

# Contracting model of the basal ganglia\*

**Benoît Girard<sup>1</sup>, Nicolas Tabareau<sup>1</sup>, Jean-Jacques Slotine<sup>2</sup> and Alain Berthoz<sup>1</sup>**

1. Laboratoire de Physiologie de la Perception et de l'Action, CNRS - Collège de France

11 place Marcelin Berthelot, 75231 Paris Cedex 05, France.

2. Nonlinear Systems Laboratory, Massachusetts Institute of Technology

Cambridge, Massachusetts, 02139, USA

## Abstract

It is thought that one role of the basal ganglia is to constitute the neural substrate of action selection. We propose here a modification of the action selection model of the basal ganglia of (Gurney et al., 2001a,b) so as to improve its dynamical features. The dynamic behaviour of this new model is assessed by using the theoretical tool of contraction analysis. We simulate the model in the standard test defined in (Gurney et al., 2001b) and also show that it performs perfect selection when presented a thousand successive random entries. From a biomimetic point of view, our model takes into account a usually neglected projection from GPe to the striatum, which enhances its efficiency.

Keywords: contraction analysis, action selection, basal ganglia, computational model

## 1 Introduction

The basal ganglia are a set of interconnected subcortical nuclei, involved in numerous processes, from motor functions to cognitive ones (Mink, 1996; Middleton and Strick, 1994). Their role is interpreted as a generic selection circuit, and they thus have been proposed to constitute the neural substrate of action selection (Mink, 1996; Krotopov and Etlinger, 1999; Redgrave et al., 1999).

The basal ganglia are included in cortico-basal ganglia-thalamo-cortical loops, five main loops have been identified in primates (Alexander et al., 1986, 1990; Kimura and Graybiel, 1995): motor, oculomotor, prefrontal (two of them) and limbic loops. Within each of these loops, the basal ganglia circuitry is organised in interacting channels, among which selection occurs. The output nuclei of the basal ganglia are tonically active and inhibitory, and thus maintain their targets under sustained inhibition. Selection occurs *via* disinhibition (Chevalier and Deniau, 1990): the removal of the inhibition exerted by one channel on its specific target circuit allows the activation of that circuit. Concerning action selection, the basal ganglia channels are thought to be associated to basic

competing actions. Given sensory and motivational inputs, the basal ganglia are thus supposed to arbitrate among these actions and to allow the activation of the winner by disinhibiting the corresponding motor circuits.

Numerous computational models of the BG have been proposed in the past (Gillies and Arbruthnott, 2000, for a review) in order to explain the operation of this disinhibition process, the most recent and complete model –in terms of anatomically identified connections accounted– is the GPR model proposed by Gurney et al. (2001a,b). Beyond its generic selection properties, explored in (Gurney et al., 2001b), the efficiency of the GPR as an action selection device has been tested in both robotic and simulated animats solving various tasks, involving execution of behavioural sequences, survival and navigation (Montes-Gonzalez et al., 2000; Girard et al., 2003, 2005).

The properties of the GPR were analytically studied at equilibrium, however the stability of this equilibrium (and thus the possibility to reach it) was not assessed. We propose to use contraction analysis (Lohmiller and Slotine, 1998) –a theoretical tool to study the dynamic behaviour of non-linear systems– in order to build a new model of the basal ganglia whose stability can be formally established. By using recent data (Parent et al., 2000) concerning the projections of a basal ganglia nucleus (the external part of the globus pallidus), we improve the quality of its selection with regards to GPR and then test this improvement in simulation. Finally, we discuss the remaining biomimetic limitations of the proposed model.

## 2 Nonlinear Contraction Analysis

Basically, a nonlinear time-varying dynamic system will be called *contracting* if initial conditions or temporary disturbances are forgotten exponentially fast, i.e., if trajectories of the perturbed system return to their nominal behaviour with an exponential convergence rate. This is an extension of the well-known *stability* analysis for linear systems with the great advantage that relatively simple conditions can still be given for this stability-like property to be verified, and furthermore that this property is preserved through basic system combinations. We also want to stress that assuming that a system is contracting, we only have to find a particular stable trajectory to be sure that the system will eventually tend to this trajectory. It is thus a way to analyse the dynamic behaviour of a model without linearised approximation.

---

\*The support of the BIBA project funded by the European Community, grant IST-2001-32115 is acknowledged.

## 2.1 The basic brick

In this section, we summarise the variational formulation of contraction analysis of (Lohmiller and Slotine, 1998), to which the reader is referred for more details. It is a way to prove the contraction of a whole system by analysing the properties of its Jacobian only. This can be seen as the basic brick of the theory, as in next sections we will often study the contraction of small components of the system and then deduce the global contraction of the system using combination rules (see section 2.2).

Consider a  $n$ -dimensional time-varying system of the form:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), t) \quad (1)$$

where  $\mathbf{x} \in \mathbb{R}^n$  and  $t \in \mathbb{R}_+$  and  $\mathbf{f}$  is  $n \times 1$  non-linear vector function which is assumed to be real and smooth in the sense that all required derivatives exist and are continuous. This equation may also represent the closed-loop dynamic of a neural network model of a brain structure.

We now restate the main result of contraction analysis, see (Lohmiller and Slotine, 1998) for details and proof.

**Theorem 1** *Consider the continuous-time system (1). If there exists a uniformly positive definite metric*

$$\mathbf{M}(\mathbf{x}, t) = \Theta(\mathbf{x}, t)^T \Theta(\mathbf{x}, t)$$

such that the generalised Jacobian

$$\mathbf{F} = (\dot{\Theta} + \Theta \mathbf{J}) \Theta^{-1}$$

is uniformly negative definite, then all the all system trajectories converge exponentially to a single trajectory with convergence rate  $|\lambda_{max}|$ , where  $\lambda_{max}$  is the largest eigenvalue of the symmetric part of  $\mathbf{F}$ . The system is said to be contracting.

**Remark.** In many cases, if the system is not properly defined, the expected metric may be hard to find. Most often, it is possible to fall into a standard combination of contracting systems just by rearranging the order of variables considered whereas the original definition of the system did not stress contraction properties.

## 2.2 Combination of contracting systems

We now present standard results on combination of contracting systems which will help us in showing that our model is contracting by analysing first contraction of each nucleus on one side and then their relative combination.

### Hierarchies

The most useful combination is the hierarchical one. Consider a virtual dynamic of the form

$$\frac{d}{dt} \begin{pmatrix} \delta \mathbf{z}_1 \\ \delta \mathbf{z}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{F}_{11} & \mathbf{0} \\ \mathbf{F}_{21} & \mathbf{F}_{22} \end{pmatrix} \begin{pmatrix} \delta \mathbf{z}_1 \\ \delta \mathbf{z}_2 \end{pmatrix}$$

The first equation does not depend on the second, so that exponential convergence of the whole system can be guaranteed (Lohmiller and Slotine, 1998). The results can be applied recursively to combinations of arbitrary size.

## Feedback Combination

Consider two contracting systems and an arbitrary feedback connection between them (Slotine, 2003). The overall virtual dynamics can be written

$$\frac{d}{dt} \begin{pmatrix} \delta \mathbf{z}_1 \\ \delta \mathbf{z}_2 \end{pmatrix} = \mathbf{F} \begin{pmatrix} \delta \mathbf{z}_1 \\ \delta \mathbf{z}_2 \end{pmatrix}$$

Compute the symmetric part of  $\mathbf{F}$ , in the form

$$\frac{1}{2} (\mathbf{F} + \mathbf{F}^T) = \begin{pmatrix} \mathbf{F}_{1s} & \mathbf{G}_s \\ \mathbf{G}_s^T & \mathbf{F}_{2s} \end{pmatrix}$$

where by hypothesis the matrices  $\mathbf{F}_{is}$  are uniformly negative definite. Then  $\mathbf{F}$  is uniformly negative definite if and only if  $\mathbf{F}_{2s} < \mathbf{G}_s^T \mathbf{F}_{1s}^{-1} \mathbf{G}_s$ , a standard result from matrix algebra (Horn and Johnson, 1985). Thus, a sufficient condition for contraction of the overall system is that

$$\sigma^2(\mathbf{G}_s) < \lambda(\mathbf{F}_1) \lambda(\mathbf{F}_2) \quad \text{uniformly } \forall \mathbf{x}, \forall t \geq 0$$

where  $\lambda(\mathbf{F}_i)$  is the contraction rate of  $\mathbf{F}_i$  and  $\sigma(\mathbf{G}_s)$  is the largest singular value of  $\mathbf{G}_s$ . Again, the results can be applied recursively to combinations of arbitrary size.

## Contraction analysis on convex regions

Consider a contracting system  $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$  maintained in a convex region  $\Omega$  (i.e. a region  $\Omega$  in which any shortest connecting line (geodesic)  $\int_{\mathbf{x}_1}^{\mathbf{x}_2} \|\delta \mathbf{x}\|$  between two arbitrary points  $\mathbf{x}_1$  and  $\mathbf{x}_2$  in  $\Omega$  is completely contained in  $\Omega$ ). Then all trajectories in  $\Omega$  converge exponentially to a single trajectory (Lohmiller and Slotine, 2000). Furthermore, the contraction rate can only be sped up by the convex constraint.

## 2.3 Our basic contracting system : the leaky integrator

In our model of basal ganglia, we will use leaky integrator models of neurons. The following equations describe the behaviour of our neurons where  $\tau$  is a time constant  $a(t)$  is the activation,  $y(t)$  is the output,  $I(t)$  represents the input of the neuron, and  $f$  is a continuous function which maintains the output in an interval.

$$\begin{cases} \tau \dot{a}(t) = -a(t) + I(t) \\ y = f(a) \end{cases}$$

This kind of neuron is basically contracting since its Jacobian is  $-\frac{1}{\tau}$  and the interval defined by the transfer function is a particular convex region.

In the rest of this paper, we will use the family of functions  $f_{\varepsilon, max}$ :

$$\begin{cases} 0 & \text{if } x \leq \varepsilon \\ x - \varepsilon & \text{if } \varepsilon \leq x \leq max + \varepsilon \\ max & \text{else} \end{cases} \quad (2)$$

## 3 Model description

The basic architecture of our model is very similar to the GPR (fig 1). We use the same leaky-integrator model of neurons as building blocks, each BG channel in each nucleus being represented by one such neuron. The input of the system is a

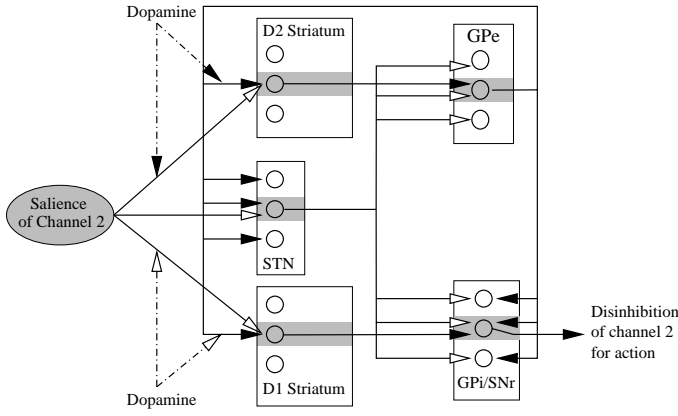


Figure 1: Basal ganglia model. Nuclei are represented by boxes, each circle in these nuclei represents an artificial leaky-integrator neuron. On this diagram, three channels are competing for selection, represented by the three neurons in each nucleus. The second channel is represented by grey shading. For clarity, the projections from the second channel neurons only are represented, they are identical for the other channels. White arrowheads represent excitations and black arrowheads, inhibitions. D1 and D2: neurons of the striatum with two respective types of dopamine receptors; STN: sub-thalamic nucleus; GPe: external segment of the globus pallidus; GPi/SNr: internal segment of the globus pallidus and substantia nigra pars reticulata.

vector of saliences, representing the propensity of each behaviour to be selected. Each behaviour in competition is associated to a specific channel and can be executed if and only if its level of inhibition decreases below a fixed threshold  $\theta$ .

An important difference between the GPR and our model is the nuclei targeted by the external part of the globus pallidus (GPe) and the nature of these projections. The GPe projects to the subthalamic nucleus (STN), the internal part of the globus pallidus (GPi) and the substantia nigra pars reticulata (SNr), but also to the striatum. Our model includes the striatum projections, which have been documented (Staines et al., 1981; Kita et al., 1999) but excluded from previous models. Moreover, the striatal terminals target the dendritic trees, while pallidal, nigral and subthalamic terminals form perineuronal nets around the soma of the targeted neurons (Sato et al., 2000). This specific organisation allows GPe neurons to influence large sets of neurons in GPi, SNr and STN (Parent et al., 2000), thus the sum of the activity of all GPe channels influences the activity of STN and GPi/SNr neurons (eqn. 5 and 7), while there is a simple channel-to-channel projection to the striatum (eqn. 3 and 4).

The striatum is one of the two input nuclei of the BG, mainly composed of GABAergic (inhibitory) medium spiny neurons. As in the GPR model, we distinguish the neurons with D1 and D2 dopamine receptors and modulate the input generated in the dendritic tree by  $\lambda$ , which here encompasses salience and GPe projections. Lateral inhibitions are also implemented, but their weights  $w_{LatD1}$  and  $w_{LatD2}$  is kept within the limits set the contraction analysis (see section 4.1). The

input to each neuron  $i$  of the D1 and D2 sub parts of the striatum is therefore defined as follows ( $N$  being the number of channels):

$$I_i^{D1} = (1 + \lambda)(S_i - w_{GPe}^{D1} y_i^{GPe}) - w_{LatD1} \sum_{\substack{j=1 \\ j \neq i}}^N I_j^{D1} \quad (3)$$

$$I_i^{D2} = (1 - \lambda)(S_i - w_{GPe}^{D2} y_i^{GPe}) - w_{LatD2} \sum_{\substack{j=1 \\ j \neq i}}^N I_j^{D2} \quad (4)$$

The up-state/down-state of the striatal medium spiny neurons is modelled, as in (Gurney et al., 2001b), by activation thresholds  $\varepsilon_{D1}$  and  $\varepsilon_{D2}$  under which the neurons remain silent.

The sub-thalamic nucleus (STN) is the second input of the basal ganglia and receives also projections from the GPe. Its glutamatergic neurons have an excitatory effect and project to the GPe and GPi. The resulting input of the STN neuron is given by:

$$I_i^{STN} = S_i - w_{GPe}^{STN} \sum_{j=1}^N y_j^{GPe} \quad (5)$$

The tonic activity of the nucleus is modelled by a negative threshold of the transfer function  $\varepsilon_{STN}$ .

The GPe is inhibitory nucleus, similarly as in the GPR, it receives channel-to-channel afferents from the striatum and a diffuse excitation from the STN:

$$I_i^{GPe} = -w_{D2}^{GPe} y_i^{D2} + w_{STN}^{GPe} \sum_{j=1}^N y_j^{STN} \quad (6)$$

The GPi and SNr are the inhibitory output nuclei of the BG, which keep their targets under inhibition unless a channel is selected. They receive channel-to-channel projections from the D1 striatum and diffuse projections from the STN and the GPe:

$$I_i^{GPi} = -w_{D1}^{GPi} y_i^{D1} + w_{STN}^{GPi} \sum_{j=1}^N y_j^{STN} - w_{GPe}^{GPi} \sum_{j=1}^N y_j^{GPe} \quad (7)$$

This model keeps the basic off-centre on-surround selecting structure, duplicated in the D1-STN-GPi/SNr and D2-STN-GPe sub-circuits, of the GPR. However, the channel specific feedback from the GPe to the Striatum helps sharpening the selection by favouring the channel with the highest salience in D1 and D2. Moreover, the global GPe inhibition on the GPi/SNr synergetically interacts with the STN excitation in order to limit the amplitude of variation of the inhibition of the unselected channels.

## 4 Mathematical results

We first analyse the contraction of the GPR model before showing under which weighting constraints our model is contracting and which sufficient salience input conditions allow it to perform “perfect selection” (output inhibition of selected channels equal to 0).

### 4.1 Contraction analysis of the GPR model

While it is difficult to refute contraction of a system as the metric in which it is contracting is not given *a priori*, we can study contraction in particular metrics for the sake of finding a contra-example which will demonstrate the non-contracting behaviour of the system.

First, remark that lateral connections on striatum ( $D_1$  and  $D_2$ ) make the model non-contracting in the identity metric when the weight of inhibition  $w_{Lat} \geq 1$ . Indeed, by computing directly the eigenvalues of the Jacobian

$$J = \begin{pmatrix} -1 & -w_{Lat} & \cdot & \cdot & -w_{Lat} \\ -w_{Lat} & -1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & -w_{Lat} \\ -w_{Lat} & \cdot & \cdot & -w_{Lat} & -1 \end{pmatrix}$$

we have  $\lambda_{max} \leq -1 + w_{Lat}$ . Unsurprisingly, when  $w_{Lat} = 1$  the system has multiple points of stability and thus the model is not contracting in any metric.

A typical example of multiple points of stability occurs when two channels, say  $i$  and  $j$ , have the same highest salience  $S_{max}$  for input. We then have a continuum of possible stable points in  $D_1$  and  $D_2$  covering the segment  $a_i + a_j = S_{max}$  with  $a_i, a_j \geq 0$ , while all the other channels being fully inhibited.

Such a situation occurs when reproducing the basic selection test proposed in (Gurney et al., 2001b). In this five-steps test (fig. 2), no channels are excited during the first one, and none of them is thus selected; then during the second one, the salience of channel 1 is increased and this channel is consequently selected; during the third one, channel 2 is provided a larger salience than channel 1, channel 1 is thus inhibited and channel 2 selected; in the fourth one, the salience of channel 1 is increased to a value equal to the salience of channel 2, *channel 1 is however not selected while channel 2 remains selected*; finally the salience of channel 1 is decreased to its initial level. Such a drawback can only be solved by reducing  $w_{Lat}$  to a value strictly inferior to 1.

Second, suppose  $w_{Lat}$  is set under 1 to avoid this specific problem, it remains to show that the  $GPe/STN$  loop is contracting. Using the feedback analysis with a scaling metric that dilates the states space of the second system involved (a key tool in the study of many feedbacks)

$$\mathbf{M} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \alpha \mathbf{I} \end{pmatrix}, \alpha > 0$$

makes us compute the maximum singular value of  $G_s$  (see section 2.2):

$$\sigma(G_s) = \max\left(\frac{\alpha}{2}, \frac{1}{2}(-\alpha w_{GPe}^{STN} + \frac{N}{\alpha} w_{STN}^{GPe})\right)$$

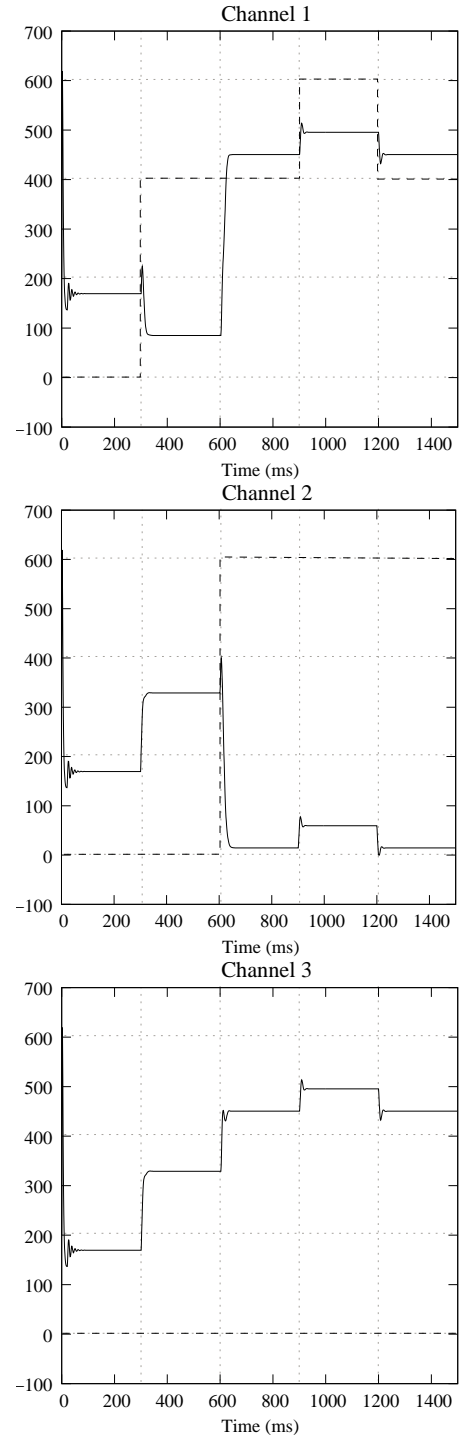


Figure 2: Simulation results (GPI/SNr inhibitory output) for the first three channels of a 6-channels system, using the Gurney et al. (2001b) test on the GPR model. During the period  $900ms < t < 1200ms$ , channels 1 and 2 have the same input saliences, and channel 2 only is selected. Dashed lines represent the input salience of the channel and solid lines represent the output of the channel.

which gives rise to the following condition on  $N$  :

$$N < \frac{4}{w_{STN}^{GPe}}(1 + w_{GPe}^{STN})$$

Analysed in the scaling metric, the contraction of the GPR is proven when  $N$  remains below this bound, which corresponds to  $N < 6$  with the parameters used in (Gurney et al., 2001b). This does not strictly demonstrate that the GPR model with lateral striatal inhibitions lower than 1 is not contracting for  $N \geq 6$ , as there might be another metric in which the analysis would give a contraction result with a different dependence on, or even an independence from,  $N$ . It however suggests that, even if the result is not conclusive, the conditions of contraction of the GPR model probably depend on  $N$ , this is the main motivation for proposing a model whose contraction is proven for less restrictive conditions.

## 4.2 Contraction of the model

The contraction of our model is demonstrated using the combination properties of contracting systems.

First, we see that every nucleus is trivially contracting with a rate  $\frac{1}{\tau}$  as no lateral connection is allowed except for the  $D_i$ 's which are contracting when  $w_{LatD_i} < 1$  with rate  $\frac{1}{\tau} - 1 + w_{LatD_i}$  (see section 4.1). Dealing with thresholds of the leaky-integrator transfer functions is transparent as it is just a particular case of contraction analysis on convex regions (see section 2.2).

Next, defining the system carefully leads to a hierarchical system of trivially contracting systems except for the loops between  $STN/GPe$  and  $D_2/GPe$ . Thus, we only have to master those loops thanks to the feedback combination analysis to guarantee contraction of the whole system.

### STN/GPe

Thanks to our reformulation of the GPe to STN projections (diffuse rather than channel-to-channel), this loop is now contracting as it is a positive/negative feedback. In other word, considering the metric

$$M_1 = \begin{pmatrix} w_{GPe}^{STN} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & w_{STN}^{GPe} \mathbf{I} \end{pmatrix}$$

leads to the generalised Jacobian

$$F = \begin{pmatrix} -\mathbf{I} & (w_{GPe}^{STN} w_{STN}^{GPe})^{\frac{1}{2}} \mathbf{1} \\ -(w_{GPe}^{STN} w_{STN}^{GPe})^{\frac{1}{2}} \mathbf{1} & -\mathbf{I} \end{pmatrix}$$

and the feedback thus disappears as the symmetrical of  $F$  is simply  $-\mathbf{I}$ .

### D2/GPe

The feedback is of the form negative/negative feedback and thus we can just try to minimise the impact of the loop by taking the average of each negative feedback. This is realised by considering the metric

$$M_2 = \begin{pmatrix} w_{D_2}^{GPe} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & w_{GPe}^{D_2} \mathbf{I} \end{pmatrix}$$

which tells us that the system is contracting as long as

$$w_{D_2}^{GPe} w_{GPe}^{D_2} < -1 + w_{LatD_2}$$

The last equation is obtained by using feedback analysis, see section 2.2 for more details.

## 4.3 Analytical results

As our model is contracting, we only have to find a particular solution to be sure that the system will eventually reach this solution. But, because this contracting system is autonomous (time-invariant), we know that this solution is an equilibrium (Slotine, 2003). Thus, it just remains to show that this equilibrium performs the awaited selection.

Naturally, as for GPR, we can show that our model is *order preserving* and that  $\sum_{j=0}^n y_j^{STN}$  is bounded. But more interestingly, we can analytically study our model in the *ideal case* when the stable state is one active neuron only, say  $i_0$ , in  $D_2$  and one inactive in  $GPe$  (necessarily the same  $i_0$ ). We call this situation *ideal case* as the selection is completely performed in the  $D_2 - STN - GPe$  loop and the rest of the model simply copies this selection.

Assuming that the salience input of the system leads to the this particular behaviour, we can obtain the following equations by solving the system of linear equations defined in section 3, using that  $a = I$  for all neurons at equilibrium.

$$\begin{aligned} \sum_{j=1}^N y_j^{STN} &= \frac{\sum_{y_j^{STN} \neq 0} (S_j + \varepsilon_{STN})}{1 + act(N-1)w_{STN}^{GPe}w_{GPe}^{STN}} \\ S_{i_0} &\geq \varepsilon_{D_2} + \frac{w_{STN}^{GPe}}{w_{D_2}^{GPe}} \sum_{j=1}^N y_j^{STN} \\ S_i &\leq \varepsilon_{D_2} + w_{LatD_2}(S_{i_0} - \varepsilon_{D_2}) \\ &+ w_{GPe}^{D_2} w_{STN}^{GPe} \sum_{j=1}^N y_j^{STN} \quad i \neq i_0 \end{aligned}$$

where  $act$  is the number of neurons of the  $STN$  whose activation is larger than  $\varepsilon_{STN}$ . Remark that when  $(N-1)w_{STN}^{GPe}w_{GPe}^{STN} = 1$ ,  $\sum_{j=0}^n y_j^{STN}$  computes essentially the mean of the active saliences.

Those equations give a range of saliences input for which the model reacts ideally, as its equilibrium corresponds to a "perfect selection", where the selected channel is completely disinhibited. Outside this range, the behaviour is more awkward as the whole system is involved in improving the partial selection made by the  $D_2 - STN - GPe$  loop. It might continue to perform "perfect selection", perform a less precise selection or behave differently, hence the simulation of section 5.2 in a wide set of input conditions.

## 5 Simulation results

Similarly to the simulations made by Gurney et al. (2001b), we used a 6-channel model. The parameters were set to the values summarised in table 1.  $w_{LatD_1}$ ,  $w_{LatD_2}$ ,  $w_{D_2}^{GPe}$  and  $w_{GPe}^{D_2}$  were set to values compatible with the constraints needed to ensure the contraction of the system (see 4.2).  $w_{GPe}^{D_1}$  and  $w_{STN}^{GPe}$  were set to values identical to  $w_{GPe}^{D_2}$  and  $w_{STN}^{GPe}$  respectively, for the sake of symmetry, whereas it is not mandatory with regards to contraction. Finally we set  $w_{D_1}^{GPe}$  to 1 rather than to 0.7 (as  $w_{D_2}^{GPe}$ ) in order to favour strong selective inhibitions over GPi and thus "perfect selections".

The simulation was programmed in C++, using the simple Euler approximation for integration, with a time step of 1ms.

Table 1: Parameters of the simulations.

|                 |      |                 |      |                     |        |
|-----------------|------|-----------------|------|---------------------|--------|
| $w_{LatD1}$     | 0.4  | $w_{D2}^{GPe}$  | 0.7  | $\tau$              | 0.003s |
| $w_{LatD2}$     | 0.4  | $w_{D1}^{GPe}$  | 1    | $\lambda$           | 0.2    |
| $w_{GPe}^{D1}$  | 1    | $w_{STN}^{GPe}$ | 0.35 | $\varepsilon_{D1}$  | 200    |
| $w_{GPe}^{D2}$  | 1    | $w_{STN}^{GPe}$ | 0.35 | $\varepsilon_{D2}$  | 200    |
| $w_{GPe}^{STN}$ | 0.35 |                 |      | $\varepsilon_{STN}$ | -150   |
| $w_{GPe}^{GPe}$ | 0.08 |                 |      |                     |        |

## 5.1 Reproduction of GPR basic selection properties

We reproduced the selection experiment of Gurney et al. (2001b), where the system is submitted a sequence of five different salience vectors. As we bounded the activity of our neurons between 0 and 1000, while Gurney et al. had an upper limit of 1, we multiplied by 1000 the input saliences for this test. Each vector is submitted to the system during 0.3s before switching to the next one in the sequence (fig. 3).

First, all saliences are null, and the system stabilises in a situation where all channels are equally inhibited. Then, the first channel receives a 400 input salience which results in perfect disinhibition of this channel ( $y_1^{GPe} = 0$ ) and increased inhibition of the others. When the second channel salience is set to 600, it becomes perfectly selected ( $y_2^{GPe} = 0$ ) while the first one is rapidly inhibited to a level identical to the one of the four last channels. During the fourth step, the salience of the first channel is increased to 600, channels 1 and 2 are therefore simultaneously selected. Finally, during the last step of the test, the salience of channel 1 is reduced to 400, which is then rapidly inhibited while the selection of channel 2 is unaffected.

Our model passes this test in satisfactory manner, its results differ with the GPR in two ways. Firstly, it tends to select channels in a sharper manner than the GPR, as it always reaches “perfect selection” ( $y_i^{GPe} = 0$ ). Secondly, the global level of inhibition in the unselected channels is subject to smaller variations, because of the regulatory effect of balance between the GPe global inhibition and the STN global excitation over the GPi.

## 5.2 1000 random vectors test

In order to test the ability of the model to perform “perfect selection” in a wide range of salience inputs and without any influence of its initial state (a property implied by contraction of the model), we fed a 6-channels system with a sequence of 1000 randomly drawn salience vectors successively. The saliences of each vectors are drawn uniformly in a 0 to 990 interval (discretisation step of 10), equal saliences are authorised within the same vector. Each vector is presented during 0.3s, at the end of this period, the “perfect selection” of the channels with maximum salience is checked along with the presence of perfectly selected channels corresponding to other salience values. Then the next random vector is presented without resetting the system. This test was conducted with our model and with a GPR model for which  $w_{LatDi}$  was set to 0.8.

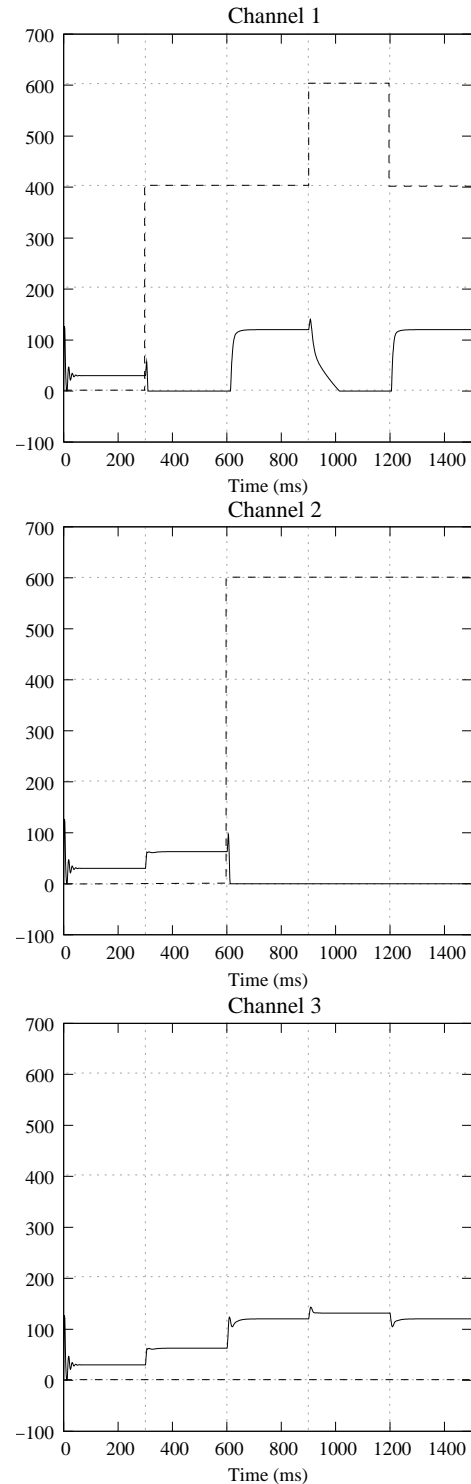


Figure 3: Simulation results (GPi/SNr inhibitory output) for the first three channels of a 6-channels system, using the Gurney et al. (2001b) test. Dashed lines represent the input salience of the channel and solid lines represent the output of the channel.

The first result of the test is that for our model, the “perfect selection” of the channels with maximum salience was not completed in only two cases out of thousand. This occurs when the maximum salience is too low to enable the activity in the striatal neurons to rise above the striatum thresholds  $\varepsilon_{D1} = \varepsilon_{D2} = 200$  and is thus unable to elicit selection, an expected result as these thresholds are thought to filter low level saliences. Concerning the GPR, processing the same 1000-vectors sequence, “perfect selection” was not obtained in 54.6% of the cases, which is quite natural as the GPR is not designed to perform “perfect selection”. The inhibitory output of the GPI/SNr of the GPR model is close 160 when the input salience vector is null, in which case no channel should be selected. We thus chose a value of the  $\theta$  threshold equal to that maximum. In that case, there is no selection in 29.3% of the cases. It seems that in this range of salience input, our model selects winning channels more efficiently than the GPR.

The second result of the test is that the model has a nice property of contrast enhancement, as the maximum can be sorted out from its competitors even if they are quite close, generating a perfect selection of the former and a strong inhibition of the latter. Indeed, simultaneous selection of the channel with maximum salience with one of its competitors happens only in 7.2% of the cases. Moreover, this only happens when the maximal salience value is high ( $\mu = 907.5$ ,  $\sigma = 71.3$ ) and when the difference between the maximal salience and the salience of the supplementary selected channel is low (45 selections with a difference of 10, 24 with a difference of 20, 2 with 30 and 1 with 40). We may thus infer that the limit of discrimination between two saliences of our model is probably inferior to a few percents.

## 6 Discussion

We proposed a new computational model of the basal ganglia exploring how their intrinsic computations operate the physiologically observed “selection by disinhibition” (Chevalier and Deniau, 1990), which is thought to be a fundamental neural substrate of action selection in vertebrates (Redgrave et al., 1999). This model shares a lot of similarities with the previously proposed GPR model (Gurney et al., 2001b), as its selection ability relies on two off-centre on-surround sub-circuits. However, it includes neglected connections from the GPe to the Striatum. Moreover, it distinguishes global projections of the GPe to the STN, GPI and SNr on the one hand and channel-to-channel ones to the Striatum on the other.

We theoretically studied the dynamic behaviour of the network and proved its stability by showing that it is contracting and has an equilibrium point, and thus always converges exponentially fast to this equilibrium. The independence of this contraction with regards to the number of channels results from the diffuse inhibitions from GPe to STN. We also showed that in an *ideal case*, implying conditions on the saliences values, this equilibrium corresponds to a *perfect selection* (where the channel corresponding to the highest salience is completely disinhibited and all others inhibited).

In order to test the selection efficiency of the model in a wider range of input conditions, we reproduced the basic se-

lection test proposed by Gurney et al. (2001b) and, above all, evaluated the quality of selection when it is given a sequence of 1000 random salience vectors. In both cases, *perfect selection* was obtained, except in the rare cases where all the components of the salience vector are too low to elicit selection. Moreover, the selectivity of the model in the second test was better than the GPR.

We modelled the projections from GPe to striatum as having a channel-to-channel selectivity. However, in their study of five pallido-striatal neurons in rats, Bevan et al. (1998) showed that their primary target seems to be the GABAergic interneurons. First, given the limited extend of this study, we cannot exclude the possibility that GPe-striatum projections also concern striatum projection neurons. Second, the GABAergic interneurons inhibit the striatum projection neurons in a relatively diffuse manner, a regulatory effect that is different from but not opposed to our selective and direct projections: it controls the activity of the whole striatum and can thus affect the contrast of the selection. An alternate version of our model derived from these results should be tested.

We omitted two extra types of documented connections. First, the STN projects to the GPe, GPI and SNr but also to the striatum (Parent et al., 2000). Intriguingly, the population of STN neurons projecting to the striatum does not project to the other targets, while the other neurons project to at least two of the other target nuclei. We could not decipher the role of this striatum-projecting population and did not include it in the current model. Its unique targeting specificity suggests it could be functionally distinct from the other STN neurons. This possibility should be explored in future work. The other missing connections concerns the fact that D1 striatal neurons probably simultaneously project to the GPI/SNr and the GPe (Wu et al., 2000), and the fact that lateral inhibition exist in GPe and SNr (Park et al., 1982; Juraska et al., 1977; Deniau et al., 1982). These additional projections were added to the GPR in an improved implementation (Gurney et al., 2004), where the lateral inhibitions of the striatum were also removed. We should add these connections and proceed to a similar test with our model, knowing that the D1-GPe projections would create a new D1-GPe loop and generate an additional constraint on the weights to ensure contraction.

The GPe to striatum connections have the previously evoked functional advantage of enhancing the quality of the selection, by silencing the unselected striatal neurons. Interestingly, the striatum is known for being a relatively silent nucleus (Wilson, 1993), a property supposed to be induced by the specific up/down state behaviour of the striatal neurons. When using simple neuron models, like leaky-integrators, it is usually difficult to reproduce this with a threshold in the transfer function only: when many channels have a strong saliences input, all the corresponding striatal neurons tend to be activated. Our model suggests that in such a case, the GPe-striatum projections may contribute to silencing the striatum.

Finally, the basal ganglia are part of cortico-basal ganglia-thalamo-cortical loops and the quality of selection of the GPR model was improved by the addition of the thalamo-cortical components (Humphries and Gurney, 2002). We plan to extend our model in a similar manner while trying to preserve its contraction properties.

## References

- Alexander, G. E., Crutcher, M. D., and DeLong, M. R. (1990). Basal ganglia-thalamocortical circuits: Parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. *Progress in Brain Research*, 85:119–146.
- Alexander, G. E., DeLong, M. R., and Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience*, 9:357–381.
- Bevan, M., Booth, P., Eaton, S., and Bolam, J. (1998). Selective innervation of neostriatal interneurons by a subclass of neurons in the globus pallidus of rats. *Journal of Neuroscience*, 18(22):9438–9452.
- Chevalier, G. and Deniau, M. (1990). Disinhibition as a basic process of striatal functions. *Trends in Neurosciences*, 13:277–280.
- Deniau, J.-M., Kitai, S., Donoghue, J., and Grofova, I. (1982). Neuronal interactions in the substantia nigra pars reticulata through axon collateral of the projection neurons. *Experimental Brain Research*, 47:105–113.
- Gillies, A. and Arbruthnott, G. (2000). Computational models of the basal ganglia. *Movement Disorders*, 15(5):762–770.
- Girard, B., Cuzin, V., Guillot, A., Gurney, K. N., and Prescott, T. J. (2003). A basal ganglia inspired model of action selection evaluated in a robotic survival task. *Journal of Integrative Neuroscience*, 2(2):179–200.
- Girard, B., Filliat, D., Meyer, J.-A., Berthoz, A., and Guillot, A. (2005). Integration of navigation and action selection in a computational model of cortico-basal ganglia-thalamocortical loops. *Adaptive Behavior: Special Issue on Artificial Rodents*, 13(2):115–130.
- Gurney, K., Humphries, M., Wood, R., Prescott, T., and Redgrave, P. (2004). Testing computational hypotheses of brain systems function: a case study with the basal ganglia. *Network: Computation in Neural Systems*, 15:263–290.
- Gurney, K., Prescott, T. J., and Redgrave, P. (2001a). A computational model of action selection in the basal ganglia. I. A new functional anatomy. *Biological Cybernetics*, 84:401–410.
- Gurney, K., Prescott, T. J., and Redgrave, P. (2001b). A computational model of action selection in the basal ganglia. II. Analysis and simulation of behaviour. *Biological Cybernetics*, 84:411–423.
- Humphries, M. D. and Gurney, K. N. (2002). The role of intra-thalamic and thalamocortical circuits in action selection. *Network: Computation in Neural Systems*, 13:131–156.
- Juraska, J., Wilson, C., and Groves, P. (1977). The substantia nigra of the rat: a golgi study. *Journal of Comparative Neurology*, 172:585–600.
- Kimura, A. and Graybiel, A., editors (1995). *Functions of the Cortico-Basal Ganglia Loop*. Springer, Tokyo/New York.
- Kita, H., Tokuno, H., and Nambu, A. (1999). Monkey globus pallidus external segment neurons projecting to the neostriatum. *Neuroreport*, 10(7):1476–1472.
- Krotopov, J. and Etlinger, S. (1999). Selection of actions in the basal ganglia thalamocortical circuits: Review and model. *International Journal of Psychophysiology*, 31:197–217.
- Lohmiller, W. and Slotine, J. (1998). Contraction analysis for nonlinear systems. *Automatica*, 34(6):683–696.
- Lohmiller, W. and Slotine, J. (2000). Nonlinear process control using contraction analysis. *American Institute of Chemical Engineers Journal*, 46(3):588–596.
- Middleton, F. A. and Strick, P. L. (1994). Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive function. *Science*, 266:458–461.
- Mink, J. W. (1996). The basal ganglia: Focused selection and inhibition of competing motor programs. *Progress in Neurobiology*, 50(4):381–425.
- Montes-Gonzalez, F., Prescott, T. J., Gurney, K. N., Humphries, M., and Redgrave, P. (2000). An embodied model of action selection mechanisms in the vertebrate brain. In Meyer, J.-A., Berthoz, A., Floreano, D., Roitblat, H., and Wilson, S. W., editors, *From Animals to Animals* 6, volume 1, pages 157–166. The MIT Press, Cambridge, MA.
- Parent, A., Sato, F., Wu, Y., Gauthier, J., Lévesque, M., and Parent, M. (2000). Organization of the basal ganglia: the importance of the axonal collateralization. *Trends in Neuroscience*, 23(10):S20–S27.
- Park, M., Falls, W., and Kitai, S. (1982). An intracellular hrp study of rat globus pallidus. i. responses and light microscopic analysis. *Journal of Comparative Neurology*, 211:284–294.
- Redgrave, P., Prescott, T. J., and Gurney, K. (1999). The basal ganglia: a vertebrate solution to the selection problem? *Neuroscience*, 89(4):1009–1023.
- Sato, F., Lavalley, P., Lévesque, M., and Parent, A. (2000). Single-axon tracing study of neurons of the external segment of the globus pallidus in primates. *Journal of Comparative Neurology*, 417:17–31.
- Slotine, J. (2003). Modular stability tools for distributed computation and control. *Journal of Adaptive Control and Signal Processing*, 17(6):397–416.
- Staines, W., Atmadja, S., and Fibiger, H. (1981). Demonstration of a pallidostriatal pathway by retrograde transport of hrp-labelled lectin. *Brain Research*, 206:446–450.
- Wilson, C. (1993). The generation of natural firing patterns in neostriatal neurons. In Arbruthnott, G. and Emson, P., editors, *Chemical signalling in the basal ganglia*, volume 99 of *Progress in Brain Research*, pages 277–297. Elsevier, Oxford.
- Wu, Y., Richard, S., and Parent, A. (2000). The organization of the striatal output system: a single-cell juxtacellular labeling study in the rat. *Neuroscience Research*, 38:49–62.