, 9 Quai Saint Bernard, 75005 Paris, France
[b] Department of Physiology and Biophysics, Washington National Primate Research Center, University of Washington, Seattle, WA 98195-7290, USA

Abstract

Dynamic recurrent neural networks composed of units with continuous activation functions provide a powerful tool for simulating a wide range of behaviors, since the requisite interconnections can be readily derived by gradient descent methods. However, it is not clear whether more realistic integrate-and-fire cells with comparable connection weights would perform the same functions. We therefore investigated methods to convert dynamic recurrent neural networks of continuous units into networks with integrate-and-fire cells. The transforms were tested on two recurrent networks derived by backpropagation. The first simulates a short-term memory task with units that mimic neural activity observed in cortex of monkeys performing instructed delay tasks. The network utilizes recurrent connections to generate sustained activity that codes the remembered value of a transient cue. The second network simulates patterns of neural activity observed in monkeys performing a step-tracking task with flexion/extension wrist movements. This more complicated network provides a working model of the interactions between multiple spinal and supraspinal centers controlling motoneurons.

Our conversion algorithm replaced each continuous unit with multiple integrate-and-fire cells that interact through delayed "synaptic potentials". Successful transformation depends on obtaining an appropriate fit between the activation function of the continuous units and the input–output relation of the spiking cells. This fit can be achieved by adapting the parameters of the synaptic potentials to replicate the input–output behavior of a standard sigmoidal activation function (shown for the short-term memory network). Alternatively, a customized activation function can be derived from the input–output relation of the spiking cells for a chosen set of parameters (demonstrated for the wrist flexion/extension network). In both cases the resulting networks of spiking cells exhibited activity that replicated the activity of corresponding continuous units. This confirms that the network solutions obtained through backpropagation apply to spiking networks and provides a useful method for deriving recurrent spiking networks performing a wide range of functions.

Keywords: Neural networks; Recurrent; Dynamic; Integrate-and-fire; Short-term memory; Wrist movement; Premotor neurons

1. Introduction

Many artificial neural networks employ units that have continuous activity to represent the firing rate of neurons and have sigmoidal activation functions to represent the units' input–output transform. These simplified representations allow powerful gradient descent algorithms such as backpropagation [2,28] to be used to derive dynamic recurrent networks that simulate many interesting biological behaviors [7,9]. A major goal of such networks is to capture the neural mechanisms that are mediated by modulated firing rates. However, real biological neurons integrate synaptic input to threshold and fire discrete spike trains. It is not clear that the network solutions obtained with "continuous" units accurately describe the behavior of networks composed of interconnected integrate-and-fire spiking cells. If so, a conversion procedure would be very useful, since there are no efficient training algorithms for deriving recurrent dynamic networks with discretely spiking cells. Algorithms for Hebbian learning with spike-time dependent

* Corresponding author. Tel.: +1-206-543-4839; fax: +1-206-685-8606.
E-mail address: fetz@u.washington.edu (E.E. Fetz).

0928-4257/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jphysparis.2004.01.017

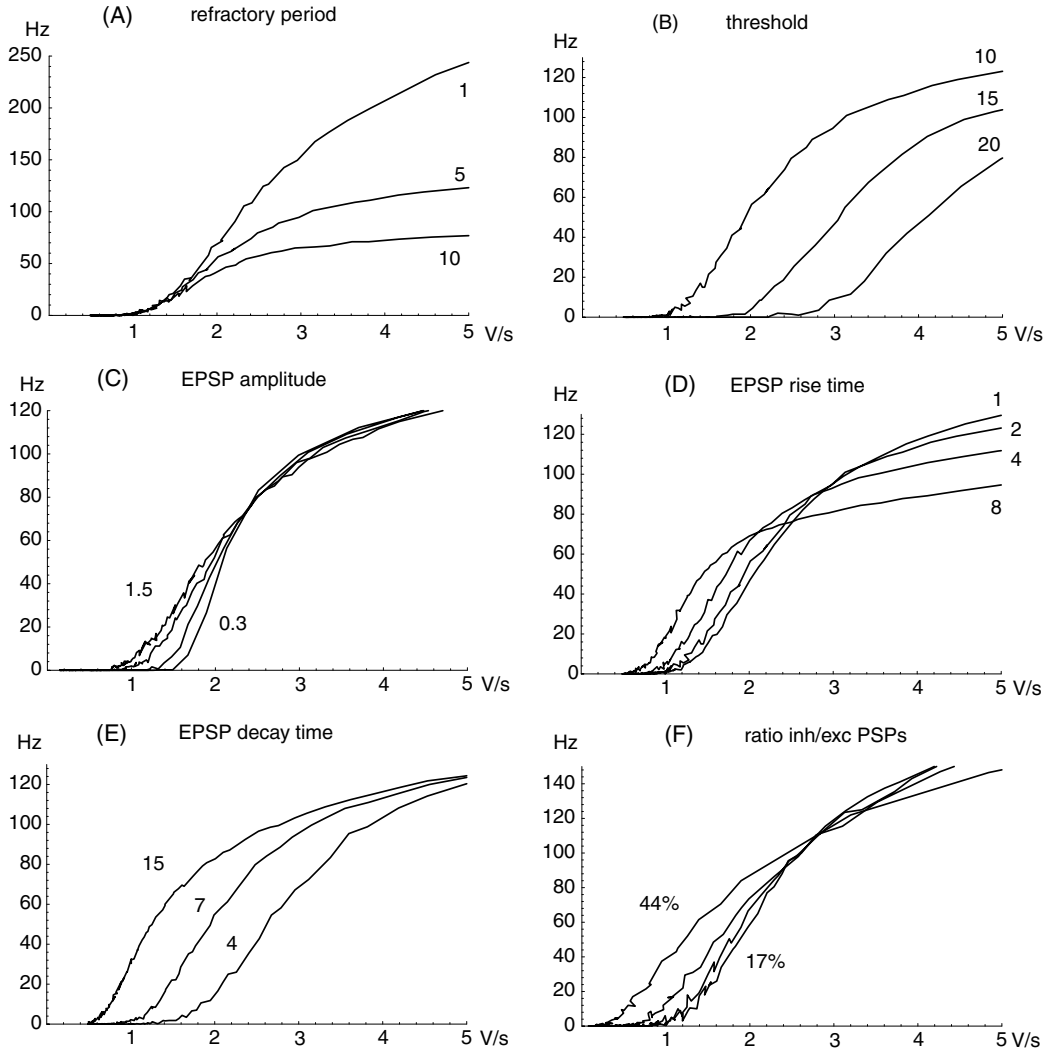

Fig. 1. Activation functions of integrate-and-fire spiking cells. Steady-state output frequency (Hz) as a function of synaptic input (V/s, 1V corresponds to 1000 one-mV EPSPs/s). Variation of (A) refractory period (1, 5, 10 ms): limits the maximum firing rate. This parameter can be used to fit the activation function to experimentally obtained *F–I* curves. (B) Threshold (10, 15, 20 mV): lowers the firing rate for a given input. (C) EPSP amplitude (0.3, 0.5, 1, 1.5 mV): changes the maximal slope. (The activation function does not shift towards higher inputs since the *x*-axis is normalized to the overall input.) (D) EPSP rise time (1, 2, 4, 8 ms): lowers the maximal firing rate and increases output before saturation starts. (E) EPSP decay time (4, 7, 15 ms): shifts the activation function towards lower input values. (F) Ratio between excitatory and inhibitory input (IPSPs in 17%, 23%, 33% and 44% of the total input): produces a shift somewhat similar to D and E.

plasticity are being developed [13,26], but these are less generally applicable to arbitrary dynamic input–output transformations. We therefore investigated whether networks of continuous units derived by backpropagation could be converted to equivalent networks of spiking cells performing the same behaviors. Of course the backpropagation algorithm is biologically unrealistic as a learning mechanism [6], but this is entirely irrelevant to its main purpose: to search efficiently for the network architectures that generate the required behaviors. Biologically plausible error-driven algorithms equivalent to backpropagation [21] and reward-based learning [19] have been developed; these find similar network solutions, albeit more slowly.

We were particularly interested in dynamic recurrent networks that simulate firing patterns observed in diverse neuronal populations in the behaving monkey. We show results for two well-documented tasks: a short-term memory task, and a wrist flexion/extension step-tracking task. The short-term memory task involves an arbitrary transient cue whose value needs to be sampled and held in memory during a variable delay period (“instructed delay”). This paradigm has been employed in numerous behavioral studies, and the task-related activity of cortical neurons has been documented in several frontal, parietal and temporal areas [1,12,14,24]. The step-tracking task is a visually guided wrist flexion/extension task for which the activity of various popu-

lations of neurons in the primate motor system has been amply documented [3,10,11,17,20,22]. We use these two examples, denoted 'short-term memory' and 'step-tracking' task, to demonstrate the feasibility of this conversion. Results indicate that under appropriate conditions the recurrent networks found by backpropagation provide valid descriptions of networks implemented with 'integrate-and-fire' cells.

2. Materials and methods

2.1. Properties of spiking cells

Our spiking cell is a single-compartment integrate-and-fire cell without conductances or electrotonic properties. It incorporates a single potential, representing the subthreshold membrane potential of a biological neuron, whose time course depends on the sum of incoming excitatory and inhibitory postsynaptic potentials (EPSP and IPSP, see Fig. 2C), and (optionally) a voltage ramp. The spiking cell integrates EPSPs and IPSPs to threshold (10 mV) for firing an action potential (whose trajectory is not modeled). The action potential produces PSPs with specified delays in all the postsynaptic target cells. During a post-spike refractory period (typically around 3 ms) any incoming PSPs are neglected and the potential is set to the resting potential (0 mV). PSPs are modelled by triangular waveforms with a given amplitude (±1 mV for EPSPs and IPSPs respectively), and rise- and fall-times [25].

2.2. Mapping procedure between continuous and spiking model

2.2.1. The input–output function of the continuous units

The backpropagation algorithm requires smooth differentiable activation functions to represent the input–output relation of continuous units. We found that the input–output function of spiking cells differs significantly from the standard symmetric sigmoidal activation function, and therefore we customized parameters to facilitate the subsequent transformation. For the *short-term memory* model we used a standard sigmoidal activation function and fitted the PSP parameters to replicate this input–output function in spiking cells. The input to a continuous unit (x) consists of the activation of all other units connected to this unit times their synaptic weights. The output of the continuous unit is given by $g(x)$, determined by the *standard sigmoid function*:

$$g(x) = \frac{1}{1 + e^{(s-x)*T}}$$

where $x =$ input, and constants $s =$ shift, and $T =$ temperature.

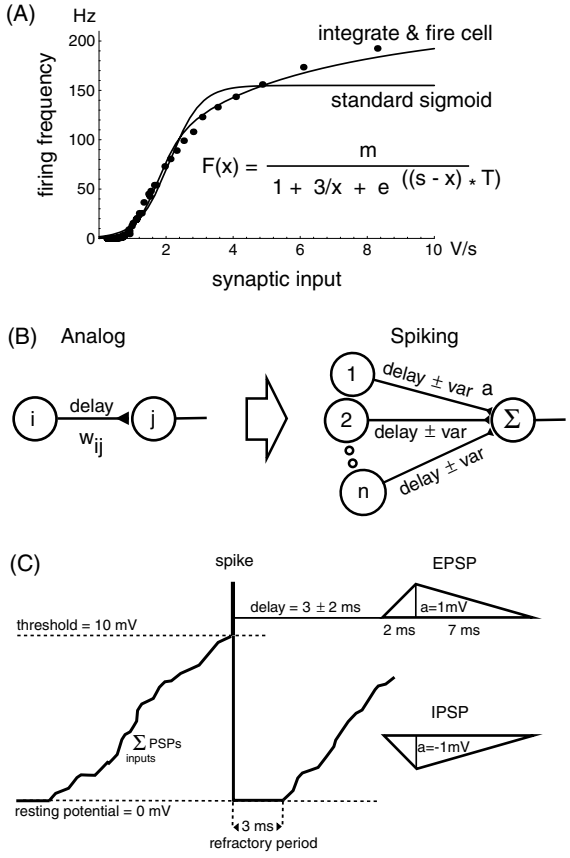

Fig. 2. Conversion from analog to spiking network. (A) Fit of continuous activation function to the input–output relation of the spiking cell. Dots: empirical values from the spiking cell; lines: sigmoids for continuous function. The maximal firing rate, m, is set to 1.0 in the activation function for training the analog network, then adjusted for in the conversion to the spiking network. (B) Transformation of analog to spiking network involves replacing each single continuous unit by a pool of spiking cells according to its maximal output weight. (C) "Membrane potential" of integrate-and-fire cell is the sum of synaptic inputs, and a spike causes delayed EPSPs or IPSPs in target cells.

For the wrist *step-tracking* model we derived a custom continuous activation function that fit the input–output relation of the spiking cell with fixed PSP rise- and fall-times, i.e. 2 and 7 ms, respectively. We used the following *custom sigmoidal activation function* for training the step-tracking analog network of continuous units (Fig. 2A).

$$\text{Custom sigmoid: } f(x) = \frac{1}{1 + (3/x) + e^{(s-x)*T}}$$

2.2.2. Characterization of the input–output function of the spiking cells

To determine the input–output relation of a single spiking cell we used 10 independent and exponentially distributed spike trains to provide excitation (or inhibition) to a test spiking cell. The EPSPs in the test cell

had the following default properties: 1 mV amplitude (−1 mV for IPSPs), 2 ms rise time, 7 ms decay time. The test cell had the following default parameters: 5 ms refractory period, 10 mV threshold and no voltage ramp. Activation curves were calculated by varying the firing rate of input spike trains in steps and plotting the resulting steady-state output rate of the test cell. Fig. 1 shows examples of the input–output curves as a function of the parameters of the test cell and its 'synaptic' inputs. In contrast to the standard sigmoidal function of a continuous unit, the activation function of spiking cells is variable and depends on the refractory period (Fig. 1A) and threshold (Fig. 1B) of the spiking cell. The activation of the test cell also varied as a function of PSP amplitude (Fig. 1C), rise time (Fig. 1D), and fall time (Fig. 1E), as well as of the ratio between inhibition and excitation (Fig. 1F).

In order to transform recurrent networks trained with the standard sigmoidal activation function, we sought spiking cell parameters that would result in a similarly sigmoidal activation curve. This involved creating activation curves using various values for the spiking cells' refractory period and voltage ramp input in combination with the connection parameters of PSP delay time, rise time, and decay time. Using the appropriate parameters each continuous unit was replaced with a set of spiking cells, as described below. Parameters that gave a symmetric sigmoid were 7 ms refractory period, 0.02 mV/ms voltage ramp, 5 ms delay, 5 ms rise, and 7 ms decay. The thresholds of the spiking cells were set at 10 mV. The long refractory period was necessary to produce a relatively flat upper asymptote (Fig. 1A); deeming this unrealistic we developed the alternate approach of deriving a custom sigmoid for more realistic parameters.

2.2.3. Transformation of the analog wrist step-tracking network to a spiking network

The following method was used to transform the trained analog wrist step-tracking network to an equivalent spiking network with identical topology, input activations, connection delays and similar weights. A single continuous unit is typically envisioned as representing mean activity of a group of neurons, so we transformed each single continuous unit into a pool of multiple spiking cells, each of which receives from and projects to all the transformed cells that the single continuous unit was connected to. The number of spiking cells within a pool is a function of the weight of the continuous unit. When the continuous unit is connected to several continuous units with different weights, the transformation produces a maximal number of spiking cells (N) determined by the strongest weight, and for lesser weights connects an appropriate subset of these cells (n) to the other target cells according to the weight of the given projection. Thus the continuous weight is matched primarily by varying the number of connected spiking cells and then fine-tuned by scaling their PSP amplitude. For weight w, the number of connections required between the pool of presynaptic spiking cells and each postsynaptic cell is determined by $n(w) = |w/(a^*m)|$ where a is the chosen PSP height (in microvolts) used in calculating the activation curve, and m is the maximal output firing frequency of a spiking cell determined by the activation curve. $n(w)$ is converted to an integer n by rounding up if $n(w) < 1$ or $n(w)/Floor[n(w)] > 1.1$; otherwise $n(w)$ is rounded down to obtain the integer number of connections, n. The magnitude of the actual PSPs used in the spiking network is equal to the desired PSP height multiplied by $n(w)/n$. To replicate the time-varying activity of the continuous units it was necessary to disperse the arrival of PSPs, in order to prevent synchronous activation of the spiking cells. This was accomplished by deriving the connections to a postsynaptic cell from randomly chosen cells in the presynaptic pool, and assigning varying conduction delays around the mean delay for the corresponding connection in the analog network.

2.3. The analog networks

To obtain the dynamic recurrent networks of continuous units we used the temporal flow algorithm [27,28], a modified version of the backpropagation algorithm that incorporates time varying activation patterns (20 and 144 time steps for the short-term memory and the step-tracking network, respectively). The learning rate (ε) was limited to a range of 10^{-1} to 10^{-4}. Starting from a configuration with initially random weights, the short-term memory network was trained for 7000 cycles and the wrist step-tracking network for 10,000 cycles.

3. Results

3.1. Short-term memory network

The short-term memory task has been modeled with recurrent networks of continuous units without sign constraint by Zipser [29,30]. Our full analog model of the short-term memory network contains a total of 27 excitatory or inhibitory units plus a bias unit with constant activity (Fig. 3). The two inputs consist of a sample gate signal (i1) and a randomly varying input corresponding to the cue (i2). Their activity is fed to 12 excitatory (a1–a12) and 12 inhibitory hidden units (b1–b12), which in turn project to the output unit (o1). The excitatory hidden units are connected among themselves (upper left quadrant of weight matrix) and to the inhibitory hidden units (upper right quadrant). Inhibitory hidden units are not interconnected among them-

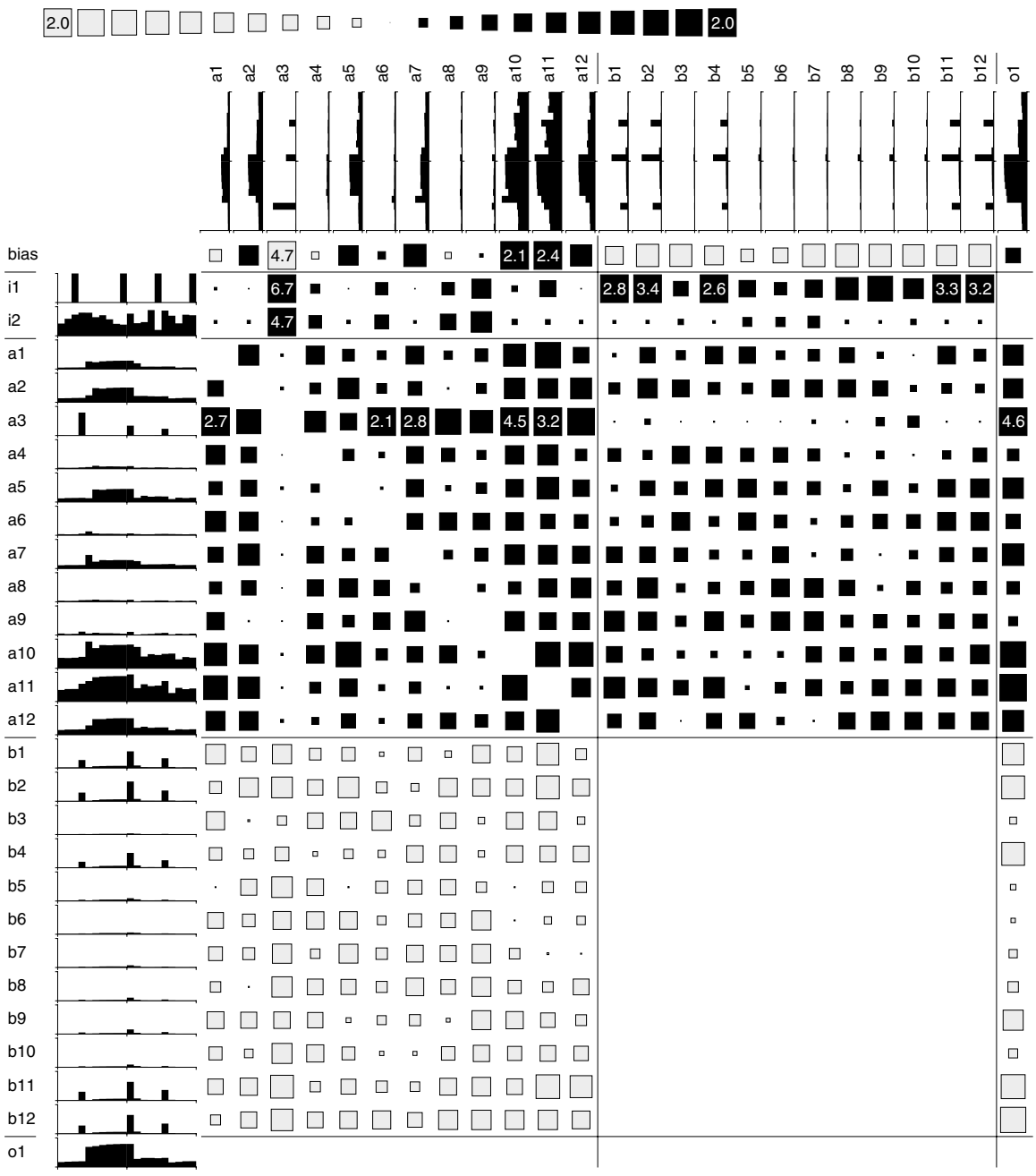

Fig. 3. Weight matrix and activity of units for the full analog short-term memory network. Names and activity of units are shown at left and along top. The bias unit has constant maximal activity. Input units: (i1: gate, i2 analog value to be remembered). Excitatory hidden units: (a1–a12). Inhibitory hidden units: (b1–b12). Output unit with short-term memory activity: (o1). The size of each square gives connection strength from row unit to column unit [in the range −2, 2, calibrated at the top, otherwise numerically]. Solid and open squares represent excitatory and inhibitory connections, respectively.

selves, but are connected to all excitatory hidden units (lower left quadrant). The output unit maintains a faithful representation of the analog value of the cue at the time of the sample gate signal. The final full network is larger than necessary because many hidden units exhibited negligible activity or weak connections and other hidden units resembled each other in activity and output weights [e.g., inhibitory hidden units]. To eliminate unnecessary and redundant units we implemented a weight decay algorithm that automatically reduced each weight by 0.001 for each of the first 5000 cycles; then extraneous units and weights were deleted and the resulting network trained for an additional 2000 cycles.

Fig. 4 (left and middle) shows the analog network reduced by the weight decay algorithm to a smaller essential configuration of only 8 continuous units with a

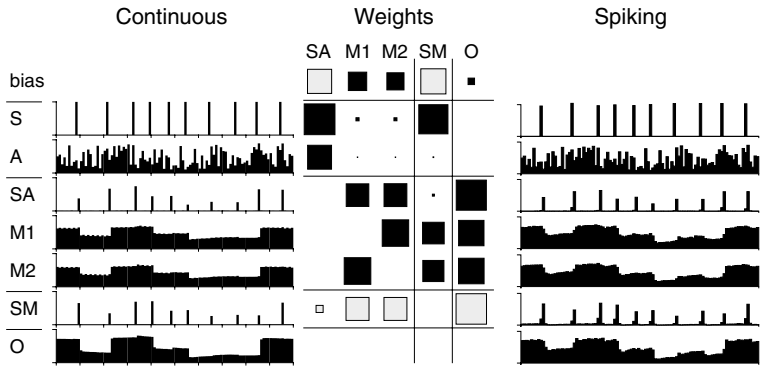

Fig. 4. *Left*: Activity of units for the reduced analog short-term memory network. Input units as in Fig. 3, renamed according to their function (S: sample gate, A: analog value to be remembered). Excitatory hidden units: (SA: sample of analog variable; M1, M2: 1st and 2nd memory unit). Inhibitory hidden unit: (SM: subtraction of memory value). Output unit with short-term memory activity: (O). *Middle*: Weight matrix corresponding to names and activity of units at left and names along top. The size of each square gives connection strength from row unit to column unit [in the range −6, 6]. *Right*: Average activity of the corresponding spiking cells after transformation from the reduced analog network (compiled from PSTHs of 1000 sweeps).

total of 29 weights among them. As in Fig. 3, the two inputs consist of a sample gate signal (S) and a randomly varying input corresponding to the cue or analog value (A). Three excitatory hidden units replace the 12 units in the full model. The first excitatory hidden unit (SA) carries a transient signal whose amplitude is proportional to the value of the cue at the time of the sample gate. Examination of the weights of the inputs to SA reveals that this activity is derived from the sum of the excitatory analog input, the excitatory gating input and the inhibitory constant bias input; the latter effectively clips the tonic activity provided by (A) leaving a transient value proportional to A. The value of (SA) is held in memory by two cross-connected hidden units (M1, M2) which also feed this value to the output unit (O). (Had we allowed self-connections, M1 and M2 could be replaced with a single self-connected unit.) The inhibitory unit (SM) subtracts the previously held value from the excitatory hidden units and the output. Thus, the network produces a constant output (O) that corresponds to the value of the input (A) at the time of the last sample gate (S).

The reduced short-term memory network with 8 units and 29 weights was transformed into a spiking network containing 39 integrate-and-fire cells and 1143 connections. Fig. 4 (right) shows the average corresponding activity of the spiking cells for the short-term memory network derived from the reduced analog network. The fact that the activity profiles of the analog and the spiking version were essentially identical confirms the success of the conversion algorithm.

3.2. The wrist step-tracking network

The much more complex analog model of the primate pre-motor system (with a total of 98 units and over 2000 weights) was designed to simulate the time-varying activity of multiple populations of neurons involved in performing an alternating wrist step-tracking task. Here we use it primarily as a biologically constrained benchmark test: we incorporated architectural constraints (including the sign of the connections) and constraints on the time-varying activity of input and representative output units. Monkeys performing this task transform a visual signal that indicates the target position (flexion or extension) into actual wrist flexion or extension. Accordingly, the model (Fig. 5) receives as input a target position and is required to provide as output the activity profiles of the biological motor units ("α" in Fig. 5) recorded in flexor or extensor muscles. This transformation is achieved within a network of four modules, each with specific connections within and among them. These modules correspond to cortical, rubral and segmental networks and to muscle afferents, and they are interconnected (with different conduction delays) in accordance with known anatomical pathways or simplifications of them.

The operation of components of the step-tracking network can be summarized as follows. The input to the network is provided to the cortical module and is represented by a step change in target position, alternating between flexion (during the first half of the 144 time steps) and extension (second half). Using the abbreviations in Figs. 6 and 7 the input is represented by both a sustained step (Sf) and a transient (Df) input to the cortical network for both flexion (Sf, Df) and extension (Se, De). A further input representing flexor (TFf) and extensor (TFe) wrist torque feedback, i.e. the afferent sensory feedback during correct task completion, is given to the respective spindle afferents (SP). The cortical and rubral modules consist of excitatory projection units (CM and RM) and local inhibitory (CL and RL) units. Within these modules, local recurrent connections among excitatory units and among inhibitory and

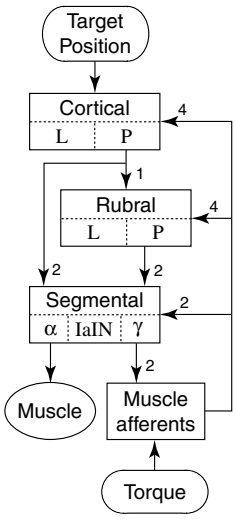

Fig. 5. Architecture of the step-tracking network. It comprises 4 different modules: cortical, rubral and segmental units and muscle afferents. The cortical and rubral modules consist of an equal number of excitatory projection units (P) and local inhibitory units (L) which are mutually interconnected (local recurrency). The target position input is relayed to the cortical module, whose projection units (P) forward their activity to the rubral and segmental population. Cortical and rubral projection units connect to all units on the segmental level, i.e. to alpha-(α) and gamma-motor (γ) units projecting to the muscle, and to Ia-inhibitory interneurons (IaIN). Muscle afferents, driven by a torque feedback delivered as input, project back to the segmental and supraspinal populations. The muscle activity is not modeled. A one step time delay applies to interactions between units of the same module and longer time delays between modules correspond to the number of time steps on projections.

excitatory units exist and represent local processing within the cortical and within the rubral module. The segmental module consists of flexor and extensor α-(MUf/e) and γ-motoneurons (GAf/e) and Ia-inhibitory interneurons (Iaf/e), all of which receive inputs from the cortical and rubral projection units. The spinal interconnections of these units are based on anatomical and physiological data [16]. The afferent module represents spindle afferent units (SPf/e), which are driven by γ-motoneurons (GAf/e) and the equivalent of applied torque feedback (TFf/e). Afferents feed back to the segmental level (i.e. to the homonymous α motor units and their corresponding Ia-inhibitory units) and supraspinal levels. The connections provide a coarse correspondence between units in the network and neurons in specific brain regions. We further constrained the network by incorporating representative profiles of physiological activity (target activity over 144 time steps) in four subpopulations of units. A summary table of these experimentally determined response classes can be found in [17]. Briefly, (i) Corticomotoneuronal cells are largely comprised of tonic and phasic-tonic cells [3]. Example units with these target profiles for the flexion period are to be found in target units (CMt1) and (CMt2) respectively, and in units (CMt3,4) for extension. (ii) Rubromotoneuronal cells fall mainly into three classes: phasic-tonic, phasic and unmodulated [20]. Example target units are: (RMt1, RMt2 and RMt5): (iii) Premotor dorsal root ganglion (DRG) afferents exclusively fall into three categories: tonic, phasic-tonic and phasic [11] and therefore all corresponding units in the network do have respective target activations (SPf1-3 for flexion, Spe1-3 for extension). (iv) the discharge patterns of single motor units (corresponding to the output units of the network) make up four classes: tonic, phasic-tonic, phasic and decrementing [22]. Accordingly, all MU-units are target units (MUf1-4 and MUe1-4).

Fig. 6 shows the activity of these units and the connection weights of a typical network solution. Emergent properties (in the following indicated in *italics*) primarily concern the weight space and the time-varying activity of hidden units, both of which will be compared to biological data. The analog network transforms the four input signals into the 8 types of experimentally observed [22] target response patterns of motor units at the output: tonic (MUf1), phasic-tonic (MUf2), decrementing (MUf3) and phasic (MUf4), generated for both flexor (MUf1-4) and extensor motor units (MUe1-4). Neither phasic-tonic nor decrementing patterns are represented as such in the network input. Correct profiles are also generated for premotor target units (Fig. 7): for cortical (CMt) and rubral (RMt) units, as well as for spindle afferents (SPf/e). Furthermore, the majority of hidden cortical (non-target) units, such as CM10, 13, 7 and others, take on *profiles of activity* akin to those observed experimentally.

These results show that the model and its particular architecture can generate appropriate activity profiles for ramp-and-hold step-tracking movements. In order to simulate this behavior, the model developed particular weights for particular groups of units, independently of the initial randomization. In the cortical module, most hidden CM units take on *unidirectional activity* (similar to the target CM units) and show *preferentially strong* excitatory connections to *multiple synergistic MUs* activated in the *same* movement phase (e.g. CM10, 13, 1, 8 for hidden units, and CMt1 and 2 for target units). This synergistic arrangement, an emergent property that allocates a specific weight configuration to hidden or target units with particular activity, holds in general for CM connections to units activated unidirectionally and in-phase, such as to LC units, RM units, GA units and to Ia-inhibitory units. In a similar way, *bi-directional* CM units are more *strongly* connected to other bidirectional units but develop more *diffuse and weaker* connections to MUs and other unidirectional units. In contrast, local inhibitory cortical units activated *unidirectionally* (LC) are preferentially connected to *out-of-phase* CM units (e.g. CL7). In the rubral module, all units show bidirectional activity except the unidirectional target units: local rubral units provide *strong*

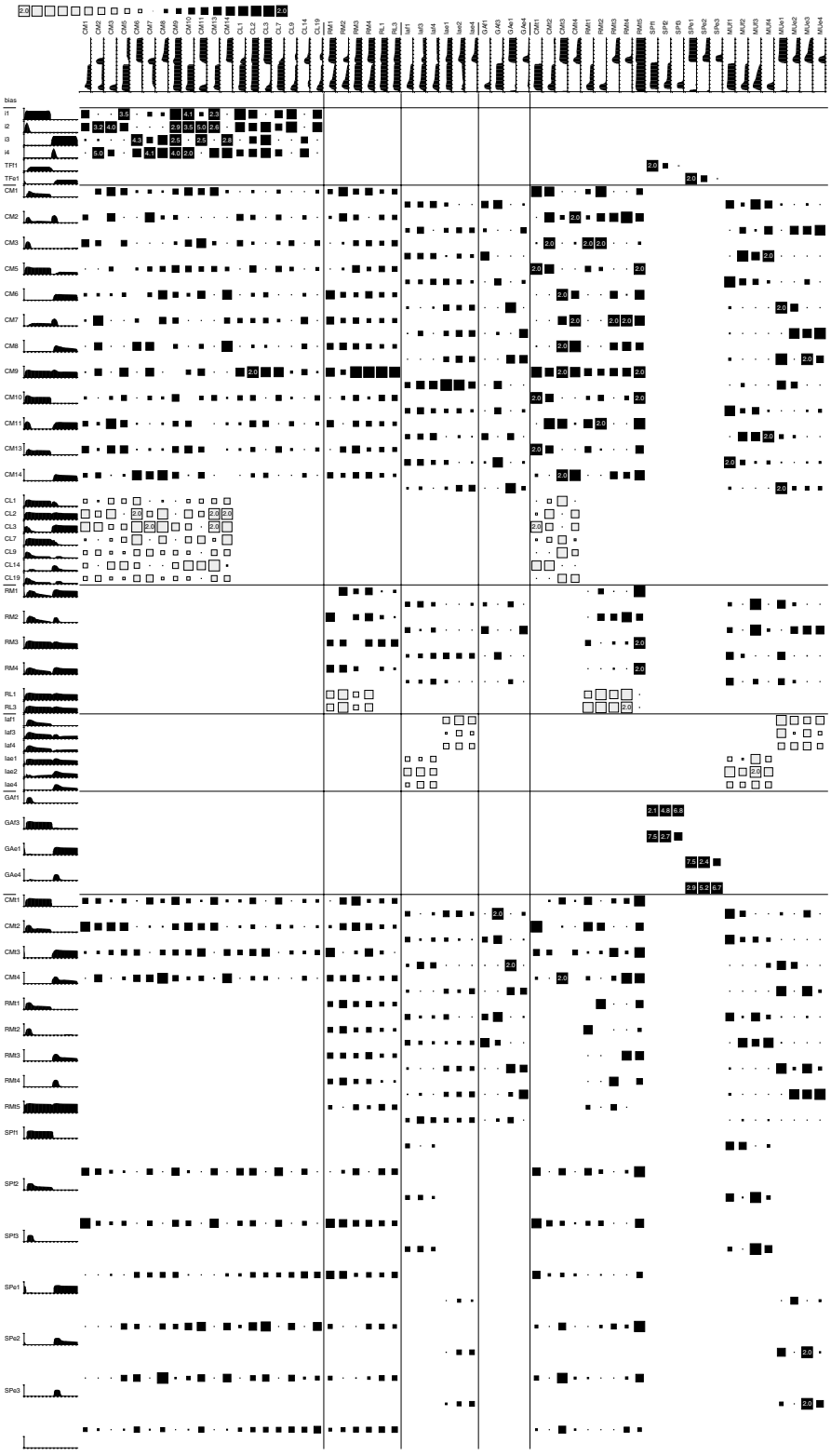

Fig. 6. Weight matrix for the analog network for the step-tracking flexion/extension of the wrist. First half of time axis: flexion, second half: extension Same conventions as in Fig. 4. For illustration purposes, we reduced the original 98 units to 64 (and 1600 weights) by eliminating units with negligible weights and by combining redundant units. Some rows of output weights are staggered for different latencies to different targets. The divergence of any unit to other units is given by its row of output weights and the convergence to a unit is given by the column of its input weights. Similarly, other connections and ratios can also be read off from this diagram. For cortical units the ratio of excitatory to inhibitory inputs ranged from 3:2 to 4:1. Abbreviations given in text.

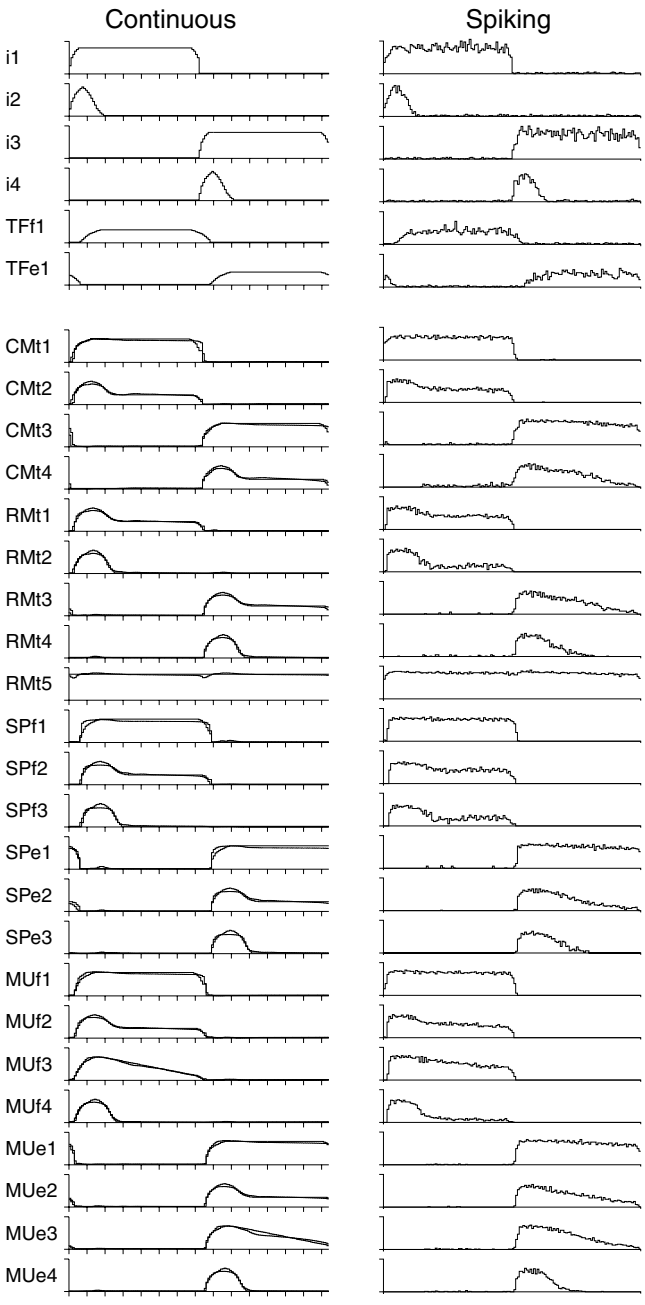

Fig. 7. Comparison between continuous unit and spiking cell activity for corresponding units in the step-tracking network. First half of time axis: flexion, second: extension. Histograms of spiking cells were derived from 30 repetitions. Abbreviations given in text.

inhibition to the unidirectional rubral target units. In the spinal module, Ia units are preferentially *unidirectional* and *strongly* inhibit the antagonist MUs, whereas the γ motor units take on strictly *unidirectional* activity, project *strongly* to the spindles (SP), which in turn project to the in-phase MUs, and provide a feedback to in-phase cortical and rubral units. In summary, the network solution demonstrates how motor units are driven by the combined activity of cortical, rubral, spinal and afferent units with specific activations and particular connections, such as the experimentally observed cortico- and rubromotoneuronal connections to specific groups of motoneurons. In particular, unidirectional CM, RM and afferent units provide excitatory inputs to in-phase MUs and form two separate positive feedback loops: one for flexion, one for extension. To prevent out-of-phase activation of MUs (from bidirectional premotor units), inhibitory cortical (CL) and spinal (Ia) units provide out-of-phase inhibition.

This analog network of 89 units and 2771 connections was transformed to a spiking network containing 480 spiking cells and 33,723 connections (on average, a cortical projection (CM) cell received 72 excitatory and 29 inhibitory connections and provided an output to 61 cells). The response-averaged activation profiles of the spiking inputs and representative cells are shown in Fig. 7. In this case all continuous units had the same custom activation function and were transformed into one type of spiking cell with the following properties: refractory period: 3 ms, threshold: 10 mV, rise- and decay-time of incoming PSPs: 2 and 7 ms respectively. During tonic firing, most spiking cells had ISI distributions that resembled Poisson events with a refractory period, i.e., they rose sharply after a refractory interval (ca. 4.8–5 ms.), representing the sum of absolute refractory period and integration times, and then fell off quasi-exponentially. The coefficient of variation for spiking cells varied between 0.2 and 2.48, a larger range than observed experimentally (0.4–1.2 [23]).

In Fig. 7 the first six traces represent the input cells driven by random spike trains modulated in proportion to the appropriate input signals, and the rest show average activity of representative spiking cells. All spiking cells in Fig. 7 show a clear unidirectional activity (either active during flexion or during extension), as do the continuous units. The only exception is RMt5, a rubral unit with a correct bidirectional tonic target activation. The activation profiles of the spiking cells resemble those of the continuous units: purely tonic and purely phasic profiles are well reproduced, whereas the distinction between phasic-tonic and decrementing profiles, a-priori more similar (cf. continuous MUf2 and MUf3), is less clear in the spiking cells.

4. Discussion

The primary reasons for transforming a recurrent neural network of continuous units into a network of integrate-and-fire cells are (1) to confirm the validity of the analog solution, (2) to exploit efficient training algorithms, such as backpropagation, for the former, and (3) to utilize these spiking networks for further investigations of temporal coding and plasticity. The transformation was applied here to two networks with substantially different complexity: a reduced network

model of a short-term memory task that retains a previously sampled value, and a more elaborate, biologically constrained network model that mimics neural activation patterns observed during step-tracking flexion/extension wrist movements in various areas of the primate motor system. We have also succeeded in transforming networks that simulated other behaviors, such as generating oscillatory activity, using the basic strategies described here.

4.1. The short-term memory network

The short-term memory network presents a neural solution for obtaining persistent activation of neurons in the absence of the stimulus, as has been observed in monkeys performing short-term memory tasks [1,12,24]. The solution depends on extracting a transient representation of the cue in the SA unit by exploiting the thresholding property of the sigmoid function—i.e., clipping the peak of the sum of S, A, and a negative bias. It was not clear whether this mechanism is a unique feature of the properties of continuous sigmoidal units. We here showed that this solution is also implemented in networks of spiking cells derived from the analog network. The average firing rates of the equivalent spiking cells are essentially similar to the activity of the continuous units.

Zipser et al. [29] also converted analog short-term memory networks to spiking networks using a fundamentally different approach. Their spiking cells fired probabilistically, in proportion to the weighted inputs from cells firing in the previous time-step, passed through an appropriately scaled sigmoid. This assured that the cells of their stochastic model exhibited firing rates generally similar to the continuous units. In contrast to this probabilistic procedure, our spiking cells interacted synaptically through delayed EPSPs and IPSPs, which were integrated to firing threshold. Thus the activity in our spiking network was generated entirely by these synaptic interactions.

It should be noted that our short-term memory networks sustain the value entirely in recurrent activity and do not employ any intrinsic memory in the units. In contrast, Compte et al. [5], investigated a recurrent biophysical network model, and proposed a role for persistent activity obtained by concurrent excitatory (NMDA receptor mediated) and inhibitory synaptic interactions.

4.2. The step-tracking wrist flexion/extension network

The simulation of the major parts of the neural premotor network during the step-tracking task elucidates the role of various units within the network and hence suggests possible functional roles of their biological counterparts (in terms of connectivity and activity). Corticomotoneuronal hidden and target units (CM, CMt) showed a variety of activation patterns, due to the combination of tonic and phasic input to the local and recurrent cortical network which included inhibitory cortical units. Many CM units had unidirectional activity and made selective and strong connections to multiple synergist motor units, and had negligible connections to antagonist motor units. This emergent property corresponds to the physiological characteristic of CM cells, which produced post-spike facilitation in multiple synergistic muscles, but not in their antagonists [8], and which facilitated different motor units within a muscle [18]. Some CM units had bidirectional, predominantly tonic, activation patterns, but these had negligible direct connections to motor units; this emergent property also corresponds to physiological evidence that bidirectional cortical cells do not produce post-spike facilitation [8]. In addition, unidirectional CM units developed strong connections to synergistic rubral units, to Ia-inhibitory units and to γ motor units. The first two kinds of connections have been shown to exist anatomically and physiologically [15,16], and evidence for cortico-fusimotor connections have been given by Clough et al. [4].

In contrast to the cortical module, most of the hidden rubral projection units had bidirectional activation patterns, and some units were unmodulated (i.e., bidirectional tonic units). Physiologically, more bidirectional neurons exist in the red nucleus than in the motor cortex and more showed co-facilitation of both flexor and extensor motoneurons [20]. The connection patterns of unidirectional rubral target units were similar to those of the unidirectional cortical projection units.

On the spinal level, Ia-inhibitory units showed unidirectional activations in-phase with their corresponding MUs, whereas the spindle afferents, with imposed unidirectional target activation patterns, developed strong connections to synergistic motor units. These kinds of connections, i.e. inhibition of out-of-phase motor units by antagonist Ia units and facilitation by in-phase agonist muscle spindles would be expected from the properties of their biological counterparts [16].

In summary, the combined input to α MUs consists primarily of excitatory unidirectional CM and RM units, excitatory feedback from strictly unidirectional spindle units and an inhibitory effect from the Ia-inhibitory units to assure the silence of the α MUs in the inactive phase. Many hidden units took on activity profiles resembling those found experimentally and the resulting weight matrix closely resembled known physiological pathways.

After conversion, the average firing rates of the corresponding spiking cells were largely similar, but not always identical to the profiles of the corresponding continuous units. The differences are partly due to the fact that a single fixed activation function of the continuous units does not adequately represent the range of

appropriate activation functions for spiking cells, which depend on the time-varying ratio between inhibitory and excitatory input and the shapes of the incoming PSPs (Fig. 1).

A second reason for minor discrepancies between continuous and spiking rate profiles is the fact that a minimal number of spiking cells was used to replace each continuous unit. This generated some synchronization, with consequences analogous to increasing PSP size. Networks with larger numbers of spiking cells could improve this problem by reducing synchronization and allowing smoother implementation of the average rates.

5. Concluding comments

These network models are based on the assumption that the essential information processing in recurrent networks is mediated by the units' firing rates. It has been suggested that the timing of spikes could potentially be used in biological networks to process information (e.g., [13]). Networks of integrate-and-fire cells derived by the procedures presented here can be used as a starting point to investigate such additional mechanisms of temporal coding. Another promising use of these spiking models is to test the effects of spike-timing dependent synaptic plasticity.

The present results could be extended in several ways to derive recurrent networks of biologically more realistic spiking cells. Custom sigmoids could be obtained by matching the input-output properties of units with biophysical properties, and used to train analog networks for subsequent conversion. The conversion also offers the possibility of using different synaptic parameters for particular subsets of units (such as the equivalents of interneurons, pyramidal neurons or motoneurons) and the use of corresponding fitted activation functions in the analog network. These steps would allow analog networks with intrinsically different continuous units to be derived with efficient gradient descent algorithms and converted to spiking networks with a higher degree of biological realism.

Acknowledgements

This work was supported in part by NIH grants NS 12542 and RR00166 and the Swiss National Science Foundation (fellowship to MAM).

References

[1] M.V. Chafee, P.S. Goldman-Rakic, Matching patterns of activity in primate prefrontal area 8a and parietal area 7ip neurons during spatial working memory task, J. Neurophysiol. 79 (1998) 2919–2940.

[2] Y. Chauvin, D.E. Rumelhart (Eds.), Back-propagation: Theory, Architectures, and Applications, Erlbaum, Hillsdale, NJ, 1995.

[3] P.D. Cheney, E.E. Fetz, Functional classes of primate corticomotoneural cells and their relation to active force, J. Neurophysiol. 44 (1980) 773–791.

[4] J.F.M. Clough, C.G. Phillips, J.D. Sheridan, The short-latency projection from the baboons' motor cortex to fusimotor neurones of the forearm and hand, J. Physiol. 216 (1971) 257–279.

[5] A. Compte, N. Brunel, P.S. Goldman-Rakic, X.J. Wang, Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model, Cerebral Cortex 10 (2000) 910–923.

[6] F.H.C. Crick, The recent excitement about neural networks, Nature 337 (1989) 129–132.

[7] E.E. Fetz, Dynamic recurrent neural network models of sensorimotor behavior, in: D. Gardner (Ed.), The Neurobiology of Neural Networks, MIT Press, Cambridge MA, 1993, pp. 165–190.

[8] E.E. Fetz, P.D. Cheney, Postspike facilitation of forelimb muscle activity by primate corticomotoneuronal cells, J. Neurophysiol. 44 (1980) 751–772.

[9] E.E. Fetz, L.E. Shupe, Dynamic models of neurophysiological systems, in: M. Arbib (Ed.), Handbook of Brain Theory and Neural Networks, MIT Press, 1995, pp. 332–335.

[10] E.E. Fetz, P.D. Cheney, K. Mewes, S. Palmer, Control of forelimb muscles activity by populations of corticomotoneuronal and rubromotoneuronal cells, Prog. Brain Res. 80 (1990) 437–449.

[11] D. Flament, P.A. Fortier, E.E. Fetz, Response patterns and postspike effects of peripheral afferents in dorsal root ganglia of behaving monkeys, J. Neurophysiol. 67 (1992) 875–889.

[12] J.M. Fuster, Inferotemporal units in selective visual attention and short-term memory, J. Neurophysiol. 64 (1990) 681–697.

[13] W. Gerstner, R. Kempter, J.L. van Hemmen, H. Wagner, A neuronal learning rule for sub-millisecond temporal coding, Nature 383 (1996) 76–78.

[14] P.S. Goldman-Rakic, Cellular basis of working memory, Neuron 14 (1995) 477–485.

[15] D.R. Humphrey, R. Gold, D.J. Reed, Sizes, laminar and topographic origins of cortical projections to the major divisions of the red nucleus in the monkey, J. Comp. Neurol. 225 (1984) 75–94.

[16] E. Jankowska, Interneuronal relay in spinal pathways from proprioceptors, Prog. Neurobiol. 38 (1992) 335–378.

[17] M.A. Maier, S.I. Perlmutter, E.E. Fetz, Response patterns and force relations of monkey spinal interneurons during active wrist movement, J. Neurophysiol. 80 (1998) 2495–2513.

[18] G.W.H. Mantel, R.N. Lemon, Cross-correlation reveals facilitation of single motor units in thenar muscles by single corticospinal neurones in the conscious monkey, Neurosci. Lett. 77 (1987) 113–118.

[19] P. Mazzoni, R.A. Andersen, M.I. Jordan, A more biologically plausible learning rule than backpropagation applied to a network model of cortical area 7a, Cerebral Cortex 1 (1991) 293–307.

[20] K. Mewes, P.D. Cheney, Primate rubromotoneuronal cells: parametric relations and contribution to wrist movement, J. Neurophysiol. 72 (1994) 14–30.

[21] R.C. O'Reilly, Biologically plausible error-driven learning using local activation differences: the generalised recirculation algorithm, Neural Comp. 8 (1996) 895–938.

[22] S.S. Palmer, E.E. Fetz, Discharge properties of primate forearm motor units during isometric muscle activity, J. Neurophysiol. 54 (1985) 1178–1193.

[23] W.R. Softky, C. Koch, The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs, J. Neurosci. 13 (1993) 334–350.

[24] J. Tanji, K. Kurata, Contrasting neuronal activity in supplementary and precentral motor cortex of monkeys, J. Neurophysiol. 53 (1985) 129–141.

[25] H.C. Tuckwell, Introduction to Theoretical Neurobiology, Vol. 1: Linear Cable Theory and Dendritic Structure, Cambridge University Press, Cambridge, 1988.
[26] M.C.W. van Rossum, G.Q. Bi, G.G. Turrigiano, Stable Hebbian learning from spike timing-dependent plasticity, J. Neurosci. 20 (2000) 8812–8821.
[27] R.L. Watrous, L. Shastri, Learning phonetic features using connectionist networks: an experiment in speech recognition, Technical Report, 1986, MS-CIS-86-87.
[28] R.J. Williams, D. Zipser, A learning algorithm for continually running fully recurrent neural networks, Neural Comp. 1 (1989) 270–280.
[29] D. Zipser, Recurrent network model of the neural mechanism of short-term active memory, Neural Comp. 3 (1991) 179–193.
[30] D. Zipser, B. Kehoe, G. Littlewort, J. Fuster, A spiking network model of short-term active memory, J. Neurosci. 13 (1993) 3406–3420.