Static and Dynamic State Feedback Control Model of Basal Ganglia – Thalamocortical Loops

A. Lőrincz [1]

International Journal of Neural Systems 8: 339-357, 1997

[1]A. Lőrincz
Department of Chemical Physics of the Institute of Isotopes
Hungarian Academy of Sciences, Budapest, P.O. Box 77, Hungary H-1525
and
Department of Adaptive Systems
Attila József University, Szeged, Dóm Square 9, Hungary H-6720
email: lorincz@iserv.iki.kfki.hu

Abstract

It is argued that a novel control architecture, the Static and Dynamic State (SDS) feedback scheme, which utilizes speed-field tracking, exhibits global stability, and allows on-line tuning by any adaptation mechanism without cancelling stability if certain structural conditions are met, can be viewed as a model of basal ganglia–thalamocortical loops since (1) the SDS scheme predicts the neuronal groups that fit neuronal classification in the supplementary motor area, the motor cortex and the putamen, (2) the structural stability conditions require parallel channels, a feature that these loops provide, and (3) the SDS scheme predicts two major disorders that can be identified as Parkinson's and Huntington's diseases. Simulations suggests that the basal ganglia work outside the realm of the stability condition allowed by the robustness of the scheme and required for increased computation speeds.

1 Introduction

Over recent years a considerable amount of research has been carried out to characterize the structure and the functional organization of the basal ganglia (BG) that comprise the caudate nucleus, the putamen, the ventral striatum – the latter three forming the striatum – the globus pallidus (GP) the substantia nigra (SN) and the subthalamic nucleus (STN). The BG have several puzzling features, including robust parallel input-output organization (Alexander et al., 1990) and high convergence ratios from input to output (Percheron et al., 1984a; Percheron et al., 1984b; Wilson, 1990; Percheron and Filion, 1991; Francois et al., 1995). The main parallel channels of the BG are the skeletomotor, the oculomotor, the associative, and the limbic channels. The present work concerns the skeletomotor channels; these receive inputs from most of the cerebral sensorimotor regions, including the supplementary motor area (SMA) and the motor cortex (MC). The output of the BG returns to the same regions via the thalamus and thus the BG form one of the stations of the basal ganglia – thalamocortical (BTC) loops. The loops are further divided into functionally segregated and highly specific parallel channels (Hoover and Strick, 1993). This high specificity concerns both the initiating and the receiving cortical areas, these areas show a degree of overlap and it seems that the information undergoes remapping within the putamen (Flaherty and Graybiel, 1991; Flaherty and Graybiel, 1993; Flaherty and Graybiel, 1994).

The parallel channels of the BG are divided into direct and indirect pathways that seem to act in an opposing manner: the direct pathway tend to enhance reentrant thalamocortical interaction by decreased inhibitory outflow from BG to thalamus whereas the indirect pathway tends to suppress that interaction by increased inhibitory outflow from BG to thalamus. These opposite effects are further emphasized by the dopamine pathways. Dopaminergic input to the putamen can be viewed as increasing the effects of the direct pathway, while decreasing the effects of the indirect pathway; in fact, dopamine seems to govern the relative contributions of the direct and indirect pathways.

Various functional and computational models have been proposed that emphasize different aspects of the basal ganglia–thalamocortical interaction. Functional models are important owing to the complexity of the BTC loops. Here, we shall overview models that deal with skeletomotor functions.

The thalamic disinhibition (TDI) model (DeLong, 1990) suggests that increasing (decreasing) the BG output shifts the motion towards hypokinetic (hyperkinetic) states. Although the model is capable of explaining several features of the BTC loops it also predicts that lesion to the BG output results in dyskinesia, which is not the case. It has been argued that this problem may be reconciled on the basis of the voltage-dependent bistable properties of thalamic neurons (Steriade and Llinas, 1988) that give rise to a different mode for thalamic neurons if their GABAergic input is strongly reduced (Alexander, 1995). However, the model predicts that lesion to the BG output recipient zones of the thalamus results in akinesia, which is not the case either.

The electrotonic coupling (EC) model of the BG (Connolly and Burns, 1993a; Connolly and Burns, 1993b) argues that the BG are responsible for driving smooth transitions of state. These works suggest that the corticostriatal arborizations implement a mapping from Cartesian coordinates (i.e., from a description formulated in terms of external space) to configurational space (i.e., to a description formulated in joint coordinates). The model proposes that the corticostriatal projections realize one of the steps of the computational sequences of control subserving goal-directed movements: the activated corticostriatal projections correspond to states of joint kinematics that are needed to achieve the planned

trajectory. The model is based on dye coupling experiments (Cepeda et al., 1989; Cepeda et al., 1991; Onn et al., 1991) indicating the possibility that electrotonic coupling via gap junctions takes place between striatal medium spiny neurons (MSNs). Then, according to the model, the striatum generates a control surface, or potential function, for guiding state changes by means of those contiguous regions which are cellular representations of state spaces. In the putamen, the somatotopic regions are suggested as being joint spaces, and the output (via the globus pallidus) is suggested as being an encoding of a trajectory vector, defining the next state change for the motor system.

The winner-lose-all (WLA) model (Berns and Sejnowski, 1995) starts from the specific architecture of the BG. It is based on the observation that the anatomy of the BG, and the connectivity of excitatory and inhibitory neurons may be interpreted as a WLA circuit. The WLA circuit contains multimodal sensory and cognitive maps of the cortex that project in a convergent manner to the striatum whereas the prefrontal cortex comprises replications of the thalamic output integrated over different time scales. The striatum receives processed sensory representation of the environment and sends two projections: a highly convergent one to the external segment of the globus pallidus (GPe), and a moderately convergent one to the internal segment of the globus pallidus (GPi). The first pool of striatal neurons that reaches the firing threshold inhibits its target in the GPi. This inhibition leads to the disinhibition of the corresponding thalamic targets and the gating of ascending information to the corresponding cortical motor neurons. At the same time, inhibition of a GPe neuron results in the disinhibition of the corresponding neuron of the STN, which then diffusively excites the other GPi neurons. This diffuse excitation is then assumed to prevent all but the first action from being selected. The cortico-STN pathway acts as negative feedback that turns off selected actions within the loop. The primary assumption for the WLA mechanism is the existence of streams of information that remain segregated from striatum to thalamus and that each of the segregated pools of neurons is devoted to a particular action.

Findings on lesions to GPi outputs and to the thalamic recipient zones of BG outputs pose similar problems to both models.

A comprehensive review of experimental findings and models of various aspects of BG can be found in a recent book (Houk et al., 1995b).

It has been argued by Lőrincz that the Static and Dynamic State (SDS) feedback control scheme (Szepesvári and Lőrincz, 1996; Szepesvári et al., 1997) fits the main features of the BTC loops, including the parallel organization of BTC channels and the opposing roles played by the direct and indirect pathways (Lőrincz, 1997). The model was based on a first order control scheme that cannot take full account of the skeletomotor function since the model should be based on a higher order control scheme. In this paper the predictions of the higher order scheme (Szepesvári and Lőrincz, 1997b) are compared to the specific features of the BTC loops.

The paper is organized as follows. First, in Section 2, the SDS scheme is reviewed, the stability theorem is stated, and the consequences are listed. Simulations on a 3-joint robotic arm demonstrate the robustness of the scheme outside the realm of the stability theorem (Sections 2.1 and 2.2), with details given in the Appendix. In the discussion section (Section 3), the biological findings and the model predictions are contrasted, including neuronal classifications in the SMA, the MC, and the putamen, the parallel organization featuring extensive remapping, and the opposing roles of the direct and the indirect pathways. The various features of the SDS model are compared with those of other models, such as the WLA, the EC, and TDI models. We shall argue that the WLA model is in disagreement with some of the findings of the Schultz group (Schultz and Romo, 1992;

Romo et al., 1992; Romo and Schultz, 1992); the EC model suffers from the 'curse of dimensionality'. In contrast the SDS model can be viewed as a mathematical framework for the TDI model that – according to the settings of model – avoids the pitfall of that model with regard to lesions to the BG output recipient zones of the ventral thalamus. Conclusions are drawn in Section 4.

2 The Static and Dynamic State feedback control scheme

The SDS architecture performs speed-field tracking. Let $D \subseteq \mathbf{R}^n$ denote the domain of the plant's state with the equation of motion given by

$$\mathbf{u} = \mathbf{A}(\mathbf{q})\dot{\mathbf{q}} + \mathbf{b}(\mathbf{q}), \tag{1}$$

where $\mathbf{q}$ is the state vector of the plant, the dot denotes temporal derivation, and $\mathbf{u} \in \mathbf{R}^m$ is the control. For the sake of notational simplicity the dependence of $\mathbf{A}$ and $\mathbf{b}$ on $\mathbf{q}$ will from now on not be explicitly represented. Let us now assume that we have an estimate of the true inverse-dynamics function $\Phi(\mathbf{q}, \dot{\mathbf{q}}) = \mathbf{A}\dot{\mathbf{q}} + \mathbf{b}$, given by $\hat{\mathbf{\Phi}}(\mathbf{q}, \dot{\mathbf{q}})$. The SDS Feedback Control equations can then be written as

$$\mathbf{u} = \mathbf{u}_{ff}(\mathbf{q}, \dot{\mathbf{q}}, \mathbf{v}(\mathbf{q})) + \mathbf{w}, \tag{2}$$

$$\dot{\mathbf{w}} = \Lambda \left(\hat{\mathbf{\Phi}}(\mathbf{q}, \mathbf{v}(\mathbf{q})) - \hat{\mathbf{\Phi}}(\mathbf{q}, \dot{\mathbf{q}}) \right), \tag{3}$$

where $\mathbf{u}_{ff}$ is the so called feedforward controller (to be specified later), $\dot{\mathbf{w}} = \mathbf{u}_{fb}$ is the so called feedback controller that gives rise to the feedback control vector $\mathbf{w}$ upon temporal integration, $\Lambda > 0$ is the gain of feedback, and the desired motion is determined by a speed-field tracking task that prescribes the speed vector $\dot{\mathbf{q}}$ of the plant as a function of the state vector:

$$\dot{\mathbf{q}} = \mathbf{v}(\mathbf{q}). \tag{4}$$

Speed-field tracking is not typical in the control literature but arises naturally if we consider stationary optimal-control problems such as path planning tasks (Hwang and Ahuja, 1992). Conventional control tasks such as *point-to-point control* and *trajectory tracking* cannot be exactly rewritten in the form of speed-field tracking and vice versa. The speed-field tracking task has the advantage that the designer can incorporate several objectives into the form of the speed-field to be tracked and hence extend the model's range of possibilities. Below, we review some of the results on the SDS control scheme (Szepesvári and Lőrincz, 1996; Szepesvári et al., 1997; Szepesvári and Lőrincz, 1997b).

If $\mathbf{M}$ is a real quadratic matrix ($\mathbf{M} : D \to \mathbf{R}^{p \times p}$, $p > 0$) then let $\mathbf{M} > 0$ denote that $\mathbf{M}$ is positive definite ($\mathbf{M}$ is said to be positive definite uniformly over D iff for all $\mathbf{q} \in D$ the term $\mathbf{M}(\mathbf{q})$ is positive definite and there exists an $\epsilon > 0$ such that $\lambda_{\min}(\mathbf{M}(\mathbf{q})) > \epsilon$ holds for all $\mathbf{q} \in D$). Similarly, if $\mathbf{M}$ is a matrix field over D, let $\mathbf{M} > 0$ denote that $\mathbf{M}$ is uniformly positive definite over D. Then the following theorem holds.

THEOREM 2.1 *Assume that the feedforward controller is given by the approximate inverse dynamics*

$$\mathbf{u}_{ff}(\mathbf{q}, \dot{\mathbf{q}}, \mathbf{v}) = \hat{\mathbf{\Phi}}(\mathbf{q}, \mathbf{v}) = \hat{\mathbf{A}}\mathbf{v} + \hat{\mathbf{b}} \tag{5}$$

and that the following assumptions hold:

1. $\mathbf{A}$ *and* $\hat{\mathbf{A}}$ *are invertible (in the generalized sense)*

2. $\mathbf{A}^{-1}$ *and* $\mathbf{D} = \hat{\mathbf{A}}^{-1}\mathbf{A}$ *are bounded away from singularities uniformly over* D

3. *the symmetrized perturbation matrix* $\mathbf{D} + \mathbf{D}^T$ *satisfies the sign-properness condition* $\mathbf{D} + \mathbf{D}^T > 0$

4. $\mathbf{A}^{-1}, \mathbf{b}, \mathbf{v}, \mathbf{D}, \hat{\mathbf{b}}$ *are bounded and have continuous, bounded derivatives over* D

Then denoting the error of tracking $\mathbf{v}(\mathbf{q})$ *by* $\mathbf{e} = \mathbf{v}(\mathbf{q}) - \dot{\mathbf{q}}$ *for all* $\epsilon > 0$ *there exists a gain* Λ *and absorption time* $T > 0$ *such that for all* $\mathbf{e}(0)$ *that satisfy* $||\mathbf{e}(0)|| < K\Lambda$ *it holds that* $||\mathbf{e}(t)|| < \epsilon$*, provided* $t > T$ *and the solution can be extended up till time* t*. Here* K *is a fixed positive constant and* $\mathbf{e}(0)$ *denotes the initial value of* $\mathbf{e}$*. Further,* $\Lambda \sim O(1/\epsilon)$*.*

The proof is based on an extension of Liapunov's second method (Szepesvári et al., 1997). It can be seen that the 'feedforward controller' is actually a static state feedback controller in this scheme, justifying the shorthand, SDS scheme. The term 'feedforward' is nevertheless kept since it describes a natural route of further generalizations.

The first order scheme , depicted in Fig. 1, can be interpreted as follows: We have an actual state ($\mathbf{q}$) and a desired speed ($\mathbf{v}$). We are equipped with the estimate of the inverse dynamics that provides us with a crude estimate of the true control vector ($\mathbf{u}_{ff} = \hat{\mathbf{\Phi}}(\mathbf{q}, \mathbf{v})$) that we term 'desired control vector'. We cannot use the desired control vector directly since it is well known that imprecise models of the inverse dynamics lead to instabilities. We utilize this control vector *together with* a correcting control vector. The correcting control vector is derived by measuring the actual speed and "asking" the inverse dynamics 'What if the actual speed were the speed that we desire?' The estimate of the inverse dynamics provides us with the 'experienced control vector' ($\hat{\mathbf{\Phi}}(\mathbf{q}, \dot{\mathbf{q}})$) when inputted with the actual state ($\mathbf{q}$) and the actual speed ($\dot{\mathbf{q}}$). The difference between the desired and actual control vectors (i.e., $\mathbf{u}_{fb}$) is then time integrated and the resulting vector $\mathbf{w}$ is used as the correcting means. These control vectors are thus developed on the basis of the approximate matrix $\hat{\mathbf{A}}$ and the approximate vector $\hat{\mathbf{b}}$. The sum of the two control vectors $\mathbf{u}_{ff}$ and $\mathbf{w}$ controls the plant which is subject to the equation of motion $\dot{\mathbf{q}} = \mathbf{A}^{-1}(\mathbf{u}_{ff} + \mathbf{w} - \mathbf{b})$. According to Theorem 2.1 the full scheme becomes a precise model of the inverse dynamics. The first order scheme served as the basis of an earlier model of the BTC loops (Lőrincz, 1997).

To cope with the control problem of systems subject to Newton's law, however, we need to extend this first order scheme to higher order schemes since it is not the speed but the acceleration that can be controlled. The higher order scheme is shown in Fig. 2. Although Fig. 1 and Fig. 2 are differently organized the only real difference is that the feedforward controller in Fig. 1 is replaced by the feedback controller in Fig. 2. The corresponding stability theorem is as follows:

THEOREM 2.2 *Assume that the feedforward controller has the form*

$$\mathbf{u}_{ff}(\mathbf{q}, \dot{\mathbf{q}}, \mathbf{v}) = \hat{\mathbf{\Phi}}(\mathbf{q}, \mathbf{v}) - \hat{\mathbf{\Phi}}(\mathbf{q}, \dot{\mathbf{q}}), \tag{6}$$

(which is the same as the input of the feedback integrator). Further, assume that the following hold:

1. $\hat{\mathbf{\Phi}}(\mathbf{q}, \dot{\mathbf{q}}) = \hat{\mathbf{A}}\dot{\mathbf{q}} + \hat{\mathbf{b}}$

2. $\mathbf{A}^T\mathbf{A}$, $\mathbf{A}^T\hat{\mathbf{A}}$ *and* $\hat{\mathbf{A}}^T\hat{\mathbf{A}}$ *are uniformly positive definite over* D *(sign-properness)*

3. $\mathbf{A}$, $\mathbf{v}$, $\mathbf{b}$ *are bounded and have uniformly bounded derivatives w.r.t.* $\mathbf{q}$ *over* D

Then for all $\Lambda > 0$ the error of tracking $\mathbf{v}(\mathbf{q})$, $\mathbf{e} = \mathbf{v}(\mathbf{q}) - \dot{\mathbf{q}}$, is eventually uniformly bounded and, further, the eventual bound b of the tracking-error can be made arbitrarily small. More specifically $b = \mathcal{O}(1/\Lambda)$, and the eventual bound for the time reaching $||\mathbf{e}|| \leq b$ is independent of Λ.

The proof of this theorem relies on a Liapunov-function approach. First of all the relation

$$(\mathbf{A} + \hat{\mathbf{A}})\mathbf{e} = \mathbf{A}\mathbf{v}(\mathbf{q}) + \mathbf{b} - \mathbf{w} \tag{7}$$

can be employed to show that $L = \frac{1}{2}\mathbf{e}^T \left[(\mathbf{A} + \hat{\mathbf{A}})^T(\mathbf{A} + \hat{\mathbf{A}})\right] \mathbf{e}$ is an appropriate semi-Liapunov function. Then the proof is complete (Szepesvári and Lőrincz, 1997b).

The subtle point of this theorem is the requirement that $\hat{\mathbf{A}}^T\hat{\mathbf{A}}$ should be uniformly positive definite over D. The theorem gives rise to a global stability result. Notice too that the particular form of the feedforward and feedback controllers makes it unnecessary to build an estimate of $\mathbf{b}$.

An important property of Theorem 2.2 is that in the error equation (Eq. 7) the r.h.s. does not depend on the approximated inverse-dynamics. This feature can be exploited to show that proof of Theorem 2.2 is not affected if $\hat{\mathbf{A}}$ and $\hat{\mathbf{b}}$ vary in time but the conditions of the theorem (most importantly the sign-properness condition) remain valid at every instant. Thus we get the following important corollary:

COROLLARY 2.3 *Suppose that the conditions of Theorem 2.2 hold and also that $\hat{\mathbf{A}} = \hat{\mathbf{A}}(t)$ and $\hat{\mathbf{b}} = \hat{\mathbf{b}}(t)$. Next assume that $\mathbf{A}^T\hat{\mathbf{A}}(t)$ and $\hat{\mathbf{A}}^T(t)\hat{\mathbf{A}}(t)$ are uniformly positive-definite over D and for all $t > 0$, and that $\hat{\mathbf{A}}(t)$ is bounded. Then the conclusions of Theorem 2.2 still hold.*

The lower bound of Λ is inversely proportional to $\inf_{t,\mathbf{q}} \lambda_{\min}(\mathbf{A}^T(\mathbf{q})\hat{\mathbf{A}}(\mathbf{q},t))$ and $\inf_{t,\mathbf{q}} \lambda_{\min}(\hat{\mathbf{A}}^T(\mathbf{q},t)\hat{\mathbf{A}}(\mathbf{q},t))$ and proportional to $\sup_{t,\mathbf{q}} \|\hat{\mathbf{A}}(\mathbf{q},t)\| + \|\hat{\mathbf{A}}(\mathbf{q},t)\|^2$. The uniform positive-definiteness conditions of the corollary follow, e.g. when $\hat{\mathbf{A}}$ is bounded away from singularities uniformly over D: an assumption often required in adaptive control (Sastry and Bodson, 1989). It is clear, too, that the stability result does not depend on the specific adaptation mechanism utilized, which is a fairly rare condition in adaptive control theory. However, one has to ensure, for example, that the sign-properness condition required for $\hat{\mathbf{A}}$ is obeyed.

2.1 Controlling a second order plant

In order to establish the similarities between the SDS scheme and the BG we overview here the control problem of a 3-joint robotic arm (Szepesvári and Lőrincz, 1997b).

Consider the following equation describing the motion of a rigid-link robot arm (see Fig. 3)

$$\mathbf{u} = \mathbf{M}(\theta)\ddot{\theta} + \mathbf{V}(\theta, \dot{\theta}) - k\dot{\theta}. \tag{8}$$

Here $\theta = (\theta_1, \theta_2, \theta_3)$ is the vector of angular positions of the robot (θ_1 is the angular position of the robot base axis, θ_2 is the angular elevation of the upper arm above horizontal, θ_3 is the angular elevation of the forearm above horizontal), $\mathbf{u} = (u_1, u_2, u_3)$ is the torque vector of actuators (u_1, u_2 and u_3 denote the torque of the base, the upper arm and the forearm actuators, respectively), $\mathbf{M}(\theta)$ is the inertia matrix, $\mathbf{V}(\theta, \dot{\theta})$ represents the Coriolis, centripetal forces and the gravity loading, and $k > 0$ is the friction coefficient. Note that this equation has the form of Eq. 1 with $\mathbf{q}^T = [\theta^T,\ \dot{\theta}^T]$, $\mathbf{A}(\mathbf{q}) = [\mathbf{0},\ \mathbf{M}(\theta)]$ and $\mathbf{b}(\mathbf{q}) = \mathbf{V}(\theta, \dot{\theta}) - k\dot{\theta}$. Unfortunately, the additive term is not bounded if $\|\dot{\theta}\|$ goes

to infinity. This means that the SDS scheme should be applied only if one can ensure *a priori* that $\dot{\theta}$ remains bounded (more specifically, a bound can be given for $\dot{\theta}$) during the control. Since the plant is a mechanical system boundedness can be achieved according to the law of energy conservation provided that the controls remain bounded, which can be achieved, for example, by some variant of the σ-modification scheme or a projection method. This condition was not implemented in the experiments, so the results once again show the robustness of the SDS scheme.

Assume that the task is defined in terms of a certain speed-field $\mathbf{s}(\theta)$ given in the configuration space, i.e., the task is to make sure that

$$\dot{\theta} = \mathbf{s}(\theta). \tag{9}$$

We admittedly abuse the terminology here in that we use the expression "speed-field" to mean a vector-field both over the configuration- and the state-spaces. The speed-field over the state-space needs another component apart from the speed-field over the configuration space, viz. the field of desired accelerations. So let this be defined as the difference between the desired and the actual speed:

$$\mathbf{a}(\theta, \dot{\theta}) = \mathbf{s}(\theta) - \dot{\theta} \tag{10}$$

In terms of $\mathbf{q}$ the desired speed-field is given by

$$\mathbf{v}(\mathbf{q})^T = [\mathbf{s}(\theta)^T,\ \mathbf{a}(\theta, \dot{\theta})^T] = [\mathbf{s}(\theta)^T,\ \mathbf{s}(\theta)^T - \dot{\theta}^T]. \tag{11}$$

Expressed in another way, $\mathbf{v}(\mathbf{q})^T = [\mathbf{v}(P_1\mathbf{q})^T,\quad \mathbf{v}(P_1\mathbf{q})^T - (P_2\mathbf{q})^T]$, where P_1 and P_2 project their arguments to the second and first coordinates, respectively (that is $P_1\mathbf{q} = \theta$ and $P_2\mathbf{q} = \dot{\theta}$). Let $\hat{\mathbf{\Phi}}(\mathbf{q}, \dot{\mathbf{q}}) = \hat{\mathbf{A}}(\mathbf{q})\dot{\mathbf{q}} + \hat{\mathbf{b}}(\mathbf{q})$, and let the 'feedforward' controller be the differencing controller of Theorem 2.2:

$$\begin{aligned} \mathbf{u}_{ff}(\mathbf{q}, \dot{\mathbf{q}}) &= \mathbf{u}_{ff}(\theta, \dot{\theta}, \ddot{\theta}) \\ &= \hat{\mathbf{\Phi}}(\mathbf{q}, \mathbf{v}(\mathbf{q})) - \hat{\mathbf{\Phi}}(\mathbf{q}, \dot{\mathbf{q}}). \end{aligned} \tag{12}$$

Additionally,

$$\begin{aligned} \hat{\mathbf{\Phi}}(\mathbf{q}, \mathbf{v}(\mathbf{q})) &= \hat{\mathbf{A}}(\mathbf{q})[\mathbf{s}(\theta),\ \mathbf{a}(\theta, \dot{\theta})] + \hat{\mathbf{V}}(\theta, \dot{\theta}) - \hat{k}\dot{\theta} \\ &= \hat{\mathbf{M}}(\theta)\mathbf{a}(\theta, \dot{\theta}) + \hat{\mathbf{V}}(\theta, \dot{\theta}) - \hat{k}\dot{\theta}. \end{aligned} \tag{13}$$

So if we let $\hat{\mathbf{\Psi}}(\theta, \dot{\theta}, \ddot{\theta}) = \hat{\mathbf{M}}(\theta)\ddot{\theta} + \hat{\mathbf{V}}(\theta, \dot{\theta}) - \hat{k}\dot{\theta}$ then $\hat{\mathbf{\Phi}}(\mathbf{q}, \mathbf{v}(\mathbf{q})) = \hat{\mathbf{\Psi}}(\theta, \dot{\theta}, \mathbf{a}(\theta, \dot{\theta}))$ and $\hat{\mathbf{\Phi}}(\mathbf{q}, \dot{\mathbf{q}}) = \hat{\mathbf{\Psi}}(\theta, \dot{\theta}, \ddot{\theta})$. Note that here the domain D is equal to $K^3 \times \mathbf{R}^3$, where K is the unit circle. Conditions *1* and *2* of Theorem 2.2 are equivalent to the conditions that $\hat{\mathbf{M}}^T\mathbf{M} > 0, \hat{\mathbf{M}}^T\hat{\mathbf{M}} > 0$, and $\mathbf{M}^T\mathbf{M} > 0$ all hold uniformly over D, and that $(\mathrm{d}/\mathrm{d}\mathbf{q})\mathbf{M}$ and $(\mathrm{d}/\mathrm{d}\mathbf{q})\mathbf{b}$ are bounded over D. These latter conditions depend only on the dynamics of the robot arm and are met. The positivity conditions are also met, e.g. when the load changes (Szepesvári et al., 1997).

It should be mentioned that one can use the 'simplified inverse-dynamics' $\hat{\mathbf{\Psi}}_0(\theta, \ddot{\theta}) = \hat{\mathbf{\Psi}}(\theta, 0, \ddot{\theta})$ because letting $\dot{\theta} = 0$ in $\hat{\mathbf{\Psi}}$ corresponds to an additive perturbation (since $\mathbf{M}$ does not depend on $\dot{\theta}$) and this can be compensated for without any additional requirements. The simplified inverse-dynamics is important for the case of learning schemes utilizing local approximators, as it means that only the configuration space (the 3-d space for the 3-joint robotic arm) and not the tangent bundle of the configuration space (i.e., the phase space, which is a 6-d space) need be discretized. This is an important property of the control scheme since the number of discretization units scales with the dimension in the exponent. The exact forms of $\mathbf{M}$ and $\mathbf{V}$ used in the computer experiments are given in the Appendix.

2.2 Robustness experiments

The robustness of the second control scheme was investigated in a series of computer experiments. Comparisons with the computed torque method showed the superiority of the SDS scheme (Szepesvári and Lőrincz, 1997b). In another series of computations the behaviour outside the realm of the theorem was studied (Szepesvári and Lőrincz, 1997a): The scheme makes extensive use of differencing. The question then arises whether the respective differences should be precisely computed or not. Constants were introduced into the control equations to simulate differencing imprecisions and point-to-point tasks with desired speed-field

$$\mathbf{s}(\theta) = \theta_0 - \theta, \tag{14}$$

acceleration field

$$\mathbf{a}(\theta, \dot{\theta}) = \mathbf{s}(\theta) - A\dot{\theta} \tag{15}$$

and control equations

$$\begin{aligned} \mathbf{v} &= B\hat{\mathbf{\Psi}}_0(\theta, \mathbf{a}(\theta, \ddot{\theta})) - C\hat{\mathbf{\Psi}}_0(\theta, \ddot{\theta}) && (16)\\ \mathbf{u} &= D\mathbf{v} + \mathbf{w} && (17)\\ \dot{\mathbf{w}} &= \Lambda\mathbf{v}. && (18) \end{aligned}$$

were considered. The sensitivity of the above control w.r.t the change of the involved coefficients A, B, C, D, and Λ was studied. The equations were sampled at a rate $\Delta = 0.001$s using Euler's method and the control time interval was of the same duration. The Euler's method was chosen to model non-predictive sampling methods in the solution. If one uses Euler's method with the given sampling rate it limits the range of the control torques (and thus the constants A, B, C, D and Λ), since otherwise the result is the total loss of computational precision, which means that $\dot{\theta}$ converges to infinity. The following values were used: $J = 0.5$ kg m^2, $M_1 = 10$ kg, $M_2 = 5$ kg, $L_1 = 0.6$ m, $L_2 = 0.8$ m, $k = 20$ kg m^2/s, $g = 9.81$ m/s^2 (see Fig. 3).

It was found that the system is relatively sensitive to coefficient A and much less sensitive to the other parameters (Szepesvári and Lőrincz, 1997a). Typical responses for different A values are shown in Fig. 4 (viz., for $A = 1.0, 0.5, 0.1, 0.0$). Even for this 'sensitive' parameter, to the naked eye there is no detectable difference if the constant of differencing is 0.5 instead of 1.0. If A is decreased further, oscillations settle in, but even for the small value ($A = 0.1$) in all the randomly generated cases tried the oscillations were damped and the point-to-point task was accomplished. One important feature of the graph is that the motion became much faster. Considering the rise-time, i.e., the time difference when the joint angles vary between 10% and 90% of the settling value, the motion is faster by a factor of about 4.0 for the case of $A = 0.1$. If A is further decreased the oscillations are not damped any more but start to grow in an exponential manner at some point, as is shown for $A = 0.0$.

3 The SDS scheme as a candidate for modelling the basal ganglia – thalamocortical loops

3.1 Basal ganglia for performing differencing

The SDS scheme has been suggested as a suitable candidate for constructing a model of higher-order motor functions of the basal ganglia – thalamocortical loops (Lőrincz, 1997), since several special properties of the SDS scheme are highlighted when dealing with these

motor areas. In what follows we shall assume that the main computational task of BG is differencing between desired and experienced parameters of motions. Consider the direct and the indirect pathways of BG. As is known, the corticostriatal projections to BG are glutamatergic and excitatory whereas the striatopallidal and the pallidofugal connections are GABAergic and inhibitory. Within the motor circuitry of the BG there are two major projections. The direct pathway projects from a certain subpopulation of striatal neurons in the putamen to the GPi. With regard to the indirect pathway, one arm has its origin in a different population of striatal medium spiny neurons, then proceeds through the GPe and the STN before reaching the GPi. The second arm of the indirect pathway is made up of GPe projections to the output nuclei themselves. All intrinsic projections are GABAergic and thus inhibitory, except the projections from the STN to the GPi, which are excitatory. We note that the full connectivity of direct and indirect pathways is more complex, see e.g. (Alexander, 1995), however the emerging picture remains the same. The two main arms of the two pathways are illustrated in Fig. 5.

These two groups of the various striatopallidal projections, i.e. the direct pathway and the indirect pathway, serve as the basis of the TDI model (DeLong, 1990) since activation of MSNs associated with the direct pathway tends to decrease BG output by directly suppressing activity at the level of GPi and SNr. In contrast, activation of MSNs associated with either arm of the indirect pathway will tend to increase BG output by increasing neuronal activity at the level of the output nuclei. In one case the BG output is increased by disinhibiting the STN with its excitatory projections to the GPi and the substantia nigra pars reticulata (SNr), whereas in the other case by directly disinhibiting the GPi and the SNr. The net result is that cortically initiated activation of the (indirect) direct pathway will tend to (suppress) enhance reentrant thalamocortical excitation by (increased) decreased inhibitory outflow from BG to thalamus.

These opposing roles of various GPe and GPi connections are further emphasized by the dopamine pathways. Dopamine is synthesized in the cell rich part of the SN, the substantia nigra pars compacta (SNc). Dopaminergic input to the putamen consists of nigrostriatal projections that originate in the SNc. At the network level dopamine has an inhibitory action on the striatal GABAergic cells projecting to the GPe, but an excitatory action on the neurons projecting directly to the GPi and the SNr. The role of dopamine could be viewed as increasing (decreasing) the effects of the direct (indirect) pathway and thus influencing the relative contributions of these pathways (West et al., 1986; Haracz et al., 1993; Kitai and Surmeier, 1993; Groves et al., 1995; Alexander, 1995).

The SDS model of BG assumes that all of the differencings of the model are performed by the basal ganglia. In other words, we assume that the desired acceleration field as well as the differencings related to the composition of the feedback controller are computed by the BG, whereas the experienced and desired directions are computed elsewhere, e.g., in the so called 'where' pathway. In this respect it is important that the major input to the BG is from the cerebral cortex and most of the cortical mantle is involved in this highly organized projection system to the striatum typically with reciprocal connections forming skeletomotor, oculomotor, associative and limbic loops. In contrast, the BG receive direct projections from parietal area 7b, i.e., from the 'where' pathway (Weber and Yin, 1984; Cavada and Goldman-Rakic, 1991) and this information received from the parietal cortex is not fed back directly to the parietal cortical areas (Selemon and Goldman-Rakic, 1985). The presumed differencing function of BG allows us to derive neuronal classification of the SMA, the MC and the putamen. It is also assumed that decisions about motor actions are formulated elsewhere and not only after but also before the initiation of the action the neuronal activities related to the desired parameters can manifest themselves.

3.2 Sign-properness and parallel channels

Before discussing neuronal classification in terms of the SDS scheme we ought to consider the sign-properness property of the scheme. The structural constraint that $\hat{\mathbf{A}}(\mathbf{q}) > 0$ should apply everywhere is a demanding requirement for the control architecture. Let us consider two examples.

Example 1. Consider the problem of mirror writing when the trajectory to be followed by the pen is observed through a mirror. The ordinary approximate inverse dynamics tuned for writing without any mirror will not be sign-proper in this case and the motion will diverge in an exponential fashion. If the components of the approximate inverse dynamics are modified accordingly and the mirror writing becomes sign-proper then writing without a mirror becomes divergent. It seems reasonable to divide the task space and utilize independent 'channels' for the two different tasks. In this case the approximate inverse dynamics should have two independent channels each activated upon recognition of the appropriate task.

Example 2. Consider the sudden task of catching a ball that will pass just above our hand. The control signal to be launched depends on several factors, including the present configuration of the body, the constraints of possible motions as well as the actual speed of parts of the body – including the hand. As has been mentioned the actual speed can be neglected owing to the fact that one can use the 'simplified inverse-dynamics' $\hat{\mathbf{\Psi}}_0(\theta, \ddot{\theta}) = \hat{\mathbf{\Psi}}(\theta, 0, \ddot{\theta})$ because letting $\dot{\theta} = 0$ in $\hat{\mathbf{\Psi}}$ corresponds to an additive perturbation (since $\mathbf{M}$ does not depend on $\dot{\theta}$) and this can be compensated for without any additional requirements. Body configurations, however, cannot be neglected. Raising the hand requires, for example, different shoulder motions depending on whether our body is in a vertical or horizontal position, etc. If any of the control components destroys the sign-properness of matrix $\hat{\mathbf{A}}(\mathbf{q})$ in any domain of $\mathbf{q}$ then the SDS scheme requires a separate sign-proper channel for control.

Both of the examples show that the SDS scheme, although robust and globally stable, still requires a complex setting for proper control: different tasks and different configurations may require different parallel channels for control.

The parallel organization is a prominent feature of the BG. Experiments in which anatomical tracers were placed into monkey brains to label simultaneously anterograde sensorimotor inputs to the striatum and retrograde striatal outputs to the pallidum showed that labelled input fibre clusters can overlap clusters of backward labelled projection neurons quite precisely (Flaherty and Graybiel, 1991; Flaherty and Graybiel, 1993; Flaherty and Graybiel, 1994). The emerging picture is that the information is dispersed to distributed modules in the striatum, called matrisomes (Malach and Graybiel, 1986; Desban et al., 1989; Selemon and Goldman-Rakic, 1990), but it can be brought together again at the next stage of processing, i.e., at the pallidum. This pattern has been observed both for striatal projections to the GPe and for projections to the GPi (Graybiel et al., 1994).

The organization of the BG can thus be viewed as a family of reentrant loops that are organized in parallel, with the matrisomes corresponding to different tasks, subtasks, or limb configurations that will be selected on the basis of the actual task to be executed. The output of the BG returns to parts of those same cortical fields by way of specific BG recipient zones in the dorsal thalamus; that is, the outputs of the matrisomes are re-collected according to the recipient zones that correspond to the cortical fields. Transneuronal retrograde tracing indicates that the BG – thalamocortical loops are organized to form multiple segregated output channels (Hoover and Strick, 1993). The connectivity structure thus looks like an extensive input-output remapping with high convergence ratios towards the

GPi (Percheron et al., 1984a; Percheron et al., 1984b; Wilson, 1990; Percheron and Filion, 1991; Francois et al., 1995), whereas the SDS scheme suggests that only parts of this remapped information are used and those parts correspond to the sign-proper feedback channels. This view is supported by the findings of Nini et al. (Nini et al., 1995) that neurons of the globus pallidus do not show correlated activity.

Although it is not the subject of the present paper, it may be worth noting that the BG, that exhibit modular structure have another type of module: the striosomes (Graybiel and C. W. Ragsdale, 1978; Graybiel, 1990; Gerfen, 1992). According to suggestions in the literature the matrisomes of the BG are trained by these striosomes that perform temporal differencing reinforcement learning (Houk, 1992; Montague et al., 1994; Houk et al., 1995a; Berns and Sejnowski, 1995). This striosomes model is based on differencing and has very much in common with the model of the matrisomes suggested by the SDS scheme including that the temporal differencing reinforcement learning equations can be reorganized to have the form of difference computing between 'expected' and 'experienced' channels.

3.3 Neuronal classification

Having explained the need of the task and configuration dependent parallel channels we shall identify the neuronal groups suggested by the SDS scheme. The desired acceleration may be expressed as the difference between the desired direction and the experienced direction. The feedforward control vector can be computed by means of subtracting the estimate of the "experienced control vector" from the estimate of the "desired control vector". The feedback control vector one notes is the time-integrated value of the feedforward control vector. Using these properties of the SDS Feedback scheme it is possible to characterize the model neurons of the control areas that prepare the signals undergoing differencing. For the learned and unperturbed motions one has the following classes:

1. Model neurons representing the desired acceleration will have higher firing rates in the so called preparatory phase, i.e., before motion is initiated, provided that motion closely follows the planned direction upon initiation.

2. The firing of model neurons representing experienced direction is motion related.

3. Model neurons that represent the desired direction have mixed characteristics.

4. Model neurons that represent the output of the approximate inverse dynamics, i.e., neurons that express the components of the control vector, behave as 'muscle like' neurons and exhibit motion related firing.

This classification of model neurons is in general agreement with the experimental findings (Tanji et al., 1980; Tanji and Kurata, 1985; Alexander and Crutcher, 1990b; Crutcher and Alexander, 1990; Lurito et al., 1991) and classification (Alexander and Crutcher, 1990b; Crutcher and Alexander, 1990). Other classes may be discovered according to the SDS model for perturbed motions. For this case the SDS model predicts that 'muscle like' neurons related to the desired and experienced control components can be dissociated. In this respect, prediction of the appearance of the experienced control components – control components that would appear provided that the experienced motion was the desired one – acts as a test of the validity of the model.

Two other features of the SDS scheme are connected with the organization into sign-proper channels: The sign of the differencing related to the desired and experienced estimates of the true inverse-dynamics function is subject to particular task settings (consider,

for example, the problem of mirror writing when the trajectory to be followed by the pen is observed through a mirror). This constraint involves separate feedback channels and recognition-based channel selection for sign-proper results. The division of the task-space and the computation of the acceleration field as well as the computation of the control components require separate neuronal subsets that compute task related and limb related components and exhibit either preparatory, or motion related, or mixed activities. This is yet another feature of neurons belonging to these loops (Alexander and Crutcher, 1990a).

3.4 Speed-field formulation of the control task

The SDS scheme, just like the EC model, offers a solution to the apparent lack of brain signals that would code motion trajectories. It seems that the brain formulates signals that specify the positions of targets in the extrapersonal space, e.g., neuronal signals of this kind can be found in the parietal association cortex (Zipser and Andersen, 1988). These signals allow the formulation of the control problem in terms of directions pointing towards target positions, i.e., in terms of speed-fields. The important feature of speed-field tracking is that motion is not specified in time. The speed-field tracking formulation is flexible as it allows motions to speed up or to slow down simply by scaling of the speed-field. It has been shown that such directional information can be used to train position-to-directional action (PDA) maps giving rise to robust population coding for the speed-field tracking task (Fomin et al., 1997).

3.5 Basal ganglia disorders

According to the SDS model of BG, BG perform differencing between the desired and experienced channels. Assume that the experienced channel is suppressed. In this case the desired channel rules the motion and accelerations are activated towards the desired direction. To give an example, consider the task of moving round a circle at constant speed. If the experienced channel is suppressed then the directional action points to the desired next position (on the circle) and the result is the acceleration along the path, i.e., the constant speed requirement will be spoilt. Moreover, this action will result in leaving the desired path since the desired path requires a force orthogonal to the path. The motion will be 'out of measure' in this respect, it resembles Huntington's disease.

Taking the other extreme, consider that the experienced channel is too strong whereas the desired channel is weak. In this case the opposite of the control vector corresponding to the experienced direction will be strongly activated and will oppose the actual experienced direction. This opposing force will slow down any motion and will thus give rise to akinesia, i.e., the main symptom of Parkinson's disease. These extremes can be predicted from the SDS model and these are the typical diseases of BG.

One can try to make quantitative modelling of these diseases by realizing that both of these diseases can be modelled as improper BG differencings. The numerical studies of Section 2.2 (Szepesvári and Lőrincz, 1997a) allow us to note some of the features of the SDS scheme. The scheme is less sensitive to the relative magnitude of the opposing control vectors. This is due to the fact that the control vectors are integrated and this integration can build up the appropriate correcting control vector. It is, however, unreasonable to expect integration without losses (such as the integration of the SDS scheme) and one can expect increased sensitivity if the integration has some kernel that limits the time of integration. It is easy to show that the stability theorem of the SDS scheme allows this generalization for an exponential kernel (Szepesvári and Lőrincz, 1997a).

On the basis of the numerical runs the SDS scheme was the most sensitive to the differencings between the desired and experienced directions used to compute the desired acceleration. The robustness of the scheme is demonstrated in Fig. 4; the naked eye cannot distinguish between the curves corresponding to $A = 1.0$ (perfect differencing) and $A = 0.5$ (the desired channel is twice as strong as the experienced channel). When decreasing parameter A further the improper differencing gives rise to damped oscillations. The other interesting property is that the motion becomes faster. The system is approaching a state that could be considered to be Huntington's disease. However, if we consider the observed appearance of dysmetria and unsmooth movement in cerebellar disorders when BG are the means of control (Kandel et al., 1991) the model indicates that BG differencing gives greater contribution to the desired channel. That is, the case of perfect differencing, where the present theorem holds, may already correspond to Parkinson's disease. This note can be taken further by claiming that the feedforward controller of learnt and practiced motions is not the BG but the cerebellum whereas the feedback controller is the BG, although the BG can play both roles if necessary. The change of the feedforward controller to another external controller can be smooth if the feedforward channel of the BG is weak and if the external feedforward controller is precise. In this respect it is important that computations conducted *without* the feedforward channel show minor differences and that differences appear at the start of the motion (Szepesvári and Lőrincz, 1997b) indicating the viability of this idea. Considering this type of robustness of the scheme and that integration and differencing are linear operations that can be interchanged we conclude that the two channels, the feedforward and the feedback channels can be unified into a *single* channel that undergoes integration with a weighted memory function, or temporal kernel. This concept of the BG that unifies the two channels by means of a weighted memory function seems to accord well with the findings that (1) activities in the supplementary motor area, the motor cortex and the putamen strongly overlap, and (2) the activities in the cortical regions precede the activity in the putamen (Alexander and Crutcher, 1990b; Crutcher and Alexander, 1990; Alexander and Crutcher, 1990a).

Consider the bistable properties of thalamic neurons (Steriade and Llinas, 1988). It has been assumed that there is shift of the function of these neurons when the GABAergic input of these neurons is drastically reduced owing to GPi lesions, and that a nonbursting relay mode appears explaining the lack of dyskinesia in this case (Alexander, 1995).

The other peculiar feature, that lesions to the thalamic recipient zones of BG outputs do not result in akinesia, can be explained by the presence of direct corticocortical connections between the SMA and the motor cortex that could represent a direct channel for the approximate inverse dynamics (without any kind of differencing – see later).

In other extreme cases the direct and the indirect pathways of BG can manifest themselves. Although the computer simulations indicate that the balancing of these channels is not crucial, nevertheless, in more heavily unbalanced cases the errors of the differencing should appear. If the contribution of the direct channel is too strong then the motion will be out of measure. Roughly speaking the motion will be ruled by the desired channel that alone acts as if the control would try to control the speed instead of the acceleration and thus the inertia of the limb is not taken into account. The expected result is that the motion overshoots giving rise to a sudden change of the desired direction and the motion overshoots again, etc. Depending on the extent of the unbalanced differencing, e.g., if the whole BG is affected or if the indirect channel is locally ineffective, etc., different forms of "out of measure" motions can be expected. This unbalancing corresponds to Huntington's disease.

If, however, the contribution of the indirect channel is too large then the experienced

motion is always counteracted according to the SDS scheme giving rise to slow motions. This case corresponds to Parkinson's disease. The main underlying reason of Parkinson's disease, according to the present model of BG, is that the desired acceleration is also a difference; it is the difference between the desired speed and the experienced speed, and that difference is computed by the BG. (Without this assumption the neuronal classification fails.) If, however, this differencing is improper and the experienced channel is relatively strong then the motion cannot be fast since large experienced speed signals overcompensate the desired speed signals and that, in turn, counteracts the desired motion up to the point that the contribution from the experienced speed becomes smaller than the contribution from the desired speed: Motion is slow under these conditions. If the contribution of the indirect channel could be reduced then Parkisonian symptoms could be ameliorated. In the case of the MPTP model of Parkinsonism, BG-cortical circuits lose their unique ability to keep pallidal neurons completely independent (Nini et al., 1995) and exhibit phase-locked oscillations at the GPi. The Parkinsonian symptoms of MPTP treated monkeys improve by STN lesions or high-frequency stimulations (Limousin et al., 1995). It was reported that these high frequency stimulations may give rise to cognitive disturbances. The improvement shown in respect of STN lesion is in agreement with the SDS model since MPTP treatment gives rise to the overactivity of excitatory STN neurons (Miller and DeLong, 1987). The high frequency stimulation can also be reconciled by means of the SDS scheme provided that this stimulation disturbs normal STN operations. The question that arises is whether cooling of the STN could ameliorate Parkinson's disease without cognitive disturbances.

3.6 Comparison of the SDS model with other BG models

The lesions to the GPi output and the thalamic recipient zones of BG output pose a problem to all of the models. In the case of the WLA and EC models when the BG represent an important link of motion control, such as the transformation of spatial coordinates to configurational space (EC) or decision making (WLA) it is hard to see how these lesions could be explained. In the case of the TDI and the SDS models the bistable property of the thalamic relay neurons offers an explanation for lesions to the GPi output. Also, for the TDI model it has been suggested that the pedunculopontine nucleus may provide another pathway to express BG functions and this can explain the effects of these thalamic lesions (Alexander, 1995). This suggestion should also work for the SDS model.

The findings for lesions to the thalamic recipient zones of BG, namely that this lesion does not give rise to akinesia, represents another problem for both the EC, the WLA, and the TDI models. The SDS model, however, offers the following solution. The SDS model makes extensive use of the approximate inverse dynamics. It assumes that the two parallel channels – the feedforward channel and the feedback channel – both undergo differencing when the desired and experienced contributions are subtracted. There is, however, a direct corticocortical connection between the SMA and the premotor areas and this may represent a direct channel of the approximate inverse dynamics not subject to any differencing. According to the numerical simulations the SDS scheme is robust against the change of relative contributions of the respective channels allowing this extra connection without risking performance qualities. This solves the problem of akinesia in this regard. The SDS model of BG can be viewed as an extension of the TDI model that formalizes how and why the thalamic disinhibition happens.

It may be worth noting here that the SDS model of BG is a robust control scheme that can adjust itself to other controllers and can work parallel with those. For example, as it

was mentioned before, the SDS can work parallel with a 'real' feedforward controller and it is not disturber by another direct channel of the approximate inverse dynamics either. We list two conditions why this type of parallel organization can be important.

1. During the evolution of a complex controller architecture the different subsystems should work together in a smooth fashion.

2. It has been suggested that the striosomes train the matrisomes of BG (Houk et al., 1995a) and the matrisomes provide an error signal to the feedforward controller, that is to the cerebellum (Lőrincz, 1997). This way the subsystems of the control architecture can work together and form a training hierarchy.

One of the drawbacks of the EC model is that BG neurons utilize local approximators (Gross and Graziano, 1995). If the role of BG is to transform spatial directions into joint space then local approximators require the discretization of a high dimensional space, a very expensive computational step since the number of local approximator units will scale with the dimension of freedom in an exponential fashion. Moreover, somewhere in the brain the dynamics has to be taken into account and that doubles the dimensionality of the problem that the EC model should take into account somehow. The SDS suggestion is less demanding in this respect since according to the scheme the simplified inverse dynamics can be used and that requires just the configuration space to be discretized. Moreover, the SDS scheme offers an alternative solution to configuration space discretization. According to the model it is satisfactory to know the directional action for the 'end effector' as long as that action is sign-proper. In other words, instead of discretizing configuration space (which is very expensive), the configuration space should be divided into sign-proper channels (this could be much less expensive since for different configurations the same approximate inverse dynamics may suffice and can be integrated into the same output channels of BG) and the directional action to each channel should be determined. Computational studies on realistic many degrees of freedom robotic arms are needed to explore this point further.

The WLA model explains several features of BG. The shortcoming that appears in the WLA model – beyond the problems mentioned that concern the lesions to the thalamus and the GPi output – is related to some of the findings of the Schultz group (Schultz and Romo, 1992; Romo et al., 1992; Romo and Schultz, 1992). According to these studies the BG output seems to reverberate for a longer time period (in the order of a second) before a voluntary action is made. In other words the output neurons can be active for a long time before the WLA effect could occur. One may thus question whether decisions are really made at the level of the BG. Moreover, the WLA decisions seem to contradict to the balancing effects of the direct and the indirect channels, unless one assumes that the BG represents a 'bang-bang' type controller. The findings of reverberating activities are in full agreement with the SDS model if – as was suggested earlier – we combine the two arms, the feedforward arm and the feedback arm, into one common arm by exchanging differencing and integration. This means that we introduce integration with a weighted memory function to be performed by neuronal populations of these arms. Then the reverberating activity represents the output of a possibly leaky integrator. The result is that SMA and MC neuronal activities precede activities in the putamen and the BG output in agreement with the experimental findings (Alexander and Crutcher, 1990b; Crutcher and Alexander, 1990; Alexander and Crutcher, 1990a). Decisions can be made according to the integrated activities and decisions would then release the desired channels and the prepared activities would start the motion.

It has been noted that the temporal differencing reinforcement learning model of the striosomes (Houk, 1992; Montague et al., 1994; Houk et al., 1995a; Berns and Sejnowski, 1995) can be rephrased as the computation of the difference between the expected reward and the experienced reward. One can thus raise the question: "Up to what extent can we consider the basal ganglia as the substrates that compute the difference between 'desired' (or 'expected', or 'planned') and 'experienced' (or 'actrual') components of motion *and* behaviour?

4 Conclusions

The higher order Static and Dynamic State (SDS) model of the basal ganglia (BG) – thalamocortical loops assumes that the control actions are formulated by means of an approximate inverse dynamics controller. The control actions are made precise by extensive differencings all of which can be formulated in terms of desired and experienced parameters. The desired acceleration is the difference between the desired speed and the experienced speed whereas the output of one stage of the controller is the difference between the desired control vector (given by the approximate inverse dynamics when inputted by the desired speed) and the experienced control vector (given by the approximate inverse dynamics when inputted by the experienced speed). The desired and experienced channels of the SDS scheme are identified with the direct and indirect pathways of BG. Numerical simulations indicated that exact balancing of the two pathways is not required and the scheme can tolerate fairly large differences. According to the model when the relative contribution of the direct pathway is large then moves 'out of measure' appear; in contrast, a relatively large contribution from the indirect pathway gives rise to slow motion with apparent friction. These extremes were identified with Huntington's disease and Parkinson's disease, respectively. The SDS model can be viewed as an extension and a mathematical framework of the thalamic disinhibition model (TDI) (DeLong, 1990) that is capable of explaining the origin of the TDI model.

The main requirement of the SDS controller is that the division of the task space (that includes joint configurations and sensory inputs) should be sign-proper, a structural condition for the controller. It has been argued that this condition on sign-properness can be taken as the origin of the parallel channels of the BTC loops and the extensive remapping as well as convergence within the BG. It was argued that the sign-proper division of the task space could be a unit saving alternative against configurational space discretization. Also, the SDS scheme allows the utilization of the simplified inverse dynamics since it does not require the discretization of the configurational space and momentum space together but the configurational space alone. Since the biological substrate seems to use a local approximator approach in these loops (Gross and Graziano, 1995) this option could offer a tremendous advantage over other methods.

It has been shown that the resulting neuronal classification precisely corresponds to the experimental findings (Alexander and Crutcher, 1990b; Crutcher and Alexander, 1990; Alexander and Crutcher, 1990a).

It has been demonstrated by means of computer simulations that the scheme is robust and can work outside the realm of the validity of the theorems. In particular it was demonstrated that perfect differencing between the desired and the experienced channels is not required.

The SDS scheme features a feedforward and a feedback channel. Computer simulations indicate that the feedback channel alone can control the 3-joint robotic arm (Szepesvári and Lőrincz, 1997a) giving rise to the observation that these two channels can be unified

into a single feedback channel by introducing a weighted memory function (a temporal kernel) that can approximate the joint effects of the feedforward and feedback channels. This modification of the scheme (1) can explain the overlapping activities of the supplementary motor area, the motor cortex, and the putamen, and (2) emphasizes the possibility of incorporating an external feedforward controller, e.g., the cerebellum, into this control scheme.

Another feature of the SDS scheme is that the scheme does not formulate the control task in terms of tarjectories but rather it designs directional actions for every point of the extrapersonal space, which is in agreement with the known properties of motion specifying signals (Zipser and Andersen, 1988).

The question was raised if the main functional task of the basal ganglia is the computation of the differences between 'desired' ('expected', 'planned') and 'experienced' ('actual') components of motion and behaviour.

5 Acknowledgments

This work was partially funded by OTKA-Hungary grants T017110 and the US-Hungarian Joint Fund Grant 519/95-A.

Appendix

The exact forms of $\mathbf{M}$ and $\mathbf{V}$ used in the computer experiments are taken from those of a benchmark problem (Anderson and Miller, III, 1992) and are detailed below with the conventions and notations closely following the cited work. (The numerical parameters are given in Section 2.1)

Let

$$\begin{aligned} a_1 &= (M_1 + M_2)L_1^2 \cos^2(\theta_2) + M_2 L_1 L_2 \cos(\theta_2)\cos(\theta_3) + M_2 L_2^2 \cos^2(\theta_3) + J \\ a_2 &= (M_1 + M_2)L_1^2 \\ a_3 &= M_2 L_1 L_2 \sin(\theta_2 + \theta_3) \\ a_4 &= M_2 L_2^2 \end{aligned}$$

and

$$H = (a_2 a_4 - a_3^2).$$

Then

$$\mathbf{M}(\theta) = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & H(a_4 - a_3) & H(a_4 - a_3) \\ 0 & H(a_4 - a_3)\, a_3/a_2 & H/a_2((a_4 - a_3)a_3 + 1) \end{pmatrix}.$$

Further, let

$$\begin{aligned} b_1 &= (2(M_1 + M_2)L_1^2 + M_2 L_1 L_2)\sin(\theta_2)\cos(\theta_2) \\ b_2 &= M_2 L_1 L_2 \cos(\theta_2)\sin(\theta_3) + 2M_2 L_2^2 \sin(\theta_3)\cos(\theta_3) \\ b_3 &= M_2 L_1 L_2 \sin(\theta_3 - \theta_2) \\ b_4 &= (M_1 - M_2)L_1^2 \sin(\theta_2)\cos(\theta_3) + M_2 L_1 L_2 \sin(\theta_2)\cos(\theta_3) \\ b_5 &= M_2 L_2^2 \sin(\theta_3)\cos(\theta_3) + M_2 L_1 L_2 \cos(\theta_2)\sin(\theta_3). \end{aligned}$$

Then

$$V(\theta,\dot{\theta}) = \begin{pmatrix} b_1\dot{\theta}_1\dot{\theta}_2 + b_2\dot{\theta}_1\dot{\theta}_3 \\ -b_3\dot{\theta}_3^2 - b_4\dot{\theta}_1^2 - (M_1+M_2)L_1\cos(\theta_2)g \\ b_3\dot{\theta}_2^2 - b_5\dot{\theta}_1^2 - M_2L_2\cos(\theta_3)g \end{pmatrix}.$$

6 Figure Captions

Figure 1 Doubling the role of the inverse-dynamics controller

For first order plants the approximate model of the inverse dynamics $\mathbf{u}_{ff} = \hat{\mathbf{\Phi}}(\mathbf{q}, \mathbf{v})$ (with $\mathbf{v}$ describing the 'desired direction') can be used to build a precise model of the inverse dynamics by doubling the role of the approximate model. A correcting control vector $\mathbf{u}_{fb} = \mathbf{u}_{ff} - \hat{\mathbf{\Phi}}(\mathbf{q}, \dot{\mathbf{q}})$ (with $\dot{\mathbf{q}}$ describing the 'experienced direction) is to be computed, time integrated, and amplified. By adding this vector to $\mathbf{u}_{ff}$ the resulting control scheme becomes a precise model of the inverse dynamics.

Figure 2 Compensatory control by using the inverse dynamics controller in differencing mode

The scheme utilizes identical feedforward and feedback controllers built from estimates of the approximate inverse dynamics. The feedforward and feedback controllers provide control vectors $\mathbf{u}_{ff}$ and $\mathbf{u}_{fb}$, respectively. The feedback control vector is time integrated and the resulting control vector is added to the feedforward control vector. The state equation of the plant gives rise to changes in the state of the plant subject to its actual state, the control vector, and the dynamics of the plant represented by. The identical feedforward and feedback controllers are detailed in the lower part of the figure. According to the scheme each controller comprises two copies of the model of the inverse dynamics. These models are inputted by the 'desired' and the 'experienced' speeds, respectively. The outputs of these models undergo differencing. (See text for further details.)

Figure 3 Idealized 3-joint robotic manipulator

Notations: L_1: length of upper arm, L_2: length of forearm, M_1: point mass between the two parts of the arm, M_2: point mass at the end of the arm, J: rotational inertia of base, $\theta_i, i = 1, 2, 3$: the angular position of manipulator base axis, the angular elevation of upper arm above horizontal, the angular elevation of forearm above horizontal, respectively.

Figure 4 Responses of the robot arm

Robustness test of the SDS controller against imprecisions of differencing are shown for the most sensitive parameter A. If $A = 1.0$ then the desired acceleration $\mathbf{a}$ derived from the desired direction $\mathbf{s}$ and the experienced direction $\dot{\theta}$ ($\mathbf{a}(\theta, \dot{\theta}) = \mathbf{s}(\theta) - A\dot{\theta}$) is properly approximated. Different A values gives rise to improper differencing. The SDS scheme is much less sensitive to errors in differencing for other components of the scheme according to the computer experiments.

Figure 5 Schematic representation of parts of the direct and the indirect pathways

Both pathways are excited by corticostriatal projections. The part of the direct pathway shown in the figure inhibits the pallidofugal connections of the internal segment of the globus pallidus (GPi) and thus disinhibits the thalamocortical connections. The part of the indirect pathway shown in the figure inhibits the projections of the external segment of the globus pallidus (GPe) and, in turn, disinhibits the excitatory projections of the subthalamic nucleus (STN) to the GPi and thus reinforces the pallidofugal connections and, in turn, suppresses the thalamocortical projections. Open and filled arrows represent excitatory and inhibitory connections, respectively.

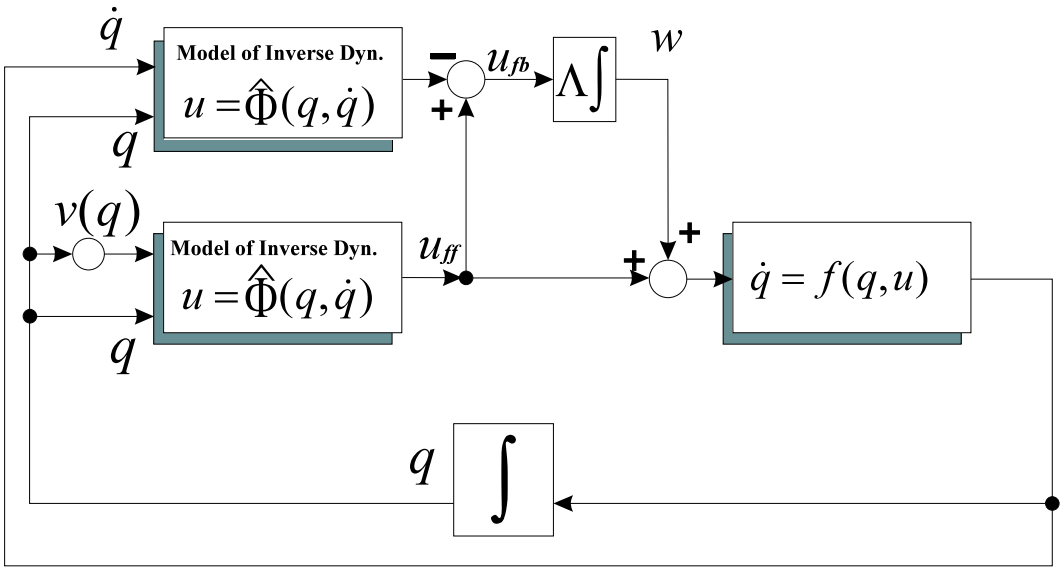

Figure 1: Doubling the role of the inverse-dynamics controller

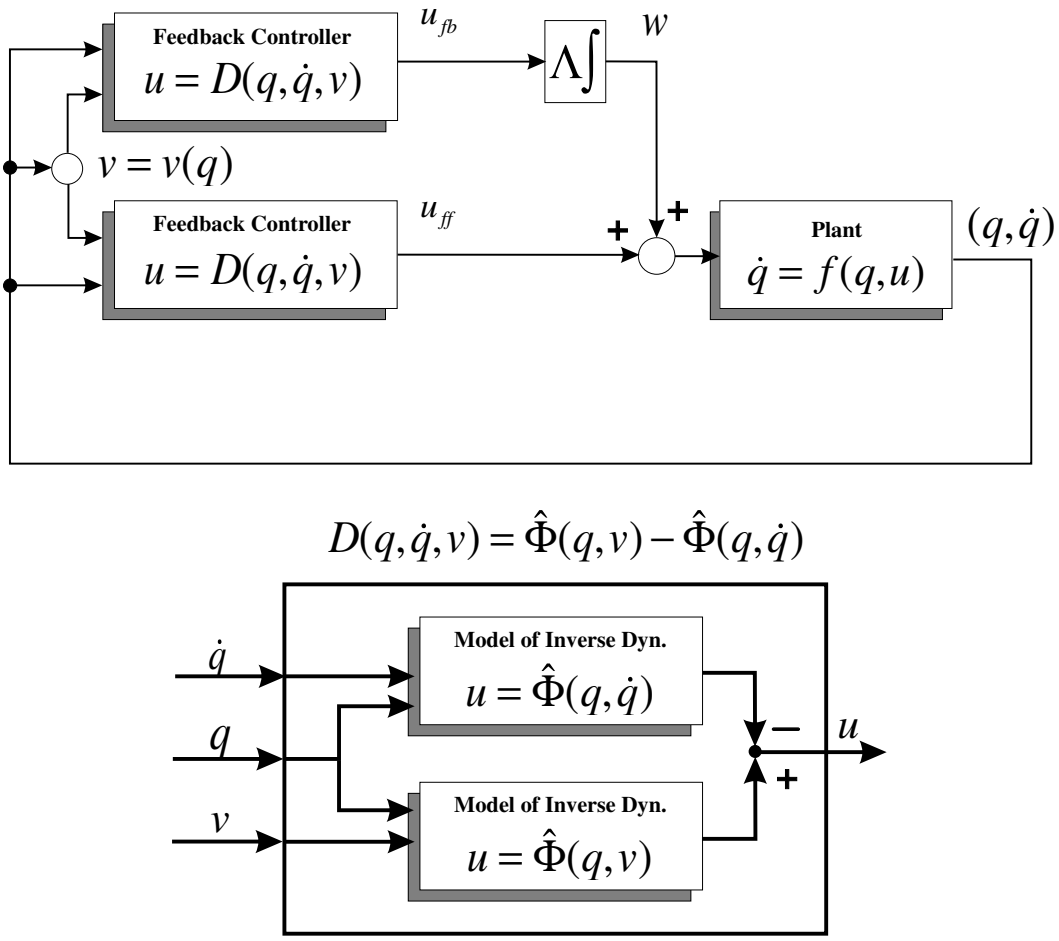

Figure 2: Compensatory control by using the inverse dynamics controller in differencing mode

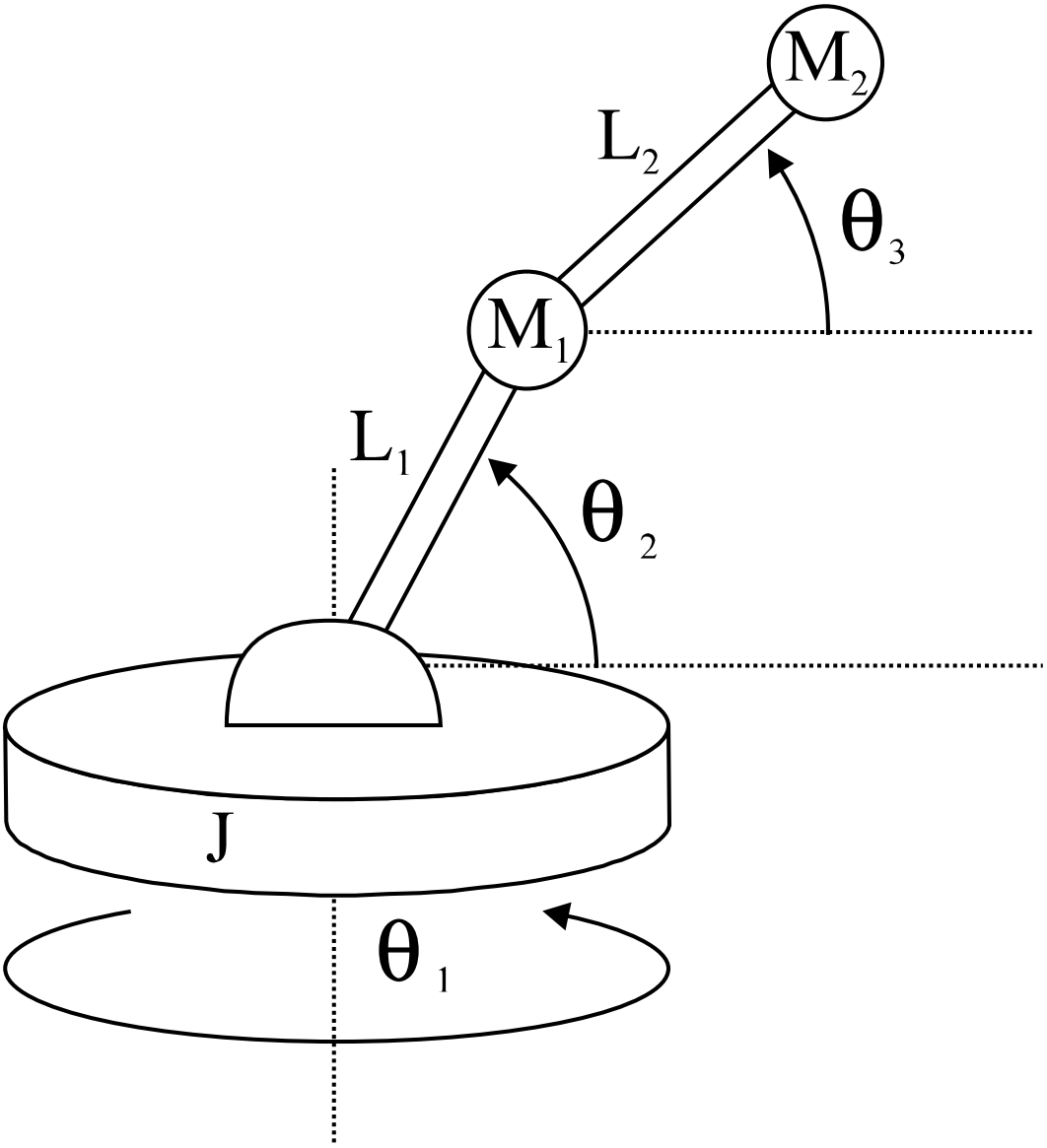

Figure 3: Idealized 3-joint robotic manipulator

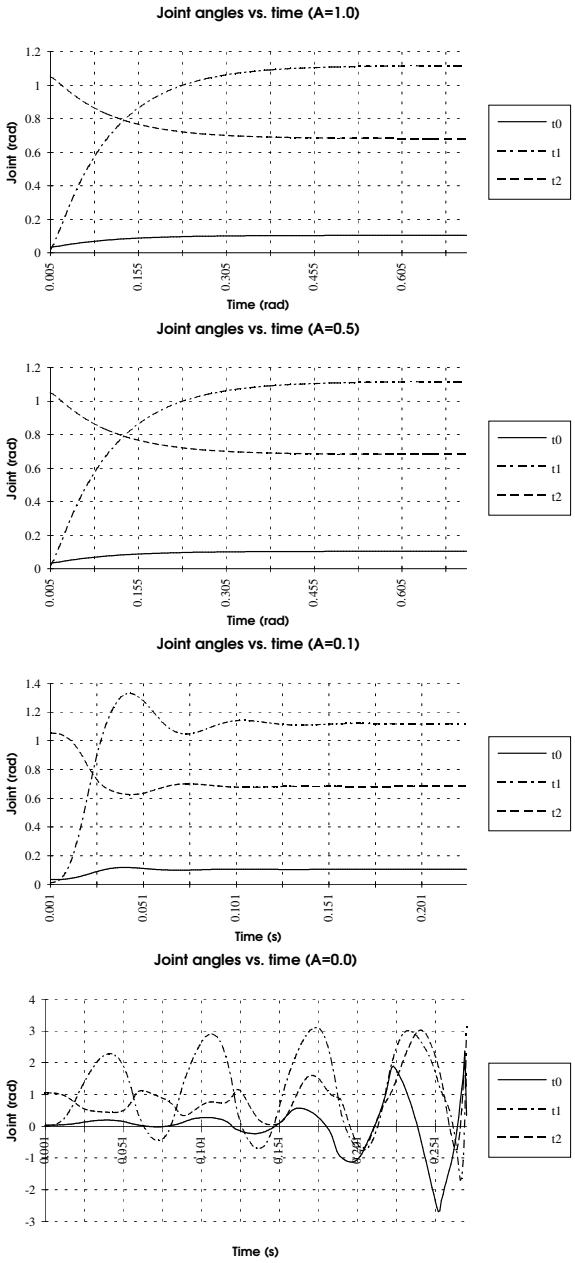

Figure 4: Responses of the robot arm

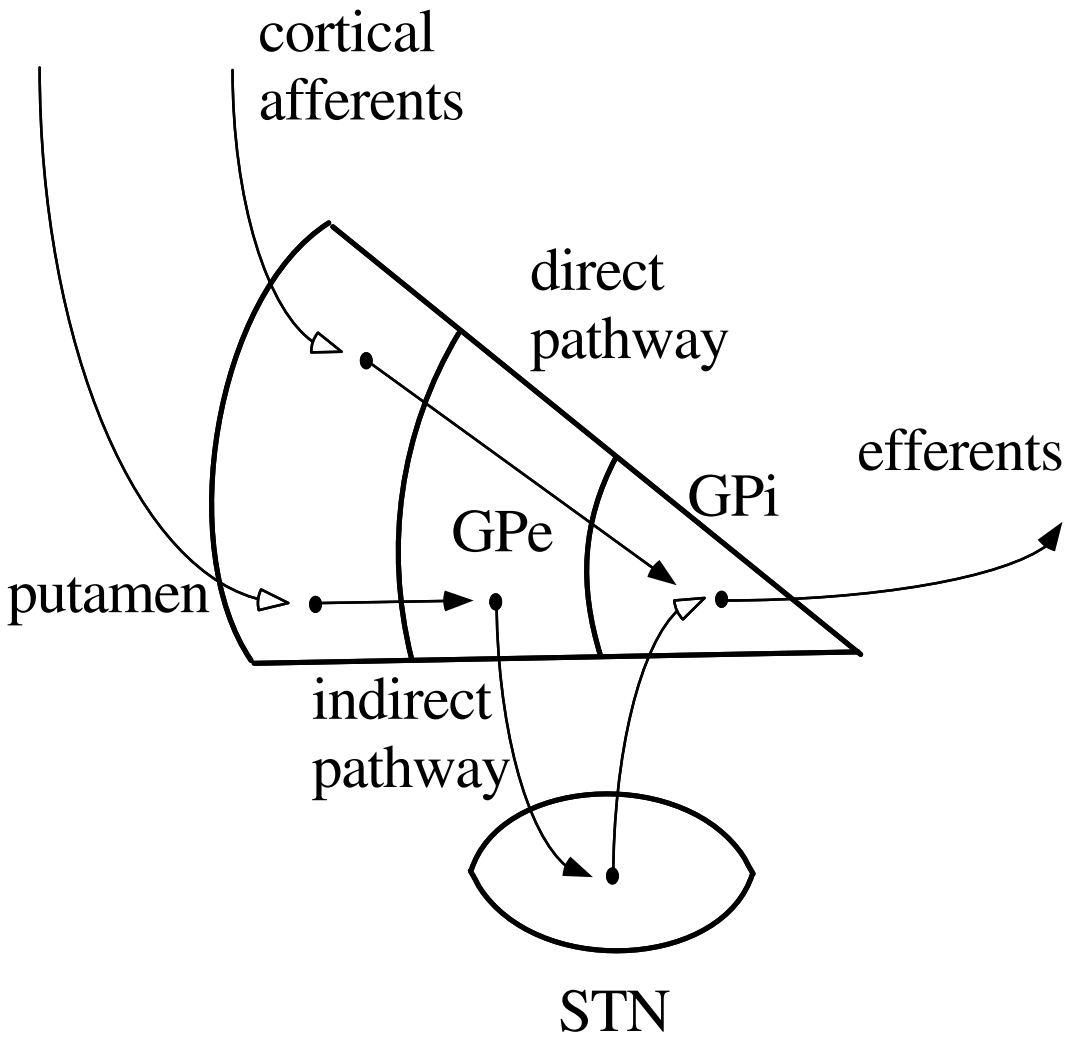

Figure 5: Schematic representation of parts of the direct and the indirect pathways

References

Alexander, G. E. (1995). Basal ganglia. In *The Handbook of Brain Theory and Neural Networks, ed. M.A. Arbib*, pages 139–144. Bradford Books / MIT Press, Cambridge, MA.

Alexander, G. E. and Crutcher, M. D. (1990a). Neural representations of the target (goal) of visually guided arm movements in three motor areas of the monkey. *Journal of Neurophysiology*, 64:164–178.

Alexander, G. E. and Crutcher, M. D. (1990b). Preparation for movement: Neural representations of intended direction in three motor areas of the monkey. *Journal of Neurophysiology*, 64:133–150.

Alexander, G. E., DeLong, M. R., and Strick, P. L. (1990). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Ann. Rev. Neuroscience*, 9:357–381.

Anderson, C. and Miller, III, W. (1992). Challenging control problems. In *Neural Networks for Control, eds. W.T. Miller, III and R.S. Sutton and P.J. Werbos*, Neural Network Modeling and Connectionism, pages 475–510. MIT Press, Cambridge.

Berns, G. S. and Sejnowski, T. J. (1995). How the basal ganglia make decisions. In *The Neurobiology of Decision Making*. Springer-Verlag, Heidelberg.

Cavada, C. and Goldman-Rakic, P. S. (1991). Topographic segregation of corticostriatal projections from posterior parietal subdivisions in the macaque monkey. *Neuroscience*, 42:683–696.

Cepeda, C., Walsh, J. P., Buchwald, N. A., and Levine, M. S. (1991). Neurophysiological maturation of cat caudate neurons: Evidence from in vitro studies. *Synapse*, 7:278–290.

Cepeda, C., Walsh, J. P., Hull, C. D., Howard, S. G., Buchwald, N. A., and Levine, M. S. (1989). Dye-coupling in the neostriatum of the rat: I. Modulation by dopamine-depleting lesions. *Synapse*, 4:229–237.

Connolly, C. I. and Burns, J. B. (1993a). A model for the functioning of the striatum. *Biological Cybernetics*, 68:535–544.

Connolly, C. I. and Burns, J. B. (1993b). A new striatal model and its relationship to basal ganglia diseases. *Neuroscience Research*, 16:271–274.

Crutcher, M. D. and Alexander, G. E. (1990). Movement-related neuronal activity selectively coding either direction or muscle pattern in three motor areas of the monkey. *Journal of Neurophysiology*, 64:151–163.

DeLong, M. R. (1990). Primate models of movement disorders of basal ganglia origin. *Trends in Neuroscience*, 13:281–285.

Desban, M., Gauchy, C., Kernel, M. L., Besson, M. J., and Glowinski, J. (1989). Three-dimensional organization of the striosomal compartment and patchy distribution of striatonigral projections in the matrix of the cat caudate nucleus. *Neuroscience*, 29:551–566.

Flaherty, A. W. and Graybiel, A. M. (1991). Corticostriatal transformations in the primate somatosensory system. projections from physiologically mapped body-parts representation. *Journal of Neurophysiology*, 66:1249–1263.

Flaherty, A. W. and Graybiel, A. M. (1993). Output architecture of the primate putamen. *Journal of Neuroscience*, 13:3222–3237.

Flaherty, A. W. and Graybiel, A. M. (1994). Input-output organization of the sensorimotor striatum in the squirrel monkey. *Journal of Neuroscience*, 14:599–610.

Fomin, T., Rozgonyi, T., Szepesvári, C., and Lőrincz, A. (1997). Self-organizing multiresolution grid for motion planning and control. *International Journal of Neural Systems*. In Press.

Francois, C., Percheron, G., and Yelnik, G. (1995). Three dimensional tracing of individual axons following biocytin injection into the striatum of macaques. In *V-th Meeting International Basal Ganglia Society*, page 19. Plenum Press, New York.

Gerfen, C. R. (1992). The neostriatal mosaic: multiple levels of compartmental organization. *Trends in Neuroscience*, 15:133–139.

Graybiel, A. M. (1990). Neurotransmitters and neuromodulators in the basal ganglia. *Trends in Neuroscience*, 13:244–254.

Graybiel, A. M., Aosaki, T., Flaherty, A. W., and Kimura, M. (1994). The basal ganglia and adaptive motor control. *Science*, 265:1826–1831.

Graybiel, A. M. and C. W. Ragsdale, J. (1978). Histochemically distinct compartments in the striatum of human, monkeys, and cat demonstrated by acetylthiocholinesteras staining. *Proc. Natl. Acad. Sci. USA*, 75:5723–5726.

Gross, C. G. and Graziano, M. S. A. (1995). Multiple representations of space in the brain. *The Neuroscientist*, 1:43–50.

Groves, P., Garcia-Munoz, J., Linder, J., Manley, M., Martone, M., and Young, S. (1995). Elements of the intrinsic organization and information processing in the neostriatum. In *Models of Information Processing in the Basal Ganglia, eds. J. C. Houk and J. L. Davis and D. G. Beiser*, pages 51–96. MIT Press, Cambridge, MA.

Haracz, J., Tschantz, J., Wang, Z., White, I., and Rebec, G. (1993). Striatal single-unit responses to amphetamine and neuroleptics in freely moving animals. *Neurosci. Biobehav. Rev.*, 17:1–12.

Hoover, J. E. and Strick, P. L. (1993). Multiple output channels in the basal ganglia. *Science*, 259:819–821.

Houk, J. C. (1992). Learning in modular networks. In *Proceedings of the Seventh Yale Workshop on Adaptive and Learning Systems*, pages 80–84. Center for Systems Science, New Haven CO.

Houk, J. C., Adams, J. L., and Barto, A. G. (1995a). A model of how the basal ganglia generate and use neural signals that predict reinforcement. In *Models of Information Processing in the Basal Ganglia, eds. J. C. Houk and J. L. Davis and D. G. Beiser*, pages 249–270. MIT Press, Cambridge, MA.

Houk, J. C., Davis, J. L., and Beiser., D. G. (1995b). In *Models of Information Processing in the Basal Ganglia*. MIT Press, Cambridge, MA.

Hwang, Y. and Ahuja, N. (1992). Gross motion planning – a survey. *ACM Computing Surveys*, 24(3):219–291.

Kandel, E., Schwartz, J., and Jessel, T. (1991). In *Principles of neural science*, page 1019. Appleton and Lange, Norwalk, Connecticut.

Kitai, S. and Surmeier, D. (1993). Cholinergic and dopaminergic modulation of potassium conductances in neostriatal neurons. In *Parkinson's Disease. From Basic Research to Treatment, eds. H. Narabayashi and T. Nagatsu and N. Yanagisawa and Y. Mizuno*, pages 44–52. Raven Press, New York.

Limousin, P., Pollak, P., Benazzouz, A., Hoffman, D., Brousolle, E., Perret, J., and Benabid, A. (1995). Bilateral subthalamic nucleus stimulation for severe Parkinson's disease. *Movement Disorders*, 10:672–674.

Lőrincz, A. (1997). Neurocontrol III: Temporal differencing models of the basal ganglia – thalamocortical loops. *Neural Network World*, pages 43–72.

Lurito, J., Georgakopoulos, T., and Georgopoulos, A. (1991). Cognitive spatial-motor processes. *Experimental Brain Research*, 87:562–580.

Malach, R. and Graybiel, A. M. (1986). Mosaic architecture of the somatic sensory-recipient sector of the cat's striatum. *Journal of Neuroscience*, 12:3436–3458.

Miller, W. and DeLong, M. (1987). Altered tonic activity of neurons in the globus pallidus and the subthalamic nucleus in the primate mptp model of Parkinsonism. In *The basal ganglia II, eds. M.B. Carpenter and A. Jayaraman*, pages 415–427. Plenum Press, New York.

Montague, P. R., Dayan, P., and Sejnowski, T. J. (1994). Foraging in and uncertain environment using predictive Hebbian learning. In *Neural Information Processing Systems 6, eds. J. D. Cowan, G. Tesauro, and J. Alspector*, pages 598–605. Morgan Kaufmann, San Francisco CA.

Nini, A., Feingold, A., Slovin, H., and Bergman, H. (1995). Neurons in the globus pallidus do not show correlated activity in the normal monkey, but phase-locked oscillations appear in the MPTP model of parkinsonism. *Journal of Neurophysiology*, 74:1800–1805.

Onn, S. P., Berger, T. W., and Grace, A. A. (1991). Dye-coupling in type I and type II striatal neurons: Alteration by apomorphine and localization to the patch/matrix. In *Proceedings of the 1991 Annual Meeting of the Society for Neuroscience*.

Percheron, G. and Filion, M. (1991). Parallel processing in the basal ganglia: Up to a point. *Trends in Neuroscience*, 14:55–56.

Percheron, G., Yelnik, G., and Francois, C. (1984a). In *The Basal Ganglia*, pages 87–105. Plenum Press, New York.

Percheron, G., Yelnik, G., and Francois, C. (1984b). A Golgi analysis of the primate globus pallidus: III. Spatial organization of the striato-pallidal complex. *Journal of Neural Computation*, 227:214–227.

Romo, R., Scarnati, E., and Schultz, W. (1992). Role of primate basal ganglia and frontal cortex in the internal generation of movements: II. Movements-related activity in the anterior-striatum. *Experimental Brain Research*, 91:385–395.

Romo, R. and Schultz, W. (1992). Role of primate basal ganglia and frontal cortex in the internal generation of movements: III. Neuronal activity in the supplementary motor area. *Experimental Brain Research*, 91:396–407.

Sastry, S. and Bodson, M. (1989). *Adaptive Control - Stability, Convergence and Robustness*. Prentice Hall, Englewood Cliffs, New Jersey.

Schultz, W. and Romo, R. (1992). Role of primate basal ganglia and frontal cortex in the internal generation of movements: I. Preparatory activity in the anterior striatum. *Experimental Brain Research*, 91:363–384.

Selemon, L. D. and Goldman-Rakic, P. S. (1985). Longitudinal topography and interdigitation of corticostriatal projections in the rhesus monkey. *Journal of Neuroscience*, 5:776–794.

Selemon, L. D. and Goldman-Rakic, P. S. (1990). Topographic intermingling of striatonigral and striatopallidal neurons in rhesus monkeys. *Journal of Comp. Neurol.*, 297:359–376.

Steriade, M. and Llinas, R. R. (1988). The functional states of the thalamus and the associated neuronal interplay. *Physiological Review*, 68:649–742.

Szepesvári, C., Cimmer, S., and Lőrincz, A. (1997). Dynamic state feedback neurocontroller for compensatory control. *Neural Networks*. In press.

Szepesvári, C. and Lőrincz, A. (1996). Inverse dynamics controllers for robust control: Consequences for neurocontrollers. In *Proceedings of International Conference on Artificial Neural Networks, eds. C. von der Malsburg, W. von Seelen, J.C. Vorbrüggen and B. Sendhoff*, pages 697–702, Bochum, Germany. Springer-Verlag, Berlin.

Szepesvári, C. and Lőrincz, A. (1997a). Unpublished observations.

Szepesvári, C. and Lőrincz, A. (1997b). Approximate inverse-dynamics based robust control using static and dynamic state feedback. In *Neural Adaptive Control Theory*, volume 2. World Scientific, Singapore. In press.

Tanji, J. and Kurata, K. (1985). Contrasting neuronal activity in supplementary and precentral motor cortex of monkeys. I. Responses to instruction determining motor responses to forthcoming signals of different modalities. *Journal of Neurophysiology*, 53:129–141.

Tanji, J., Tanuguchi, K., and Saga, T. (1980). Supplementary motor area: Neuronal response to motor instructions. *Journal of Neurophysiology*, 43:60–68.

Weber, J. T. and Yin, T. C. T. (1984). Subcortical projections of the inferior parietal cortex (area 7) in the stump-tailed monkey. *J. Comp. Neurol.*, 224:206–230.

West, M., Michael, A., Knowles, S., Chapin, J., and Woodward, D. (1986). Striatal unit activity and the linkage between sensory and motor events. In *Basal Ganglia and Behavior: Sensory Aspects of Motor Functioning, eds. J.S. Schneider and T.I. Linsky*, pages 27–35. Hans Huber, Bern.

Wilson, C. J. (1990). Basal ganglia. In *The Synaptic Organization of the Brain*, pages 279–316. Oxford University Press, New York.

Zipser, D. and Andersen, R. A. (1988). A backpropagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature*, 331:679–684.