Neural Networks 21 (2008) 628–641

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

2008 Special Issue

Where neuroscience and dynamic system theory meet autonomous robotics: A contracting basal ganglia model for action selection

B. Girard [a,*], N. Tabareau [a], Q.C. Pham [a], A. Berthoz [a], J.-J. Slotine [b]

[a] *Laboratoire de Physiologie de la Perception et de l'Action, UMR7152, CNRS - Collège de France, 11 place Marcelin Berthelot, 75231 Paris Cedex 05, France*

[b] *Nonlinear Systems Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

ARTICLE INFO

Article history:
Received 15 March 2007
Received in revised form
7 March 2008
Accepted 7 March 2008

Keywords:
Action selection
Basal ganglia
Computational model
Autonomous robotics
Contraction analysis

ABSTRACT

Action selection, the problem of choosing what to do next, is central to any autonomous agent architecture. We use here a multi-disciplinary approach at the convergence of neuroscience, dynamical system theory and autonomous robotics, in order to propose an efficient action selection mechanism based on a new model of the basal ganglia. We first describe new developments of contraction theory regarding locally projected dynamical systems. We exploit these results to design a stable computational model of the cortico-baso-thalamo-cortical loops. Based on recent anatomical data, we include usually neglected neural projections, which participate in performing accurate selection. Finally, the efficiency of this model as an autonomous robot action selection mechanism is assessed in a standard survival task. The model exhibits valuable dithering avoidance and energy-saving properties, when compared with a simple if-then-else decision rule.

1. Introduction

Action selection is the problem of motor resource allocation an autonomous agent is faced with, when attempting to achieve its long-term objectives. These may vary from survival and reproduction to delivering letters to researchers' offices, depending on the nature of the considered agent (animal, robot, etc.). Action selection is a topic of interest in various disciplines, including ethology, artificial intelligence, psychology, neuroscience, autonomous robotics, etc. We address here the question of action selection for an autonomous robot, using a computational model of brain regions involved in action selection, namely the cortico-baso-thalamo-cortical loops. In order to avoid unwanted dynamical behaviors resulting from a highly recurrent network, we use contraction analysis (Lohmiller & Slotine, 1998) to obtain a rigorous proof of its stability. The efficiency of this action selection mechanism (ASM) is assessed using a standard minimal survival task in a robotic simulation.

The basal ganglia are a set of interconnected subcortical nuclei common to all vertebrates and involved in numerous processes, from motor functions to cognitive ones (Middleton & Strick, 1994; Mink, 1996). Their role is interpreted as a generic selection circuit, and they have been proposed to form the neural substrate of action selection (Krotopov & Etlinger, 1999; Mink, 1996; Redgrave, Prescott, & Gurney, 1999). The basal ganglia are included in cortico-baso-thalamo-cortical loops (Fig. 1), five main loops have been identified in primates (Alexander, Crutcher, & DeLong, 1990; Alexander, DeLong, & Strick, 1986; Kimura & Graybiel, 1995): one motor, one oculomotor, two prefrontal and one limbic loop. Within each of these loops, the basal ganglia circuitry is organized in interacting channels, among which selection occurs. Depending on the considered loop, this selection may concern, for example, the target of an upcoming saccadic movement, the target of a reaching movement or the piece of information to be stored in working memory. The output nuclei of the basal ganglia are inhibitory and tonically active, and thus maintain their targets under sustained inhibition. Selection occurs *via* disinhibition (Chevalier & Deniau, 1990): the removal of the inhibition exerted by one channel on its specific target circuit allows the activation of that circuit. When considering action selection, the basal ganglia channels are thought to be associated to competing action primitives. 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.

The considered network contains a large number of closed loops, from the large cortico-baso-thalamo-cortical loop, to small loops formed by the interconnections between nuclei within the basal ganglia and between the thalamus and the cortex. A system with such a structure may exhibit varied dynamical behaviors, some of which should be avoided by an ASM, like reaching a standstill state which does not depend anymore on the

* Corresponding author. Tel.: +33 1 44 27 13 91; fax: +33 1 44 27 13 82.
E-mail address: benoit.girard@college-de-france.fr (B. Girard).

0893-6080/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2008.03.009

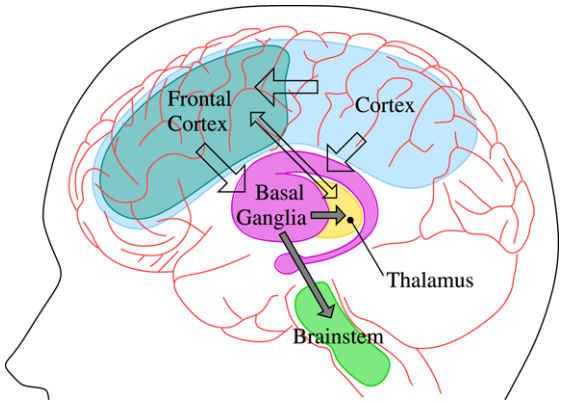

Fig. 1. Cortico-baso-thalamo-cortical loops. The basal ganglia receive inputs from the whole cortex, but establish loops with the frontal areas only. Shaded arrows: inhibitory projections.

external input. This motivates the use of a theoretical framework to study the dynamics of basal ganglia models. We propose to use contraction analysis (Lohmiller & Slotine, 1998) in order to guide the design of a new model of the basal ganglia whose stability can be formally established. Contraction analysis is a theoretical tool used to study the dynamic behavior of nonlinear systems. Contraction properties are preserved through a number of particular combinations, which is useful for a modular design of models.

Numerous computational models of the BG have been proposed in order to investigate the details of the operation of the basal ganglia disinhibition process (see Gillies & Arbruthnott, 2000; Gurney, Prescott, Wickens, & Redgrave, 2004, for recent reviews). Among these, the model proposed by Gurney, Prescott, and Redgrave (2001a, 2001b) (henceforth the GPR model) has been successfully tested as an action selection mechanism for autonomous agents (Girard, Cuzin, Guillot, Gurney, & Prescott, 2003; Girard, Filliat, Meyer, Berthoz, & Guillot, 2005; Montes-Gonzalez, Prescott, Gurney, Humphries, & Redgrave, 2000; Prescott, Montes-Gonzalez, Gurney, Humphries, & Redgrave, 2006). In particular, it was shown to be able to solve a minimal survival task, and, compared with a simpler winner-takes-all mechanism, displayed dithering avoidance and energy-saving capabilities.

We present here an action selection mechanism based on a contracting computational model of the basal ganglia (or CBG). In order to adapt the contraction theory to the analysis of rate-coding artificial neural networks, we first extend it to locally projected dynamical systems (Section 2). Using the resulting neuron model and contraction constraints on the model's parameters, we build a computational model of the basal ganglia including usually neglected neural connections (Section 3). We then check the selection properties of the disembodied model and compare them to those of the GPR, so as to emphasize the consequences of using contraction analysis (Section 4). We finally test its efficiency in a survival task similar to the one used to evaluate the GPR (Girard et al., 2003), and emphasize its dithering avoidance and energy-saving properties by comparing it to a simple if-then-else decision rule (Section 5).

Preliminary versions of the basal ganglia computational model were presented in Girard, Tabareau, Berthoz, and Slotine (2006) and Girard, Tabareau, Slotine, and Berthoz (2005).

2. Nonlinear contraction analysis for rate-coding neural networks

Basically, a nonlinear time-varying dynamic system is said to be *contracting* if initial conditions or temporary disturbances are forgotten exponentially fast, that is, if any perturbed trajectory returns to its nominal behavior with an exponential convergence rate. Contraction is an extension of the well-known *stability* analysis for linear systems. It has the desirable feature of being preserved through hierarchical and particular feedback combinations. Thus, as we will see below, contraction analysis is an appropriate tool to study stability properties of rate-coding neural networks.

In addition, when a system is contracting, it is sufficient to find a particular bounded trajectory to be sure that the system will eventually tend to this trajectory. Thus contraction theory is a convenient way to analyze the dynamic behavior of a system without linearized approximations.

2.1. Contraction theory

We summarize the differential formulation of contraction analysis presented in Lohmiller and Slotine (1998). Contraction analysis is a way to prove the exponential stability of a nonlinear system by studying the properties of its Jacobian. Consider an n-dimensional time-varying system of the form:

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

where $\mathbf{x} \in \mathbb{R}^n$ and $t \in \mathbb{R}_+$ and $\mathbf{f}$ is a $n \times 1$ nonlinear vector function which is assumed in the remainder of this paper to be real and smooth, in the sense that all required derivatives exist and are continuous. This equation may also represent the closed-loop dynamics of a neural network model of a brain structure. We recall below the main result of contraction analysis (see Lohmiller and Slotine (1998), for a proof and more details).

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

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

such that the generalized Jacobian

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

is uniformly negative definite, then 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 symmetric part of a matrix $\mathbf{A}$ is $\mathbf{A}_s = 1/2(\mathbf{A} + \mathbf{A}^{\mathrm{T}})$. A matrix $\mathbf{A}(\mathbf{x}, t)$ is uniformly positive definite if there exists $\beta > 0$ such that

$$\forall \mathbf{x}, t \quad \lambda_{\min}(\mathbf{A}(\mathbf{x}, t)) \geq \beta.$$

2.2. Neural networks and locally projected dynamical systems

Networks of leaky integrators are widely used to model the behavior of neuronal assemblies (Dayan & Abbott, 2001). A leaky-integrator network is usually described by the following set of equations

$$\tau_i \dot{x}_i = -x_i(t) + \sum_{j \neq i} K_{ji} x_j(t) + I(t)$$

where $x(t)$ is the synaptic current of a neuron, τ_i its time constant, K_{ji} the synaptic projection weight from neuron j to neuron i and $I(t)$ the input coming from an external source. Next, $x(t)$ is converted into a non-negative firing rate $y(t)$ using a transfer function, for instance

$$y(t) = \max(x(t), 0) = [x(t)]_+.$$

Another way to enforce non-negativity of the firing rate is to use through *locally projected dynamical systems* (lPDS in short). These systems were introduced in Dupuis and Nagurney (1993)

and further analyzed in Zhang and Nagurney (1995). Related ideas can be found in the standard parameter projection method in adaptive control (Ioannou & Sun, 1996; Slotine & Coetsee, 1986). A lPDS is given by

$$\dot{\mathbf{x}} = \mathbf{\Pi}_{\mathbf{\Omega}}(\mathbf{x}, \mathbf{f}(\mathbf{x}, t)) \tag{2}$$

where $\mathbf{\Omega}$ is a convex subset of the state space and $\mathbf{\Pi}_{\mathbf{\Omega}}$ is the vector-projection operator on $\mathbf{\Omega}$ given by

$$\mathbf{\Pi}_{\mathbf{\Omega}}(\mathbf{x}, \mathbf{v}) = \lim_{h \to 0^+} \frac{\mathbf{P}_{\mathbf{\Omega}}(\mathbf{x} + h\mathbf{v}) - \mathbf{x}}{h}.$$

In the above equation, $\mathbf{P}_{\mathbf{\Omega}}$ denotes the point-projection operator on the convex $\mathbf{\Omega}$ defined as

$$\mathbf{P}_{\mathbf{\Omega}}(\mathbf{x}) = \operatorname*{argmin}_{\mathbf{y} \in \mathbf{\Omega}} \|\mathbf{x} - \mathbf{y}\|.$$

Intuitively, if $\mathbf{x}$ is in the interior of $\mathbf{\Omega}$ then $\mathbf{\Pi}_{\mathbf{\Omega}}(\mathbf{x}, \mathbf{v}) = \mathbf{v}$. If $\mathbf{x}$ is on the boundary of $\mathbf{\Omega}$, then $\mathbf{\Pi}_{\mathbf{\Omega}}(\mathbf{x}, \mathbf{v})$ is the maximal component of $\mathbf{v}$ that allows the system to remain within $\mathbf{\Omega}$. In particular, it is easy to see that any trajectory starting in $\mathbf{\Omega}$ remains in $\mathbf{\Omega}$.

Note that Eq. (2) does not define a classical ordinary differential equation since its right-hand side can be discontinuous due to the projection operator. However, under some conditions on $\mathbf{f}$ and $\mathbf{\Omega}$ (similar to the Cauchy–Lipschitz conditions for classical ordinary differential equations, see Dupuis and Nagurney (1993) and Filippov (1988) for more details), existence, uniqueness and some qualitative properties can be established for the solutions of (2). For our purpose, we recall here that any solution $\mathbf{x}$ of (2) is continuous and right differentiable for all t. In the remainder of this article, we make the additional assumption that the set of time instants when $\mathbf{x}(t)$ is not differentiable has measure zero.

Within the above framework, the dynamics of a neural network can now be given in the matrix form as

$$\dot{\mathbf{x}} = \mathbf{\Pi}_{\mathbb{H}_n}(\mathbf{x}, \mathbf{W}\mathbf{x} + \mathbf{I}(t)) \tag{3}$$

where $\mathbf{x}(t) = (x_1(t), \ldots, x_n(t))^{\mathrm{T}}$ is the states of the neurons, $\mathbf{W}$ is the $n \times n$ matrix whose diagonal elements represent the leaking rate of the neurons and whose non-diagonal elements represent the synaptic projection weight, $\mathbf{I}(t)$ is the vector of external inputs. Finally, $\mathbb{H}_n$ is a *regular n-cube* defined as follows

Definition 1. A *regular n-cube* $\mathbb{H}_n$ is a subset of $\mathbb{R}^n$ defined by

$$\mathbb{H}_n = \{(x_1, \ldots, x_n)^{\mathrm{T}} \in \mathbb{R}^n : \forall i,\ m_i \le x_i \le M_i\}$$

where $m_1, \ldots, m_n, M_1, \ldots, M_n \in \mathbb{R}$.

Intuitively, a regular n-cube is an n-cube whose edges are parallel to the axes.

In practice, networks of leaky integrators described by lPDS as above and their classical counterparts with transfer functions show very similar behavior. However, the stability properties of lPDS networks can be rigorously established through contraction theory (see the next section), which makes them interesting from a theoretical viewpoint.

2.3. Contraction analysis of locally projected dynamical system on regular n-cubes

Contraction analysis for systems subject to convex constraints has already been discussed in Lohmiller and Slotine (2000). However, in that work, the projection applied to constrain the system in the convex region depends on the metric which makes the original system contracting. Thus, we cannot use this result here since our projection operator must not depend on the neural network

Since the contraction condition is local, a lPDS can only be contracting if the original, un-projected, system is contracting within $\mathbf{\Omega}$. The converse implication is not true in general, because the projection operator can deeply modify the system's behavior along the boundary of $\mathbf{\Omega}$. We now introduce some definitions in order to be able to state this converse implication in some particular cases.

Definition 2. Let $\mathbf{x} \in \delta\mathbf{\Omega}$ where $\delta\mathbf{\Omega}$ denotes the boundary of $\mathbf{\Omega}$. The set of *inward normals* to $\mathbf{\Omega}$ at $\mathbf{x}$ is defined as

$$N_{\mathbf{\Omega}}(\mathbf{x}) = \{\mathbf{n} : \forall \mathbf{y} \in \mathbf{\Omega}, \mathbf{n}^{\mathrm{T}}(\mathbf{x} - \mathbf{y}) \le 0\}.$$

If $\mathbf{x} \in \mathbf{\Omega} - \delta\mathbf{\Omega}$ then we set $N_{\mathbf{\Omega}}(\mathbf{x}) = \{\mathbf{0}\}$.

Definition 3. A metric $\mathbf{M}$ is said to be *compatible* with a convex set $\mathbf{\Omega}$ if there exists a coordinate transform $\mathbf{\Theta}$ such that $\mathbf{\Theta}^{\mathrm{T}}\mathbf{\Theta} = \mathbf{M}$ and

$$\forall \mathbf{x} \in \delta\mathbf{\Omega}, \forall \mathbf{n} \in N_{\mathbf{\Omega}}(\mathbf{x}), \quad \mathbf{\Theta}\mathbf{n} \in N_{\mathbf{\Theta}\mathbf{\Omega}}(\mathbf{\Theta}\mathbf{x}).$$

In this case, we say that $\mathbf{\Theta}$ is a square root of $\mathbf{M}$ which is *compatible* with $\mathbf{\Omega}$.

We can give a simple sufficient condition for a metric to be compatible with a regular n-cube.

Proposition 1. *Any diagonal positive definite metric* $\mathbf{M}$ *is compatible with any regular n-cube* $\mathbb{H}_n$.

Proof. Let $\mathbf{x} = (x_1, \ldots, x_n)^{\mathrm{T}} \in \delta\mathbb{H}_n$. An inward normal $\mathbf{n} = (n_1, \ldots, n_n)^{\mathrm{T}}$ to $\mathbb{H}_n$ at $\mathbf{x}$ is characterized by

$$\begin{cases} n_i \ge 0 & \text{if } x_i = m_i \\ n_i \le 0 & \text{if } x_i = M_i \\ n_i = 0 & \text{if } m_i < x_i < M_i. \end{cases}$$

Since $\mathbf{M}$ is diagonal and positive definite, one has $\mathbf{M} = \operatorname{diag}(d_1^2, \ldots, d_n^2)$ with $d_i > 0$. Consider the coordinate transform $\mathbf{\Theta} = \operatorname{diag}(d_1, \ldots, d_n)$. Clearly, $\mathbf{\Theta}^{\mathrm{T}}\mathbf{\Theta} = \mathbf{M}$ and $\mathbf{\Theta}\mathbb{H}_n$ is a regular n-cube with minimal values $d_1 m_1, \ldots, d_n m_n$ and maximal values $d_1 M_1, \ldots, d_n M_n$. It follows from the characterization above that $\mathbf{\Theta}\mathbf{n} = (d_1 n_1, \ldots, d_n n_n)^{\mathrm{T}} \in N_{\mathbf{\Theta}\mathbb{H}_n}(\mathbf{\Theta}\mathbf{x})$. □

We also need another elementary result.

Lemma 1. *Let* $\mathbf{x} \in \mathbf{\Omega}$ *and* $\mathbf{v} \in \mathbb{R}^n$. *There exists* $\mathbf{n}(\mathbf{x}, \mathbf{v}) \in N_{\mathbf{\Omega}}(\mathbf{x})$ *such that*

$$\mathbf{\Pi}_{\mathbf{\Omega}}(\mathbf{x}, \mathbf{v}) = \mathbf{v} + \mathbf{n}(\mathbf{x}, \mathbf{v}).$$

Proof. Let $\mathbf{y} \in \mathbf{\Omega}$. We need to show that $A_{\mathbf{y}} = (\mathbf{\Pi}_{\mathbf{\Omega}}(\mathbf{x}, \mathbf{v}) - \mathbf{v})^{\mathrm{T}}(\mathbf{x} - \mathbf{y}) \le 0$. By the definition of $\mathbf{\Pi}_{\mathbf{\Omega}}$, one has

$$A_{\mathbf{y}} = \lim_{h \to 0^+} \frac{1}{h}(\mathbf{P}_{\mathbf{\Omega}}(\mathbf{x} + h\mathbf{v}) - (\mathbf{x} + h\mathbf{v}))^{\mathrm{T}}(\mathbf{x} - \mathbf{y}).$$

Next, introduce the terms $\mathbf{P}_{\mathbf{\Omega}}(\mathbf{x} + h\mathbf{v})$ and $h\mathbf{v}$ into $(\mathbf{x} - \mathbf{y})$

$$\begin{aligned} A_{\mathbf{y}} = \lim_{h \to 0^+} \frac{1}{h}[&(\mathbf{P}_{\mathbf{\Omega}}(\mathbf{x} + h\mathbf{v}) - (\mathbf{x} + h\mathbf{v}))^{\mathrm{T}}(\mathbf{P}_{\mathbf{\Omega}}(\mathbf{x} + h\mathbf{v}) - \mathbf{y}) \\ &+ (\mathbf{P}_{\mathbf{\Omega}}(\mathbf{x} + h\mathbf{v}) - (\mathbf{x} + h\mathbf{v}))^{\mathrm{T}}(\mathbf{x} + h\mathbf{v} - \mathbf{P}_{\mathbf{\Omega}}(\mathbf{x} + h\mathbf{v})) \\ &+ (\mathbf{P}_{\mathbf{\Omega}}(\mathbf{x} + h\mathbf{v}) - (\mathbf{x} + h\mathbf{v}))^{\mathrm{T}}(-h\mathbf{v})]. \end{aligned}$$

The first term in the above equation is non-positive by the property of the point-projection operator. The second term is the negative of a distance and thus is also non-positive. As for the third term, observe that

$$\lim_{h \to 0^+} (\mathbf{P}_{\mathbf{\Omega}}(\mathbf{x} + h\mathbf{v}) - (\mathbf{x} + h\mathbf{v}))^{\mathrm{T}}\mathbf{v} = (\mathbf{P}_{\mathbf{\Omega}}(\mathbf{x}) - \mathbf{x})^{\mathrm{T}}\mathbf{v} = 0$$

since $\mathbf{x} \in \mathbf{\Omega}$. □

We can now state the following theorem

Theorem 2. *Let $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$ be a dynamical system which is contracting in a constant metric $\mathbf{M}$ compatible with a convex set $\mathbf{\Omega}$. Then the lPDS $\dot{\mathbf{x}} = \mathbf{\Pi}_{\mathbf{\Omega}}(\mathbf{x}, \mathbf{f}(\mathbf{x}, t))$ is also contracting in the same metric and with the same contraction rate.*

Proof. Let $\mathbf{\Theta}$ be a square root of $\mathbf{M}$ compatible with $\mathbf{\Omega}$. Consider $\mathbf{z} = \mathbf{\Theta}\mathbf{x}$. By Lemma 1, the system $\mathbf{z}$ is described by

$$\dot{\mathbf{z}} = \mathbf{\Theta}\mathbf{\Pi}_{\mathbf{\Omega}}(\mathbf{x}, \mathbf{f}(\mathbf{x})) = \mathbf{F}(\mathbf{z}) + \mathbf{\Theta}\mathbf{n}(\mathbf{x}, \mathbf{f}(\mathbf{x})) \tag{4}$$

where $\mathbf{F}(\mathbf{z}) = \mathbf{\Theta}\mathbf{f}(\mathbf{\Theta}^{-1}\mathbf{z})$.

Consider two particular trajectories of (4) $\mathbf{z}_1$ and $\mathbf{z}_2$. Denote by Δ the squared distance between $\mathbf{z}_1$ and $\mathbf{z}_2$

$$\Delta(t) = \|\mathbf{z}_1(t) - \mathbf{z}_2(t)\|^2 = (\mathbf{z}_1(t) - \mathbf{z}_2(t))^{\mathrm{T}}(\mathbf{z}_1(t) - \mathbf{z}_2(t)).$$

When Δ is differentiable, we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\Delta &= 2(\mathbf{z}_1 - \mathbf{z}_2)^{\mathrm{T}}(\dot{\mathbf{z}}_1 - \dot{\mathbf{z}}_2)\\ &= 2(\mathbf{z}_1 - \mathbf{z}_2)^{\mathrm{T}}(\mathbf{F}(\mathbf{z}_1) + \mathbf{\Theta}\mathbf{n}(\mathbf{x}_1, \mathbf{f}(\mathbf{x}_1)) - (\mathbf{F}(\mathbf{z}_2)\\ &\quad + \mathbf{\Theta}\mathbf{n}(\mathbf{x}_2, \mathbf{f}(\mathbf{x}_2)))).\end{aligned}$$

Since the metric is compatible with $\mathbf{\Omega}$, $\mathbf{\Theta}\mathbf{n}(\mathbf{x}_i, \mathbf{f}(\mathbf{x}_i)) \in N_{\mathbf{\Theta}\mathbf{\Omega}}(\mathbf{z}_i)$ for $i = 1, 2$. Next, by the definition of inward normals, we have $(\mathbf{z}_1 - \mathbf{z}_2)^{\mathrm{T}}\mathbf{\Theta}\mathbf{n}(\mathbf{x}_1, \mathbf{f}(\mathbf{x}_1)) \leq 0$ and $-(\mathbf{z}_1 - \mathbf{z}_2)^{\mathrm{T}}\mathbf{\Theta}\mathbf{n}(\mathbf{x}_2, \mathbf{f}(\mathbf{x}_2)) \leq 0$, from which we deduce

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\Delta &\leq 2(\mathbf{z}_1 - \mathbf{z}_2)^{\mathrm{T}}(\mathbf{F}(\mathbf{z}_1) - \mathbf{F}(\mathbf{z}_2))\\ &\leq -2\lambda\Delta(t)\end{aligned}$$

where $\lambda > 0$ is the contraction rate of $\mathbf{f}$ in the metric $\mathbf{M}$.

Since the set of time instants when $\Delta(t)$ is not differentiable has measure zero (see Section 2.2), one has

$$\forall t \geq 0, \quad \Delta(t) = \int_0^t \left(\frac{\mathrm{d}}{\mathrm{d}t}\Delta\right)\mathrm{d}t \leq -2\lambda\int_0^t \Delta(s)\mathrm{d}s$$

which yields by Grönwall's lemma

$$\forall t \geq 0, \quad \Delta(t) \leq \Delta(0)\mathrm{e}^{-2\lambda t}$$

i.e.

$$\forall t \geq 0, \quad \|\mathbf{z}_1(t) - \mathbf{z}_2(t)\| \leq \|\mathbf{z}_1(0) - \mathbf{z}_2(0)\|\mathrm{e}^{-\lambda t}. \quad \square$$

2.4. Combination of contracting systems

One of our motivations for using contraction theory is that contraction properties are preserved under suitable combinations (Lohmiller & Slotine, 1998). This allows both stable aggregation of contracting systems, and variation or optimization of individual subsystems while preserving overall functionality (Slotine & Lohmiller, 2001). We present here three standard combinations of contracting systems which preserve both contraction of the system and diagonality of the metric. Then, constructing our neural network as a lPDS using only those three combinations will give rise to a contracting system in a diagonal metric.

2.4.1. Negative feedback combination

Consider two coupled systems

$$\dot{\mathbf{x}}_1 = \mathbf{f}_1(\mathbf{x}_1, \mathbf{x}_2, t)$$
$$\dot{\mathbf{x}}_2 = \mathbf{f}_2(\mathbf{x}_1, \mathbf{x}_2, t).$$

Assume that system i ($i = 1, 2$) is contracting with respect to $\mathbf{M}_i = \mathbf{\Theta}_i^{\mathrm{T}}\mathbf{\Theta}_i$, with rate λ_i. Assume furthermore that the two systems are connected by *negative feedback* (Tabareau & Slotine, 2006). More precisely, the Jacobian matrices of the couplings verify

$$\mathbf{\Theta}_1\mathbf{J}_{12}\mathbf{\Theta}_2^{-1} = -k\mathbf{\Theta}_2\mathbf{J}_{21}^{\mathrm{T}}\mathbf{\Theta}_1^{-1}$$

with k a positive constant. Hence, the Jacobian matrix of the unperturbed global system is given by

$$\mathbf{J} = \begin{pmatrix}\mathbf{J}_1 & -k\mathbf{\Theta}_1^{-1}\mathbf{\Theta}_2\mathbf{J}_{21}^{\mathrm{T}}\mathbf{\Theta}_1^{-1}\mathbf{\Theta}_2\\ \mathbf{J}_{21} & \mathbf{J}_2\end{pmatrix}.$$

Consider the coordinate transform

$$\mathbf{\Theta} = \begin{pmatrix}\mathbf{\Theta}_1 & \mathbf{0}\\ \mathbf{0} & \sqrt{k}\mathbf{\Theta}_2\end{pmatrix}$$

associated with the metric $\mathbf{M} = \mathbf{\Theta}^{\mathrm{T}}\mathbf{\Theta} > \mathbf{0}$. After some calculations, one has

$$\begin{aligned}\left(\mathbf{\Theta}\mathbf{J}\mathbf{\Theta}^{-1}\right)_s &= \begin{pmatrix}\left(\mathbf{\Theta}_1\mathbf{J}_1\mathbf{\Theta}_1^{-1}\right)_s & \mathbf{0}\\ \mathbf{0} & \left(\mathbf{\Theta}_2\mathbf{J}_2\mathbf{\Theta}_2^{-1}\right)_s\end{pmatrix}\\ &\leq \max(-\lambda_1, -\lambda_2)\mathbf{I}.\end{aligned} \tag{5}$$

The augmented system is thus contracting with respect to the metric $\mathbf{M}$, with rate $\min(\lambda_1, \lambda_2)$.

2.4.2. Hierarchical combination

We first recall a standard result in matrix analysis (Horn & Johnson, 1985). Let $\mathbf{A}$ be symmetric matrix in the form

$$\mathbf{A} = \begin{pmatrix}\mathbf{A}_1 & \mathbf{A}_{21}^{\mathrm{T}}\\ \mathbf{A}_{21} & \mathbf{A}_2\end{pmatrix}.$$

Assume that $\mathbf{A}_1$ and $\mathbf{A}_2$ are positive definite. Then $\mathbf{A}$ is positive definite if

$$\sigma^2(\mathbf{A}_{21}) < \lambda_{\min}(\mathbf{A}_1)\lambda_{\min}(\mathbf{A}_2)$$

where $\sigma(\mathbf{A}_{21})$ denotes the largest singular value of $\mathbf{A}_{21}$. In this case, the smallest eigenvalue of $\mathbf{A}$ satisfies

$$\begin{aligned}\lambda_{\min}(\mathbf{A}) \geq\ & \frac{\lambda_{\min}(\mathbf{A}_1) + \lambda_{\min}(\mathbf{A}_2)}{2}\\ & - \sqrt{\left(\frac{\lambda_{\min}(\mathbf{A}_1) - \lambda_{\min}(\mathbf{A}_2)}{2}\right)^2 + \sigma^2(\mathbf{A}_{21})}.\end{aligned}$$

Consider now the same set-up as in Section 2.4.1, except that the connection is now *hierarchical* and upper bounded. More precisely, the Jacobians of the couplings verify

$$\mathbf{J}_{12} = \mathbf{0}, \quad \sigma^2(\mathbf{\Theta}_2\mathbf{J}_{21}\mathbf{\Theta}_1^{-1}) \leq K.$$

Hence, the Jacobian matrix of the augmented system is given by

$$\mathbf{J} = \begin{pmatrix}\mathbf{J}_1 & \mathbf{0}\\ \mathbf{J}_{21} & \mathbf{J}_2\end{pmatrix}.$$

Consider the coordinate transform

$$\mathbf{\Theta}_\epsilon = \begin{pmatrix}\mathbf{\Theta}_1 & \mathbf{0}\\ \mathbf{0} & \epsilon\mathbf{\Theta}_2\end{pmatrix}$$

associated with the metric $\mathbf{M}_\epsilon = \mathbf{\Theta}_\epsilon^{\mathrm{T}}\mathbf{\Theta}_\epsilon > \mathbf{0}$. After some calculations, one has

$$\left(\mathbf{\Theta}\mathbf{J}\mathbf{\Theta}^{-1}\right)_s = \begin{pmatrix}\left(\mathbf{\Theta}_1\mathbf{J}_1\mathbf{\Theta}_1^{-1}\right)_s & \frac{1}{2}\epsilon(\mathbf{\Theta}_2\mathbf{J}_{21}\mathbf{\Theta}_1^{-1})^{\mathrm{T}}\\ \frac{1}{2}\epsilon\mathbf{\Theta}_2\mathbf{J}_{21}\mathbf{\Theta}_1^{-1} & \left(\mathbf{\Theta}_2\mathbf{J}_2\mathbf{\Theta}_2^{-1}\right)_s\end{pmatrix}.$$

Set now $\epsilon = \sqrt{\frac{2\lambda_1\lambda_2}{K}}$. The augmented system is then contracting with respect to the metric $\mathbf{M}_\epsilon$, with rate λ verifying

$$\lambda \geq \frac{1}{2}\left(\lambda_1 + \lambda_2 - \sqrt{\lambda_1^2 + \lambda_2^2}\right).$$

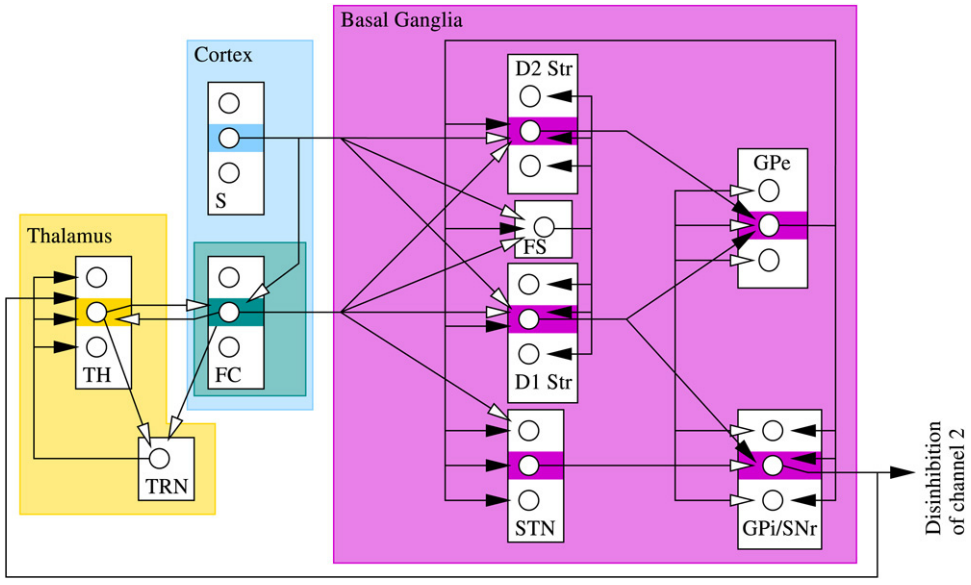

Fig. 2. Basal ganglia model. Nuclei are represented by boxes, each circle in these nuclei represents an artificial rate-coding neuron. In this diagram, three channels are competing for selection, represented by the three neurons in each nucleus. The second channel is represented by colored shading. For 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: subthalamic nucleus; GPe: external segment of the globus pallidus; GPi/SNr: internal segment of the globus pallidus and substantia nigra pars reticulata.

2.4.3. Small gains

In this section, we require no specific assumption on the form of the couplings

$$\mathbf{J} = \begin{pmatrix} \mathbf{J}_1 & \mathbf{J}_{12} \\ \mathbf{J}_{21} & \mathbf{J}_2 \end{pmatrix}.$$

As for negative feedback, consider the coordinate transform

$$\boldsymbol{\Theta}_k = \begin{pmatrix} \boldsymbol{\Theta}_1 & \mathbf{0} \\ \mathbf{0} & \sqrt{k}\boldsymbol{\Theta}_2 \end{pmatrix} \quad k > 0$$

associated with the metric $\mathbf{M}_k = \boldsymbol{\Theta}_k^{\mathrm{T}}\boldsymbol{\Theta}_k > \mathbf{0}$. After some calculations, one has

$$\left(\boldsymbol{\Theta}_k \mathbf{J} \boldsymbol{\Theta}_k^{-1}\right)_s = \begin{pmatrix} \left(\boldsymbol{\Theta}_1 \mathbf{J}_1 \boldsymbol{\Theta}_1^{-1}\right)_s & \mathbf{A}_k^{\mathrm{T}} \\ \mathbf{A}_k & \left(\boldsymbol{\Theta}_2 \mathbf{J}_2 \boldsymbol{\Theta}_2^{-1}\right)_s \end{pmatrix}$$

where $\mathbf{A}_k = \frac{1}{2}\left(\sqrt{k}\boldsymbol{\Theta}_2 \mathbf{J}_{21} \boldsymbol{\Theta}_1^{-1} + \frac{1}{\sqrt{k}}\left(\boldsymbol{\Theta}_1 \mathbf{J}_{12} \boldsymbol{\Theta}_2^{-1}\right)^{\mathrm{T}}\right)$. Following the result stated at the beginning of Section 2.4.2, if

$$\min_k \sigma^2(\mathbf{A}_k) < \lambda_1 \lambda_2$$

then the augmented system is contracting with respect to the metric $\mathbf{M}_k$ for some k, with rate λ verifying

$$\lambda \geq \frac{\lambda_1 + \lambda_2}{2} - \sqrt{\left(\frac{\lambda_1 - \lambda_2}{2}\right)^2 + \min_k \sigma^2(\mathbf{A}_k)}.$$

3. Model description

Rather than using standard leaky-integrator rate-coding neurons, we use the very similar local projected dynamical system model defined by Eq. (3), where each component of the state vector $\mathbf{x}$ is an artificial rate-coding neuron representing the discharge rate of populations of real neurons. Each competing BG channel in each nucleus is represented by one such neuron, and the corresponding thalamic nucleus and cortical areas are also subdivided into identical channels (Fig. 2). The convergence of cortical sensory inputs on the striatum channels is encoded, for simplicity, by a vector of saliences (one salience per channel). Each salience represents the propensity of its corresponding channel to be selected. Each behavior in competition is associated to a specific channel and can be executed if and only if its level of inhibition decreases below a the inhibition level at rest $y^{\mathrm{GPi}}_{\mathrm{Rest}}$ (ie. the SNr/GPi output when the salience vector is null).

The main difference of our architecture with the recent GPR proposal (Gurney et al., 2001a) is the nuclei targeted by the external part of the globus pallidus (GPe) and the nature of these projections. In our model, the GPe projects to the subthalamic nucleus (STN), the internal part of the globus pallidus (GPi) and the substantia nigra pars reticulata (SNr), as well as to the striatum, as documented in Bevan, Booth, Eaton, and Bolam (1998), Kita, Tokuno, and Nambu (1999) and Staines, Atmadja, and Fibiger (1981). 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, Lavallee, Lévesque, & Parent, 2000). This specific organization 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 (Eqs. (9) and (11)), while there is a simple channel-to-channel projection to the striatum Eqs. (6) and (7).

The striatum is one of the two input nuclei of the BG. It is mainly composed of GABAergic (inhibitory) medium spiny neurons (MSN). As in the GPR model, we distinguish among them, those with D1 and D2 dopamine receptors and modulate the input generated in the dendritic tree by the dopamine level γ, which here encompasses salience, frontal cortex feedback and GPe projections.

Using the formulation of Eq. (3), the ith neurons ($i \in [1, N]$, with N the number of channels) of the D1 and D2 subparts of the striatum are defined as follows

$$(\mathbf{Wx} + \mathbf{I}(t))_{\mathrm{D1}_i} = \frac{1}{\tau}\left((1+\gamma)(w^{\mathrm{D1}}_{\mathrm{FC}} x_i^{\mathrm{FC}} - w^{\mathrm{D1}}_{\mathrm{GPe}} x_i^{\mathrm{GPe}} + w^{\mathrm{D1}}_{S} S_i(t)) - w^{\mathrm{D1}}_{\mathrm{FS}} x^{\mathrm{FS}} + I_{\mathrm{D1}}\right) \quad (6)$$

$$(\mathbf{Wx} + \mathbf{I}(t))_{\mathrm{D2}_i} = \frac{1}{\tau}\left((1-\gamma)(w^{\mathrm{D2}}_{\mathrm{FC}} x_i^{\mathrm{FC}} - w^{\mathrm{D2}}_{\mathrm{GPe}} x_i^{\mathrm{GPe}} + w^{\mathrm{D2}}_{S} S_i(t)) - w^{\mathrm{D2}}_{\mathrm{FS}} x^{\mathrm{FS}} + I_{\mathrm{D2}}\right) \quad (7)$$

Table 1
Parameters of the simulations

N	6	τ	40 ms	τ_{STN}	5 ms	τ_{FS}	5 ms	τ_{FC}	80 ms
τ_{TH}	5 ms	τ_{TRN}	5 ms	γ	0.2	w^{D2}_{GPe}	1	w^{GPe}_{D2}	0.4
w^{D1}_{GPe}	1	w^{GPe}_{D1}	0.4	w^{FS}_{GPe}	0.05	w^{D1}_{FS}	0.5	w^{D2}_{FS}	0.5
w^{GPe}_{STN}	0.7	w^{STN}_{GPe}	0.45	w^{GPi}_{GPe}	0.08	w^{GPi}_{STN}	0.7	w^{GPi}_{D1}	0.4
w^{TH}_{TRN}	0.35	w^{TRN}_{TH}	0.35	w^{TH}_{FC}	0.6	w^{FC}_{TH}	0.6	w^{TRN}_{FC}	0.35
w^{TH}_{GPi}	0.18	w^{STN}_{FC}	0.58	w^{D1}_{FC}	0.1	w^{D2}_{FC}	0.1	w^{FS}_{FC}	0.01
I_{D1}	−0.1	I_{D2}	−0.1	I_{STN}	0.5	I_{GPi}	0.1	I_{GPe}	0.1

where $S(t)$ is the salience input vector, and where the negative constant inputs I_{D1} and I_{D2}, which keep the neurons silent when the inputs are not strong enough, model the up-state/down-state property of the MSNs.

The striatum also contains a small proportion of phenotypically diverse interneurons (Tepper & Bolam, 2004). We include here the fast spiking GABAergic interneurons (FS), that we model roughly as a single population exerting feedforward inhibition on the MSN (Tepper, Koós, & Wilson, 2004), and modulated by GPe feedback (Bevan et al., 1998)

$$(\mathbf{W}\mathbf{x}+\mathbf{I}(t))_{FS} = \frac{1}{\tau_{FS}}\sum_{j=1}^{N}\left(w^{FS}_{FC}x^{FC}_j - w^{FS}_{GPe}x^{GPe}_j + w^{FS}_{S}S_j(t)\right). \tag{8}$$

The subthalamic nucleus (STN) is the second input of the basal ganglia and also receives diffuse projections from the GPe, as explained above. Its glutamatergic neurons have an excitatory effect and project to the GPe and GPi. The resulting input of the STN neuron is given by

$$(\mathbf{W}\mathbf{x}+\mathbf{I}(t))_{STN_i} = \frac{1}{\tau_{STN}}\left(w^{STN}_{FC}x^{FC}_i - w^{STN}_{GPe}\sum_{j=1}^{N}x^{GPe}_j + I_{STN}\right) \tag{9}$$

where the constant positive input I_{STN} models the tonic activity of the STN.

The GPe is an inhibitory nucleus, it receives channel-to-channel afferents from the whole striatum (Wu, Richard, & Parent, 2000), and a diffuse excitation from the STN

$$(\mathbf{W}\mathbf{x}+\mathbf{I}(t))_{GPe_i} = \frac{1}{\tau}\left(-w^{GPe}_{D1}x^{D1}_i - w^{GPe}_{D2}x^{D2}_i + w^{GPe}_{STN}\sum_{j=1}^{N}x^{STN}_j + I_{GPe}\right) \tag{10}$$

where the constant positive input I_{GPe} models the tonic activity of the GPe.

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

$$(\mathbf{W}\mathbf{x}+\mathbf{I}(t))_{GPi_i} = \frac{1}{\tau}\left(-w^{GPi}_{D1}x^{D1}_i + w^{GPi}_{STN}\sum_{j=1}^{N}x^{STN}_i - w^{GPi}_{GPe}\sum_{j=1}^{N}x^{GPe}_j + I_{GPi}\right) \tag{11}$$

where the constant positive input I_{GPi} models the tonic activity of the GPi/SNr.

Finally, the thalamus (TH) forms an excitatory loop with the frontal cortex (FC), these two modules representing different thalamus nuclei and cortical areas, depending on the cortico-baso-thalamo-cortical loop considered. The thalamus is moreover under a global regulatory inhibition of the thalamic reticular nucleus (TRN, represented by a single population of neurons) and a channel-specific selective inhibition from the basal ganglia

$$(\mathbf{W}\mathbf{x}+\mathbf{I}(t))_{TH_i} = \frac{1}{\tau_{TH}}\left(w^{TH}_{FC}x^{FC}_i - w^{TH}_{TRN}x^{TRN} - w^{TH}_{GPi}x^{GPi}_i\right) \tag{12}$$

$$(\mathbf{W}\mathbf{x}+\mathbf{I}(t))_{FC_i} = \frac{1}{\tau_{FC}}\left(w^{FC}_{S}S_i + w^{FC}_{TH}x^{TH}_i\right) \tag{13}$$

$$(\mathbf{W}\mathbf{x}+\mathbf{I}(t))_{TRN} = \frac{1}{\tau_{TRN}}\left(\sum_i w^{TRN}_{FC}x^{FC}_i + w^{TRN}_{TH}x^{TH}_i\right). \tag{14}$$

This model keeps the basic off-center on-surround selecting structure, duplicated in the D1-STN-GPi/SNr and D2-STN-GPe subcircuits, of the GPR. However, the channel-specific feedback from the GPe to the Striatum helps in sharpening the selection by favoring 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. The inhibitory projections of the BG onto the thalamo-cortical excitatory loop limits the amplification of the unselected channels and thus favors a selective amplification of the winning channels. In such an architecture, the frontal cortex preserves the information from all channels but amplifies selectively the winning channel, in a sort of attention "spotlight" process, while the subcortical target circuits of the BG are under very selective inhibition, ensuring that motor commands do not interfere.

4. Disembodied model results

We first analyze the contraction of the contracting basal ganglia model (CBG) and its selection properties in simple disembodied tests before evaluating it as an ASM in a simulated robot.

Similarly to the simulations made by Gurney et al. (2001b), we used a 6-channel model. The parameters of the model were hand-tuned in order to obtain a selective system and respecting the local contraction constraints defined below, their values are summarized in Table 1. The simulation was programmed in C++, using the simple Euler approximation for integration, with a time step of 1 ms.

4.1. Contraction analysis of the model

According to the theory developed in Section 2.3, our model is contracting if the non-projected dynamics (which are linear) are contracting in a diagonal metric. To find this metric, we will use the three combinations presented in Section 2.4 that preserve diagonality.

Remark that each separated nucleus is trivially contracting in the identity metric because there is no lateral connection. The contracting rate of each nucleus is $\frac{1}{\tau}$, where τ is the common time constant of the N neurons of the nucleus. Thus, the metric $\mathbf{M}_{BG}$ of the basal ganglia is constituted of the blocks $\kappa_{GPe}\mathbf{I}$, $\kappa_{STN}\mathbf{I}$, $\kappa_{D1}\mathbf{I}$, $\kappa_{D2}\mathbf{I}$, $\kappa_{FS}1$ and $\kappa_{GPi}\mathbf{I}$. Similarly, the thalamic metric $\mathbf{M}_{TH}$ is constituted of

the blocks $\kappa_{\mathrm{FC}}\mathbf{I}$, $\kappa_{\mathrm{TH}}1$ and $\kappa_{TRN}\mathbf{I}$. The resulting metric for the whole system $\mathbf{M}_{\mathrm{CBG}}$ combines $\mathbf{M}_{BG}$ and $\mathbf{M}_{\mathrm{TH}}$ in the following way

$$\mathbf{M}_{\mathrm{CBG}} = \begin{pmatrix} \mathbf{M}_{BG} & 0 \\ 0 & \alpha\mathbf{M}_{\mathrm{TH}} \end{pmatrix}.$$

4.1.1. Analysis of the basal ganglia

- $\kappa_{\mathrm{GPe}} = 1$.
  We can set κ_{GPe} to any value as there is no combination at this stage. The current contracting rate is $\frac{1}{\tau}$.
- $\kappa_{\mathrm{STN}} = w^{\mathrm{GPe}}_{\mathrm{STN}}/w^{\mathrm{STN}}_{\mathrm{GPe}}$.
  We use negative feedback. The contracting rate remains unchanged
- $\begin{cases} \kappa_{\mathrm{D1}} = w^{\mathrm{GPe}}_{\mathrm{D1}}/((1+\gamma)w^{\mathrm{D1}}_{\mathrm{GPe}}) \\ \kappa_{\mathrm{D2}} = w^{\mathrm{GPe}}_{\mathrm{D2}}/((1-\gamma)w^{\mathrm{D2}}_{\mathrm{GPe}}). \end{cases}$
  We use small gains to show that the system constituted by the STN, GPe, striatum D1 and D2 is contracting when

$$\left((1+\gamma)w^{\mathrm{GPe}}_{\mathrm{D1}}w^{\mathrm{D1}}_{\mathrm{GPe}}\right)^2 + \left((1-\gamma)w^{\mathrm{GPe}}_{\mathrm{D2}}w^{\mathrm{D2}}_{\mathrm{GPe}}\right)^2 < 1 \tag{15}$$

  with a contracting rate

$$\frac{1}{\tau}\left(1 - \sqrt{((1+\gamma)w^{\mathrm{GPe}}_{\mathrm{D1}}w^{\mathrm{D1}}_{\mathrm{GPe}})^2 + ((1-\gamma)w^{\mathrm{GPe}}_{\mathrm{D2}}w^{\mathrm{D2}}_{\mathrm{GPe}})^2}\right).$$

- $\kappa_{\mathrm{FS}} = w^{\mathrm{D1}}_{\mathrm{FS}}/w^{\mathrm{FS}}_{\mathrm{GPe}}$.
  Again by use of small gains.
- $\kappa_{\mathrm{GPi}} = 1/(\tau\sigma(\mathbf{G}))^2$

where $\sigma(\mathbf{G})$ is the largest singular value of the matrix of projections on GPi and τ is the slowest time constant of neurons in the basal ganglia. This constant is set by using hierarchical combination.

Thus we can guarantee the contraction of the basal ganglia as soon as condition (15) is satisfied.

4.1.2. Analysis of the thalamus

- $\kappa_{\mathrm{TH}} = 1$.
  We can set κ_{TH} to any value as there is no combination at this stage. The current contracting rate is $\frac{1}{\tau_{\mathrm{TH}}}$.
- $\kappa_{\mathrm{GPe}} = w^{\mathrm{TH}}_{TRN}/w^{TRN}_{\mathrm{TH}}$.
  We use negative feedback. The contracting rate remains unchanged
- $\kappa_{\mathrm{FC}} = \sqrt{{w^{\mathrm{TH}}_{\mathrm{FC}}}^2 + N{w^{TRN}_{\mathrm{FC}}}^2}/w^{\mathrm{FC}}_{\mathrm{TH}}$.
  We use small gains to show that the thalamo-cortical module is contracting when

$$w^{\mathrm{FC}}_{\mathrm{TH}}\left(w^{\mathrm{TH}}_{\mathrm{FC}} + \sqrt{{w^{\mathrm{TH}}_{\mathrm{FC}}}^2 + N{w^{TRN}_{\mathrm{FC}}}^2}\right) < 1. \tag{16}$$

  Remark that this condition depends on N. This would not have been the case if we had modeled the *TRN* by N channels instead of 1.

Thus we can guarantee the contraction of the thalamus as soon as condition (16) is satisfied.

It remains to examine the large loop between the thalamus and the basal ganglia involving projections of the GPi and the FC. Again, we use small gains to set α.

$$\alpha = \sqrt{\frac{\tau_{FCtx}\kappa_{\mathrm{GPi}}\left({w^{\mathrm{STN}}_{\mathrm{FC}}}^2 + {w^{\mathrm{D1}}_{\mathrm{FC}}}^2 + {w^{\mathrm{D2}}_{\mathrm{FC}}}^2 + n{w^{\mathrm{FS}}_{\mathrm{FC}}}^2\right)}{\tau_{\mathrm{TH}}\kappa_{\mathrm{FC}}\,{w^{\mathrm{TH}}_{\mathrm{GPi}}}^2}}.$$

Proposition 2. *Let* $\mathbf{M}_{\mathrm{CBG}} = \mathbf{\Theta}^{\mathrm{T}}_{\mathrm{CBG}}\mathbf{\Theta}_{\mathrm{CBG}}$ *be the diagonal metric defined above. By Theorem 2, if the generalized Jacobian* $\mathbf{\Theta}_{\mathrm{CBG}}\mathbf{W}\mathbf{\Theta}^{-1}_{\mathrm{CBG}}$ *is negative definite, the dynamical system* $\dot{\mathbf{x}} = \mathbf{\Pi}_{\mathbb{H}_n}(\mathbf{x}, \mathbf{W}\mathbf{x} + \mathbf{I}(t))$ *describing the cortico-baso-thalamo-cortical loop model is contracting with a rate* $|\lambda_{\max}|$*, where* $\lambda_{\max}$ *is the largest eigenvalue of* $\mathbf{\Theta}_{\mathrm{CBG}}\mathbf{W}\mathbf{\Theta}^{-1}_{\mathrm{CBG}}$.

Table 2
Value of the constants defining the metric M_{CBG} for the set of parameters of our simulation

κ_{GPe}	κ_{STN}	κ_{D1}	κ_{D2}	κ_{FS}	κ_{GPi}	κ_{TH}	κ_{TRN}	κ_{FCtx}	α
1	0.441	0.577	0.707	1	0.104	1	1	5.282	0.253

At this stage, we have provided an algebraic definition of the metric $\mathbf{M}_{\mathrm{CBG}}$. Unfortunately, the complexity of the induced generalized Jacobian prevents us from giving a global algebraic condition on the projection weights for the generalized Jacobian to be negative definite. This is not of major incidence as we can compute numerically, for any instance of the weights, the eigenvalues of the symmetric part of the generalized Jacobian and check that they are all negative.

Table 2 gives the numerical value of the constants defining the metric M_{CBG} for the set of parameters of our simulation (see Table 1). Using the free software *Octave*, we compute in that case the eigenvalues of the generalized Jacobian and obtain that our model is contracting with contracting rate of 2.20.

Notice that computing the maximum real part of the eigenvalue of the non-projected dynamics (which are linear) gives an upper bound of the contracting rate. For the set of parameters of our simulation, this upper bound is 2.59. It is remarkable that being forced to use diagonal metrics in our proof (which discards a huge set of metrics) has not decreased much the contracting rate.

4.2. Basic selection test

We first reproduced the selection test of Gurney et al. (2001b) with our model and with the GPR model version presented in Prescott et al. (2006). In this test, a specific sequence of five different salience vectors (represented by the dashed lines in Fig. 3) is submitted to a 6-channel version of the BG model, in order to show the basic selection properties of the system. Here, we submitted each vector to the system during 2 s before switching to the next one in the sequence.

During the CBG simulation (Fig. 3, top row), with the first vector of null saliences, the system stabilizes in a state where all channels are equally inhibited ($x^{\mathrm{GPi}}_i = 0.095$). Then, the first channel receives a 0.4 input salience which results in a clear disinhibition of this channel ($x^{\mathrm{GPi}}_1 = 0.014$) and increased inhibition of the others. When the second channel salience is set to 0.6, it becomes perfectly selected ($x^{\mathrm{GPi}}_2 = 0$) 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 0.6, channels 1 and 2 are therefore simultaneously selected ($x^{\mathrm{GPi}}_1 = x^{\mathrm{GPi}}_2 = 0.03$). Finally, during the last step of the test, channel 1 has its salience reduced to 0.4, and it is then rapidly inhibited, while channel 2 returns to perfect selection ($x^{\mathrm{GPi}}_2 = 0$). The CBG thus passes this test in a satisfactory manner: the channels with the highest saliences are always selected while the others are inhibited.

The GPR simulation (Fig. 3, bottom row) is qualitatively quite similar, excepted during the fourth step of the sequence (emphasized with an asterisk): while the salience of channel 1 increases from 0.4 up to 0.6 (the same salience as that of channel 2), channel 2 remains selected and channel 1 is fully inhibited (its level of inhibition is higher than the inhibition at rest). The inputs in channels 1 and 2 being exactly the same, this difference in their selection state is clearly caused by the initial conditions of the system (i.e. the fact that channel 2 was selected before). This example of a dependence on the initial conditions clearly shows that the GPR model is not contracting.

Indeed, as we have seen in Section 2.3, a rate-coding neural network is contracting only if its non-projected dynamics are contracting in a diagonal metric. But a linear system is stable if and

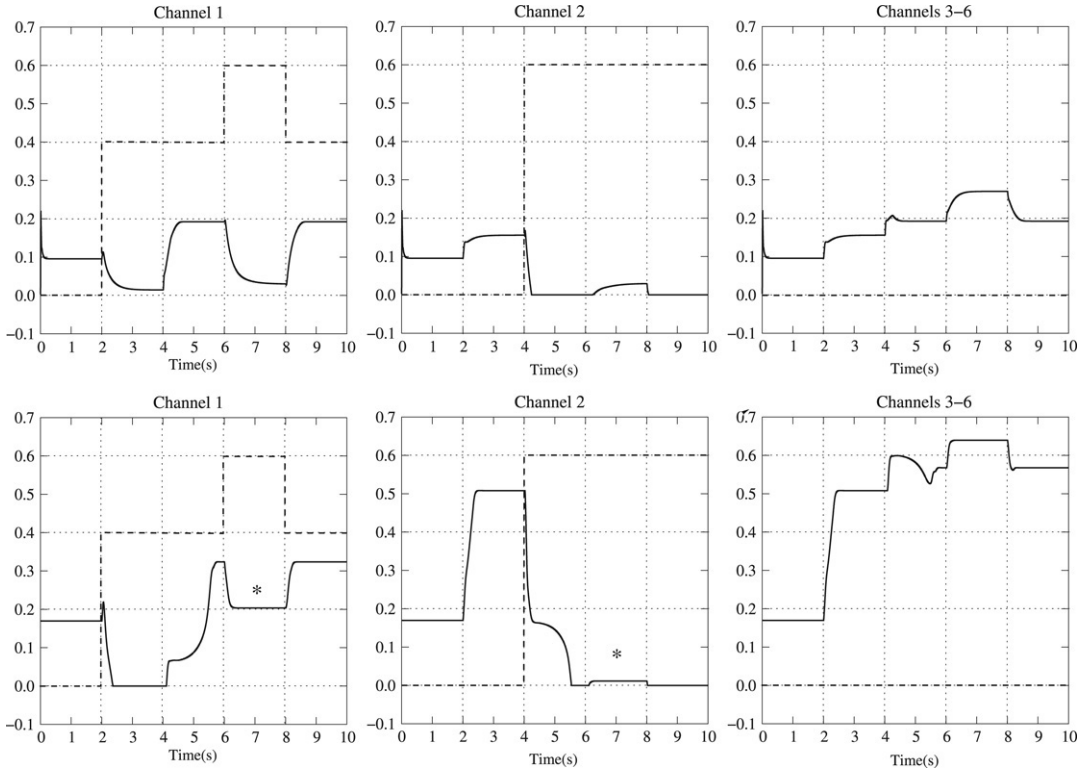

Fig. 3. Variation of the GPi/SNr inhibitory output during the Gurney et al. (2001b) test applied to (top) the CBG and (bottom) the GPR. Dashed lines represent the input salience of the channel and solid lines represent the output of the channel. Note that during the fourth step (6 s $< t <$ 8 s), channels 1 and 2 are selected by the CBG, while the GPR selects channel 2 only (asterisk).

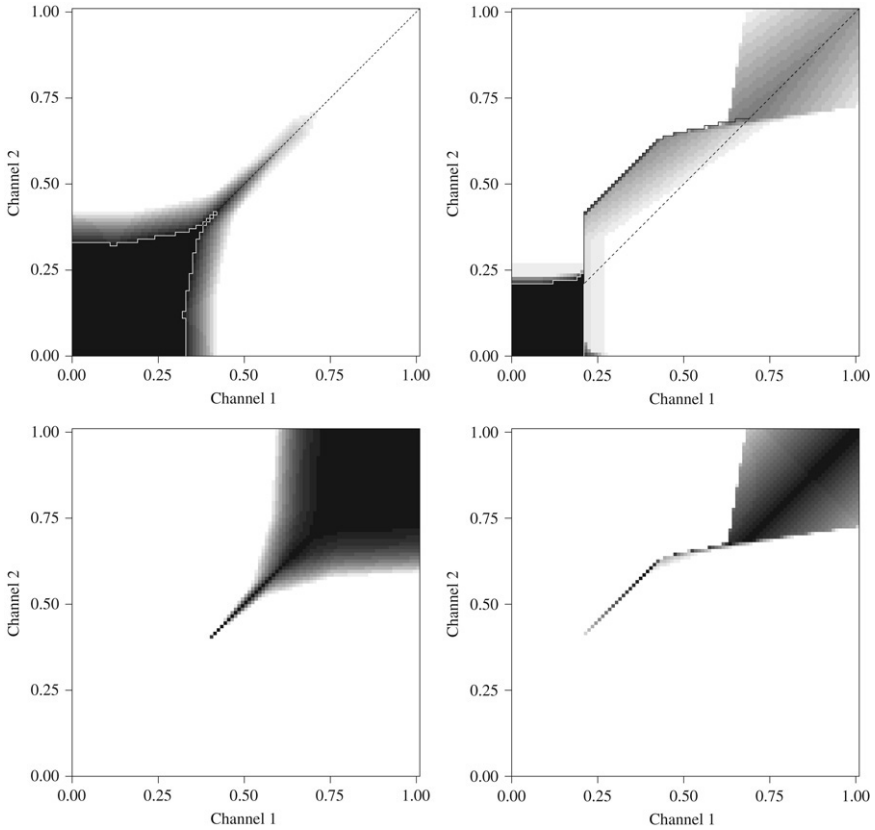

Fig. 4. Efficiency (top) and distortion (bottom) in the winning channel for a systematic salience-space search for the CBG (left) and the GPR (right). Top: black to white gradient represents increasing efficiency (from 0 to 1); bottom: black to white gradient represents decreasing distortion (from 1 to 0), maximal distortion corresponding to simultaneous selection of both channels is thus in black. White line: limit beyond which no selection occurs; dashed black line: diagonal representing equal saliences. For the GPR efficiency (top right), note the hysteresis area between the dashed and the full black lines. See the text for further explanations.

only if all its eigenvalues have a negative real part. Computing the eigenvalues of the linear part of the GPR reveals that $N-1$ of them have a positive real part (namely 10.387). We can thus conclude that the GPR is not contracting.

4.3. Systematic salience search test

This first result is however not surprising, as revealed by the systematic salience search experiment performed in Prescott et al. (2006), and that we also reproduced with both the GPR and the CBG. In this experiment, the first two channels of the ASM are put in competition in the following manner: the first channel salience is increased from 0 to 1 in steps of 0.01, and for each of these steps, the salience of the second channel is also gradually increased from 0 to 1 in steps of 0.01. The system is run to convergence between all step increases. The internal state of the model is not reset between each channel 2 salience increase, but only for channel 1 steps. This means that the test evaluates the selection response of the system with one channel salience fixed while the other one gradually increases.

In order to evaluate the response of the ASM to this experiment, four numerical values are computed. First, the efficiencies of the selection of channels 1 and 2, equivalent to the percentage of disinhibition, are computed as follows:

$$e_i = [1 - y_i^{\mathrm{GPi}}/y_{\mathrm{Rest}}^{\mathrm{GPi}}]_+ \tag{17}$$

with i the index of the channel, y_i^{GPi} the output of the ith GPi neuron and $y_{\mathrm{Rest}}^{\mathrm{GPi}}$ the output inhibition of all channels when all saliences are null. The absolute efficiency of the selection is defined as the efficiency of the winning channel:

$$e_w = \max_i e_i. \tag{18}$$

Finally, the distortion of the selection, which is null when only the winning channel is disinhibited and increasing with the disinhibition of its competitors, is defined by:

$$d_w = 2\frac{\sum_i e_i - e_w}{\sum_i e_i}. \tag{19}$$

The results of the experiment are summarized by the e_w and d_w graphs (Fig. 4), where the value of each of these variables is represented with regard to the corresponding channel 1 (abscissa) and channel 2 (ordinate) saliences. First observe that the GPR results we obtain with 6 channels are very similar to those presented in Prescott et al. (2006) for a 5-channel GPR. Concerning e_w (top row), whereas, for the CBG, the selection switches from channel 1 to channel 2 as soon as the salience of channel 2 is larger than the salience of channel 1 (when it crosses the diagonal in dashed black), for the GPR, this switch is delayed until much higher values are reached (when it crosses the black line). As previously noted, this hysteresis effect is a direct consequence of the non-contraction of the GPR.

Note that when high saliences are in competition, the GPR tends to partially select both channels ($e_w < 1$ and $d_w > 0$), while the CBG fully disinhibits both channels ($e_w = 1$ and d_w close to 1). Which behavior is preferable for an ASM is not decided.

Is the GPR's strong dependence on initial conditions a good feature for an ASM? Prescott et al. (2006) argue that it allows behavioral persistence, and that in their experiment, the robot takes advantage of it to avoid dithering between actions. We do not claim that there is a definitive answer to the question. Nevertheless, in the next section, we describe the evaluation of the CBG in a minimal survival task in which the robot also avoids dithering, despite its contracting ASM. This shows that this dependence on initial conditions is not necessary from the point of view of dithering avoidance.

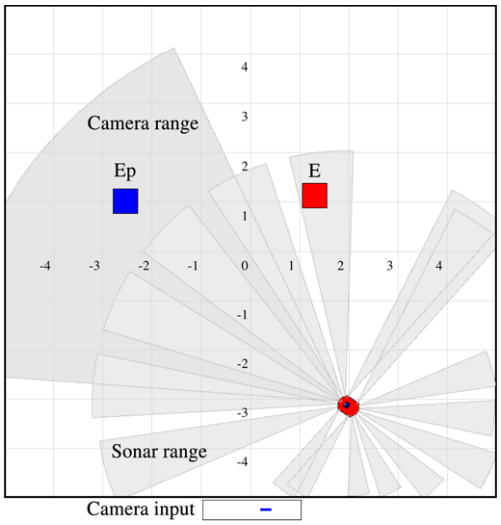

Fig. 5. Experimental set-up. Blue square: Potential Energy resource; red square: Energy resource. The light gray surfaces represent the field of view of the sonars, and the darker one the field of view of the camera. The corresponding camera image is represented at the bottom.

5. Minimal survival task

5.1. Materials and methods

The suitability of the model for action selection in an autonomous robot has been tested in simulation with the same minimal survival task previously used to evaluate the GPR model (Girard et al., 2003). In order to emphasize its properties, and in particular those resulting from the selective feedback loop, its performance was compared to a simple if-then-else decision rule (ITE, fully described in Appendix A).

In such a task, the robot has to go back and forth between locations containing two different kind of resources, in order to keep its energy level above 0. The robot has two internal variables, namely *Energy* and *Potential Energy*, taking values between 0 and 1, and an artificial metabolism, which couples them as follows:

- The Energy (E) is continuously decreasing, with a constant consumption rate (0.01 Energy unit per second). When it reaches 0, the robot has run out of energy and the ongoing trial is interrupted. To prevent this, the robot has to regularly acquire Energy by activating the *ReloadOnE* action on an Energy resource. Note that ReloadOnE only *transforms* Potential Energy into Energy (0.2 units of E_p are transformed into 0.2 units of E each second), thus Potential Energy has to be also reloaded.
- The Potential Energy (E_p) is a sort of Energy storage, it can be acquired by activating the *ReloadOnEp* action on a Potential Energy resource, and is consumed in the transformation process only.

In this version of the task, the experiments are run in simulation using the Player/Stage robot interface and robot simulator (Gerkey, Vaughan, & Howard, 2003). The simulated robot is a 40 × 50 cm wheeled robot with differential steering, similar to the Activ-Media Pioneer 2DX (Fig. 5), equipped with a ring of 16 sonars and a camera. The sonar sensors have a maximum range of 5 m and a view angle of 15°, the camera has a resolution of 200 × 40 pixels and a view angle of 60° and uses a color-blob-finding vision device to track the position of red and blue objects. The experiment takes place in a 10 × 10 m arena, containing one Energy and one Potential Energy resource (Fig. 5). These resources are represented by colored 50 × 50 cm objects (respectively red and blue), and do not constitute obstacles (as if they were suspended above the

arena). They are randomly positioned in the arena for each trial, with the constraint that their center is at least 1 m away from the walls.

The robot has to select from among seven possible actions:

- ReloadOnE (ROE) and ReloadOnEp (ROE_p) affect the robot's survival as previously described. These actions are effective if the robot is facing the corresponding resource and is close enough (45° of the camera field of view is occupied by the resource).
- Wander (W) activates random accelerations, decelerations and turning movements.
- Rest (R) stops the robot, which is a disadvantage as the robot has to continuously explore the arena to find resources, but Rest also halves the rate of Energy consumption (0.005 unit per s), which promotes long survival. Consequently, it should be activated when there is no risk (i.e. when both internal variables reach high levels) in order to minimize the Potential Energy extracted from the environment to survive.
- AvoidObstacle (AO) uses data from the 6 front sonars and the 2 central rear sonars in order to avoid collisions with walls.
- ApproachE (AE) and ApproachEp (AE_p) use the color-blob-finder in order to orient and displace the robot towards the corresponding resource if it is visible.

The action selection mechanisms base their decisions on the following variables:

- E, E_p,$(1-E)$ and $(1-E_p)$, which provide the amount (or lack of) Energy and Potential Energy,
- *seeEBlob* and *seeEpBlob*, which are set to 1 if a red (resp. blue) object is in the camera input, and to 0 otherwise,
- *onEBlob* and *onEpBlob*, which are set to 1 if a red (resp. blue) object is larger than 150 pixels (i.e. close enough to allow the use of the corresponding resource), and to 0 otherwise,
- *SFR* and *SFL* are the values of the front-right and front-left sonar sensors, measured in meters, taking values between 0 and 5.

For the CBG, the detailed salience computation using these variables is given in Appendix B.

The action selection mechanisms receive new sensory data every 100 ms, and must then provide an action selection for the next 100 ms. Concerning the ITE, it is simply done by executing the decision rule once with the latest data. Concerning the CBG, the selection is made using the output inhibition resulting from the computation of 100 simulation steps of 1ms, using the latest sensory data. A given action is then considered selected if the inhibition of the corresponding channel is below the inhibition at rest y^{GPi}_{Rest} (as defined previously). In the case of multiple channel disinhibition, the following action combination rules have been defined:

- Rest is effective if and only if it is the only disinhibited action,
- ReloadOnE and ReloadOnEp are effective if and only if the robot does not move,
- The other movement-generating actions can be co-activated. In that case, the efficiency of selection (as defined by Eq. (17)) is used to weight the contributions of each action to the final motor command.

The comparison between the CBG and the ITE is made according to the following protocol: 20 random resource positions are drawn and, for each model, 20 trials are run using the same set of positions. The robot begins the experiment with a full battery ($E = 1$) and no Potential Energy storage ($E_p = 0$), this allows a maximal survival duration of 1 min 40 s if no reloading action occurs. Unless the robot runs out of energy ($E = 0$), the trial is stopped after 15 min.

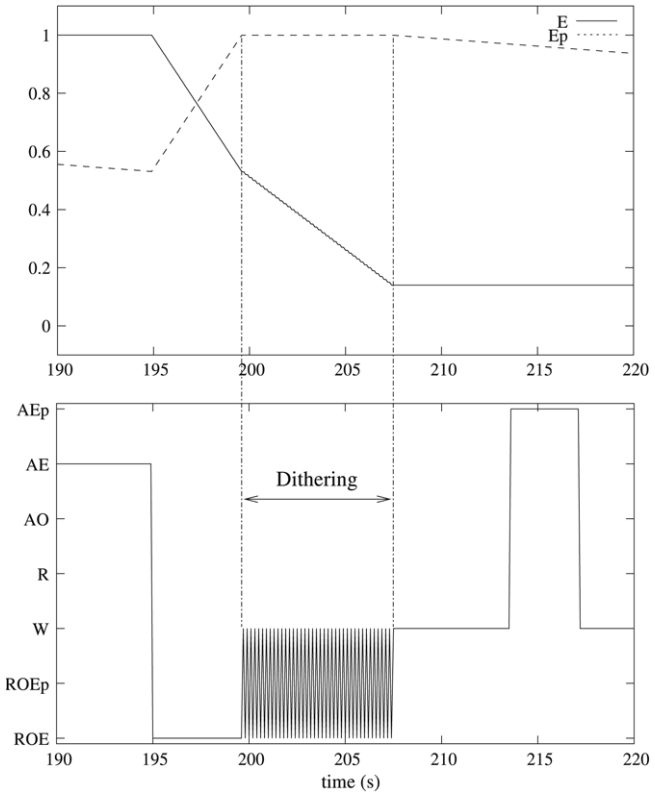

Fig. 6. Typical dithering of the ITE between the ReloadOnEnergy and Wander actions. Top: levels of Energy (dashed line) and Potential Energy (full line); bottom: selected action. Note how during the dithering period, more than 0.3 units of E_p are wasted in about 7 s, while they should have allowed 30 s of survival.

5.2. Results

The first result is that the CBG and the ITE algorithm have similar survival performance. They are both able to survive the trial in a majority of cases, but can be subject to premature Energy shortage. This is expected, because their ability to find resources is limited by the camera range and field of view, as well as by the random exploration action. The average survival duration is 687 s ($\sigma = 244$) for the CBG and 737 s ($\sigma = 218$) for the ITE, and the two-tailed Kolmogorov–Smirnov test confirms that the two sets of survival durations are not drawn from significantly different distributions ($D_{KS} = 0.2, p = 0.771$). From an action selection point of view, the comparison of the two mechanisms is thus fair: despite the fact that they were tuned independently, they both achieve similar survival performance.

Nevertheless, *a clear behavioral difference between the two mechanisms was observed*, which has significant repercussions on their ability to store Potential Energy and on the Potential Energy extracted from the environment. Indeed, while the CBG may use its feedback loops in order to persist in action execution, the ITE was deliberately deprived of any memory. This was done in order to investigate the effects of this persistence property. The ITE exhibits behavioral dithering in a critical and frequent situation: when the robot fully reloads its Energy, it activates the Wander action, but after 100 ms of Wander execution, some Energy has been consumed and the robot has not moved much. In most cases, it is still on the Energy resource, and if it still has spare E_p, ReloadOnE is activated again. This repeats until there is no E_p left or until, in a sequence of small movements, the robot has left the resource (see Fig. 6). This dithering generates a strong energy dissipation: 100 ms of Wander consumes 0.001 units of Energy, and during the following 100 ms, ReloadOnEnergy consumes 0.02 units of E_p while E, being bounded by 1, increases by 0.001 only.

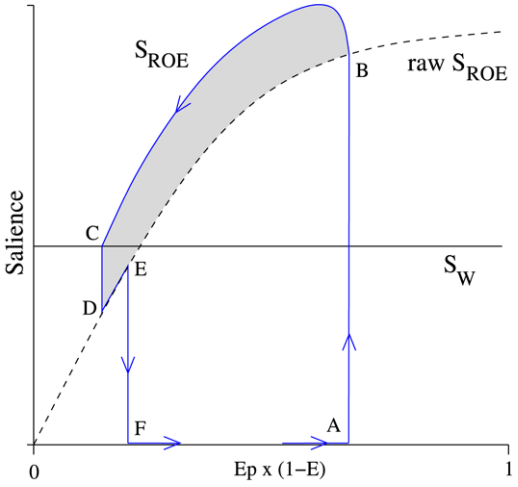

Fig. 7. Hysteresis in the variation of the salience of ReloadOnEnergy for the CBG. Black dashed line: variation of S_{ROE} with regard to $(E_p \times (1 - E))$, with $onEBlob = 1$ and without the persistence term (raw S_{ROE}); blue line: variation of S_{ROE}; shaded area: S_{ROE} increase resulting from the frontal cortex feedback; black line: salience of Wander (S_W). Explanations are given in the text.

On the contrary, in the same situation, the CBG takes advantage of a hysteresis effect caused by the positive feedback from the frontal cortex to the basal ganglia to avoid dithering.

Indeed, the salience of ROE is defined by: $S_{\mathrm{ROE}} = 950 \times f(4 \times onEBlob \times E_p \times (1 - E)) + 0.6 \times x^{\mathrm{FC}}_{\mathrm{ROE}}$ (where f is a sigmoid transfer function, see Appendix B). Consequently, when the robot has a lack of Energy and reaches an Energy resource, *onEBlob* jumps from 0 to 1 and S_{ROE} also jumps from 0 (Fig. 7, point A) to a level depending on the current E and E_p internal states (Fig. 7, point B) situated on the raw S_{ROE} curve (Fig. 7, dashed line). In the case depicted in Fig. 7, S_{ROE} is then much higher than S_W, and ROE is thus selected. As a consequence, the corresponding thalamo-cortical channel is disinhibited, leading to an amplification of the salience, fed back to the basal ganglia thanks to the cortical output $x^{\mathrm{FC}}_{\mathrm{ROE}}$ (this bonus is represented by the shaded area over the raw S_{ROE} curve on Fig. 7).

While the robot reloads, S_{ROE} decreases with $(E_p \times (1 - E))$, but because of the $x^{\mathrm{FC}}_{\mathrm{ROE}}$ salience bonus, it follows the blue trajectory down to point C, where Wander is selected again. The deselection of ROE shuts off the $x^{\mathrm{FC}}_{\mathrm{ROE}}$ signal, causing an immediate decrease to point D. As soon as the robot activates Wander, Energy is consumed and S_{ROE} increases again, along the raw S_{ROE} curve. However, at point D, $S_{\mathrm{ROE}} < S_W$, and as long as the robot manages to leave the resource before S_{ROE} exceeds S_W (points E and F, when the *OnEBlob* variable jumps from 1 to 0), no dithering occurs.

This observation is not trivial, as it has a direct consequence on the global E_p storage of the ITE: both CBG and ITE keep high levels of E_p (between 0.9 and 1) more than 50% of the time (Fig. 8, right), but for the rest of the time, the ITE level is very low (0–0.1) much more often (almost 20% of the time) than the CBG. Moreover, the CBG activates the Rest action often enough to extract, on average, less Potential Energy from the environment ($0.93 \times 10^{-2}\mathrm{Ep\,s^{-1}}, \sigma = 0.30 \times 10^{-3}$) than the basic rate ($1 \times 10^{-2}\mathrm{Ep\,s^{-1}}$). On the contrary, the dissipation of energy caused by the dithering of the ITE generates a much higher Potential Energy extraction rate ($1.17 \times 10^{-2}\mathrm{Ep\,s^{-1}}, \sigma = 1.17 \times 10^{-3}$). The two-tailed Kolmogorov–Smirnov test reveals that the E_p consumption rates measured for the CBG and the ITE (Fig. 9) are drawn from different distributions ($D_{KS} = 0.95, p < 0.001$). The ITE dithering thus generates so much dissipation that it has to extract extra Potential Energy from the environment, despite its use of the Sleep action to lower its consumption, while the CBG exploits as much as possible this possibility to limit Potential Energy extraction.

6. Discussion

We proposed a new action selection mechanism for an autonomous robot, using a multi-disciplinary approach combining computational neuroscience and dynamic system theory. This study proved fruitful in the three considered domains:

- We proposed an extension of the contraction theory to locally projected dynamical systems, which was necessary to study the stability of rate-coding neural networks.

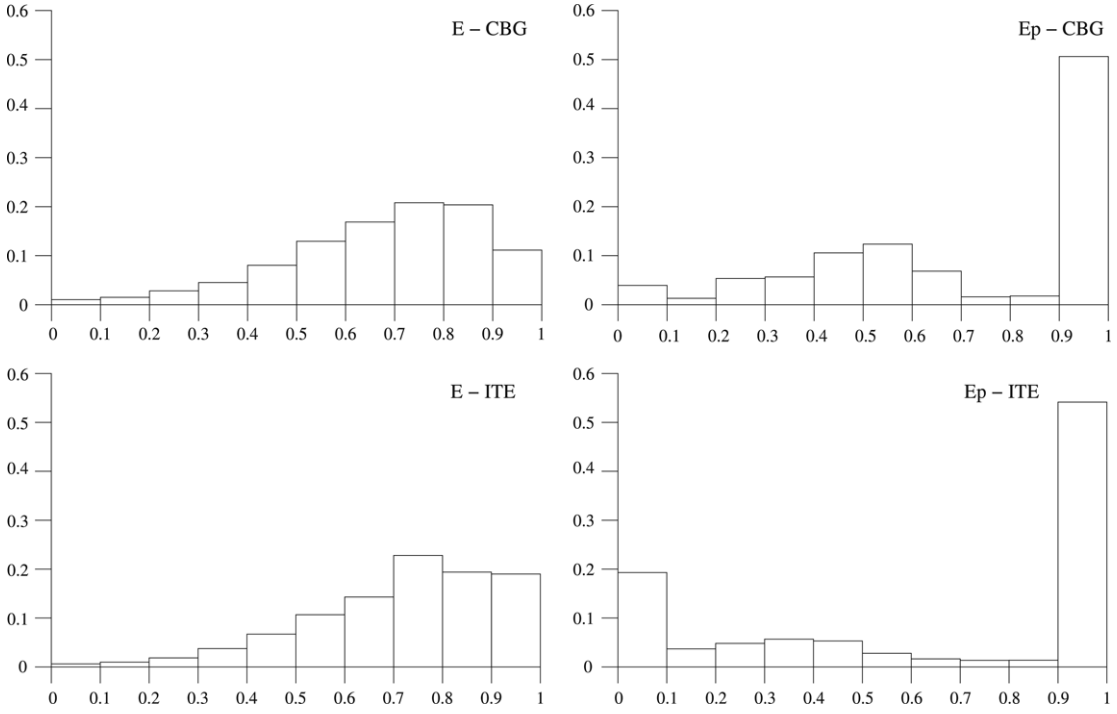

Fig. 8. Histograms of Energy (left) and Potential Energy (right) for the CBG (top) and the ITE (bottom), cumulated over all trials.

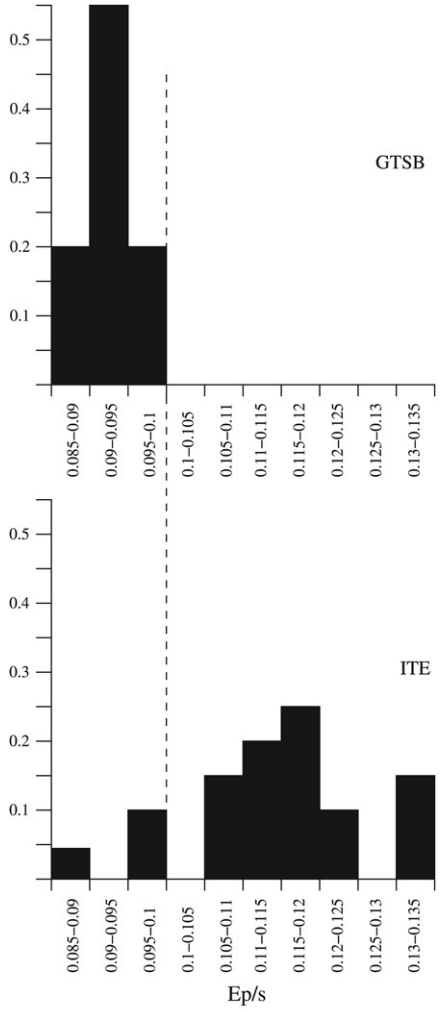

Fig. 9. Potential Energy consumption rate. These histograms represent the average E_p consumption rate computed for each trial. Top: BG model; bottom: ITE; the dashed line shows the Energy consumption rate of all actions except Rest (0.001 E/s).

- As a consequence, we proposed a modified rate-coding artificial neuron model.
- Using these results, we designed a stable model of the cortico-baso-thalamo-cortical loops (CBG) using previously neglected anatomical data.
- After having tested this model offline, we integrated it in a simulated robot confronted to a standard survival task to assess its efficiency as an action selection mechanism.

6.1. Dynamic systems

In this paper, we have investigated the stability properties of locally projected dynamical systems (lPDS) using nonlinear contraction theory. In particular, we have given a sufficient condition for a general non-autonomous (i.e. with time-varying inputs) lPDS to be *globally exponentially stable*. By contrast, Zhang and Nagurney (1995) only studied the stability of a fixed equilibrium point in autonomous lPDS. Thus, the novelty of our theoretical result should be noticed.

Locally projected dynamical systems have attracted great interest since they were introduced in 1993 by Dupuis and Nagurney. Indeed, this theory is central to the study of oligopolistic markets, traffic networks, commodity production, etc (Dupuis & Nagurney, 1993). As we demonstrated in this article, this theory has also proved to be a valuable tool for establishing rigorous stability properties of neural networks. In this respect, further development of the theory as well as its application to numerous problems in theoretical neuroscience may represent exciting subjects of research.

6.2. Neuroscience

The CBG shares a number of similarities with the previously proposed GPR model (Gurney et al., 2001b), as its selection ability relies on two off-center on-surround subcircuits. However, it includes neglected connections from the GPe to the Striatum, which provide additional selectivity. It also considers the possible role of global projections of the GPe to the STN, GPi and SNr as a regulation of the activity in the whole basal ganglia.

We omitted two types of documented connections in the current CBG model. First, the STN projects not only 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 STN neurons project to at least two of the other target nuclei (GPe, GPi or SNr). 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. To our knowledge, no modeling study has yet proposed a functional interpretation of this connection, a question that should be explored in future works. The other missing connections concern the fact that lateral inhibitions exist in GPe and SNr (Deniau, Kitai, Donoghue, & Grofova, 1982; Juraska, Wilson, & Groves, 1977; Park, Falls, & Kitai, 1982). These additional projections were added to a version of the GPR (Gurney, Humphries, Wood, Prescott, & Redgrave, 2004) and seemed to enhance its selectivity. We might add these connections and proceed to a similar test with the CBG.

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 (DeLong et al., 1984), a property supposed to be induced by the specific up/down state behavior 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 salience 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.

The proposed model includes the modulatory role of the dopamine (DA) in the BG selection process only, which corresponds to the tonic level of dopaminergic input from the ventral tegmental area and the substantia nigra pars compacta (VTA and SNc). The effects of the variation of this tonic DA level on the selection abilities of the BG has been examined in detail for the GPR (Gurney et al., 2001b), and compared with the symptoms of Parkinson's disease.

The role of the phasic dopamine activity in reinforcement learning, through the adaptation of the cortico-striatal synapses, is beyond the scope of our study. Nevertheless, such an extension of the CBG could allow the online adaptation of the saliences, which are here hand-tuned. The existing models of reinforcement learning in the BG are based on the temporal difference (TD) learning algorithm (Houk, Adams, & Barto, 1995; Joel, Niv, & Ruppin, 2002). These TD models are composed of two cooperating circuits: a *Critic* dedicated to learning to predict future reward given the current state, and an *Actor*, using the Critic's predictions to choose the most appropriate action. Our model can then be considered as an Actor circuit, more anatomically detailed than those usually used (simple winner-takes-all, without persistence properties). The first attempts at using detailed Actor models

in TD architectures for tasks requiring a single motivation have been conducted (Frank, Santamaria, O'Reilly, & Willcutt, 2007; Khamassi, Girard, Berthoz, & Guillot, 2004; Khamassi, Lachèze, Girard, Berthoz, & Guillot, 2005). Note however that the use of the current TD-learning models would not necessarily be straightforward in our case: we had to use relatively complex salience computations (see Appendix B), in order to solve our relatively simple task. This is caused by its multi-motivational nature, quite common in action selection problems, but which has been given only little attention in RL-related works (Dayan, 2001; Konidaris & Barto, 2006).

6.3. Autonomous robotics

While early action selection mechanisms were based on a purely engineering approach (Pirjanian, 1999), progress in the understanding of the physiology of the brain regions involved in action selection now allows the investigation of biomimetic action selection mechanisms. Indeed, basal ganglia models – variations of the GPR – and reticular formation models have already been used as action selection mechanisms for autonomous robots (Girard et al., 2003, 2005; Humphries, Gurney, & Prescott, 2005; Montes-Gonzalez et al., 2000; Prescott et al., 2006).

We showed here that the CBG may exploit its cortical feedback to exhibit behavioral persistence and thus dithering avoidance, one of the fundamental properties of efficient ASMs (Tyrrell, 1993). In our experiment, this promotes energy storage and reduces energy consumption. These properties, which clearly provide a survival advantage, were also highlighted for the GPR when tested in a similar experiment (Girard et al., 2003). Thus, comparing the GPR and the CBG in exactly the same task could reveal some subtle differences which were not identified yet. Moreover, in the current version of the CBG, these cortico-striatal feedback connections are strictly channel to channel, the possible sequence generation effects that could result from cross channel connections probably deserves additional attention.

The contraction property of the CBG also provides a fundamental advantage for an autonomous robot. It provides a theoretical certainty regarding its stability of operation, whatever the sequences of input might be. For an autonomous agent confronted with a uncontrolled environment, where all possible sequences of inputs may happen, it seems to be essential. Of course, contraction analysis does not say anything about the pertinence of the resulting stable behavior, hence leading the necessity of verifying the CBG selection properties. However, the fact that stability issues have already been evoked for previous GPR versions (Girard et al., 2005; Prescott et al., 2006) confirms that such a rigorous proof is useful.

Acknowledgment

B.G. and N.T. acknowledge the partial support of the European Community Neurobotics project, grant FP6-IST-001917.

Appendix A. If-Then-Else decision rule

The If-Then-Else decision tree is the following:

if $E_p < 1$ and $onEpBlob = true$ **then**
  ReloadOnEp
else if $E < 1$ and $E_p > 0$ and $onEBlob = true$ **then**
  ReloadOnE
else if $E < 0.8$ and $E_p > 0$ and $seeEBlob = true$ **then**
  ApproachE
else if $E_p < 0.8$ and $seeEpBlob = true$ **then**
  ApproachEp
else if $E > 0.7$ and $E_p > 0.7$ **then**
  Rest
else if $SFL < 1$ or $SFR < 1$ or ($SFL < 1.5$ and $SFR < 1.5$) **then**
  AvoidObstacle
else
  Wander
end if

Appendix B. Robot CBG saliences

Using the sigmoid transfer function

$$f(x) = \frac{2}{1 + e^{-4x}} - 1$$

the saliences of each action (including the frontal cortex feedback) are:

$$S_{ROE} = 950 \times f(4 \times onEBlob \times E_p \times (1 - E)) + 0.6 \times x^{FC}_{ROE}$$
$$S_{ROE_p} = 750 \times f(4 \times \text{onEpBlob} \times (1 - E_p)) + 0.2 \times x^{FC}_{ROE_p}$$
$$S_W = 380$$
$$S_{SI} = 550 \times f(2 \times \max(Ep \times E - 0.5, 0))$$
$$S_{AO} = 950 \times f(2 \times (\max(1.5 - SFL, 0) + \max(1.5 - \text{SFR}, 0))) + 0.2 \times x^{FC}_{AO}$$
$$S_{AE} = 750 \times f(\text{seeEBlob} \times E_p \times (1 - E) \times (1 - onEBlob)) + 0.2 \times x^{FC}_{AE}$$
$$S_{AE_p} = 750 \times f(seeEpBlob \times (1 - E_p) \times (1 - \text{onEpBlob})) + 0.2 \times x^{FC}_{AE_p}.$$

References

Alexander, G. E., Crutcher, M. D., & 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., & 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., & 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., & Deniau, M. (1990). Disinhibition as a basic process of striatal functions. *Trends in Neurosciences*, *13*, 277–280.

Dayan, P. (2001). Motivated reinforcement learning. In T. Leen, T. Dietterich, & V. Tresp (Eds.), *Neural information processing systems*: *Vol. 13*. Cambridge, MA: The MIT Press.

Dayan, P., & Abbott, L. (2001). *Theoretical neuroscience: Computational and mathematical modeling of neural systems*. MIT Press.

DeLong, M., Georgopoulos, A., Crutcher, M., Mitchell, S., Richardson, R., & Alexander, G. (1984). Functional organization of the basal ganglia: Contributions of single-cell recording studies. *Ciba Foundation Symposium*, *107*, 64–82.

Deniau, J.-M., Kitai, S., Donoghue, J., & Grofova, I. (1982). Neuronal interactions in the substancia nigra pars reticulata through axon collateral of the projection neurons. *Experimental Brain Research*, *47*, 105–113.

Dupuis, P., & Nagurney, A. (1993). Dynamical systems and variational inequalities. *Annals of Operations Research*, *44*(1), 7–42.

Filippov, A. F. (1988). *Differential equations with discontinuous righthand sides*. Kluwer Academic Pub.

Frank, M., Santamaria, A., O'Reilly, R., & Willcutt, E. (2007). Testing computational models of dopamine and noradrenaline dysfunction in attention deficit/hyperactivity disorder. *Neuropsychopharmacology*, *32*, 1583–1599.

Gerkey, B., Vaughan, R., & Howard, A. (2003). The player/stage project: Tools for multi-robot and distributed sensor systems. In *11th International conference on advanced robotics* (pp. 317–323).

Gillies, A., & Arbruthnott, G. (2000). Computational models of the basal ganglia. *Movement Disorders*, *15*(5), 762–770.

Girard, B., Cuzin, V., Guillot, A., Gurney, K. N., & 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., & Guillot, A. (2005). Integration of navigation and action selection in a computational model of cortico-basal ganglia-thalamo-cortical loops. *Adaptive Behavior*, *13*(2), 115–130.

Girard, B., Tabareau, N., Berthoz, A., & Slotine, J.-J. (2006). Selective amplification using a contracting model of the basal ganglia. In F. Alexandre, Y. Boniface, L. Bougrain, B. Girau, & N. Rougier (Eds.), *NeuroComp* (pp. 30–33).

Girard, B., Tabareau, N., Slotine, J.-J., & Berthoz, A. (2005). Contracting model of the basal ganglia. In J. Bryson, T. Prescott, & A. Seth (Eds.), *Modelling natural action selection: Proceedings of an international workshop* (pp. 69–76). Brighton, UK: AISB Press.

Gurney, K., Humphries, M., Wood, R., Prescott, T., & 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., Wickens, J., & Redgrave, P. (2004). Computational models of the basal ganglia: From membranes to robots. *Trends in Neurosciences*, *27*, 453–459.

Gurney, K., Prescott, T. J., & 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., & Redgrave, P. (2001b). A computational model of action selection in the basal ganglia. II. Analysis and simulation of behaviour. *Biological Cybernetics*, *84*, 411–423.

Horn, R., & Johnson, C. (1985). *Matrix analysis*. Cambridge University Press.

Houk, J. C., Adams, J. L., & Barto, A. G. (1995). A model of how the basal ganglia generate and use neural signals that predict reinforcement. In J. C. Houk, J. L. Davis, & D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (pp. 249–271). Cambridge, MA: The MIT Press.

Humphries, M., Gurney, K., & Prescott, T. (2005). Is there an integrative center in the vertebrate brain-stem? A robotic evaluation of a model of the reticular formation viewed as an action selection device. *Adaptive Behavior*, *13*(2), 97–113.

Ioannou, P., & Sun, J. (1996). *Robust adaptive control*. Upper Saddle River, NJ, USA: Prentice Hall, Inc.

Joel, D., Niv, Y., & Ruppin, E. (2002). Actor-critic models of the basal ganglia: New anatomical and computational perspectives. *Neural Networks*, *15*(4–6).

Juraska, J., Wilson, C., & Groves, P. (1977). The substancia nigra of the rat: A golgi study. *Journal of Comparative Neurology*, *172*, 585–600.

Khamassi, M., Girard, B., Berthoz, A., & Guillot, A. (2004). Comparing three critic models of reinforcement learning in the basal ganglia connected to a detailed actor part in a s-r task. In F. Groen, N. Amato, A. Bonarini, E. Yoshida, & B. Krse (Eds.), *Proceedings of the eighth international conference on intelligent autonomous systems*, IAS8. (pp. 430–437). Amsterdam, The Netherlands: IOS Press.

Khamassi, M., Lachèze, L., Girard, B., Berthoz, A., & Guillot, A. (2005). Actor-critic models of reinforcement learning in the basal ganglia: From natural to artificial rats. *Adaptive Behavior*, *13*(2), 131–148.

Kimura, A., & Graybiel, A. (Eds.) (1995). *Functions of the cortico-basal ganglia loop*. Tokyo, New York: Springer.

Kita, H., Tokuno, H., & Nambu, A. (1999). Monkey globus pallidus external segment neurons projecting to the neostriatum. *Neuroreport*, *10*(7), 1476–1472.

Konidaris, G., & Barto, A. (2006). An adaptive robot motivational system. In S. Nolfi, G. Baldassarre, R. Calabretta, J. Hallam, D. Marocco, J.-A. Meyer, O. Miglino, & D. Parisi (Eds.), *LNAI*: *Vol. 4095*. *From animals to animats 9: Proceedings of the 9th international conference on the simulation of adaptive behavior* (pp. 346–356). Berlin, Germany: Springer.

Krotopov, J., & Etlinger, S. (1999). Selection of actions in the basal ganglia thalamocortical circuits: Review and model. *International Journal of Psychophysiology*, *31*, 197–217.

Lohmiller, W., & Slotine, J. (1998). Contraction analysis for nonlinear systems. *Automatica*, *34*(6), 683–696.

Lohmiller, W., & Slotine, J. (2000). Nonlinear process control using contraction analysis. *American Institute of Chemical Engineers Journal*, *46*(3), 588–596.

Middleton, F. A., & 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., & Redgrave, P. (2000). An embodied model of action selection mechanisms in the vertebrate brain. In J.-A. Meyer, A. Berthoz, D. Floreano, H. Roitblat, & S. W. Wilson (Eds.), *From animals to animats 6*: *Vol. 1* (pp. 157–166). Cambridge, MA: The MIT Press.

Parent, A., Sato, F., Wu, Y., Gauthier, J., Lévesque, M., & 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., & Kitai, S. (1982). An intracellular HRP study of rat globus pallidus. I. responses and light microscopic analysis. *Journal of Comparative Neurology*, *211*, 284–294.

Pirjanian, P. (1999). *Behavior coordination mechanisms — state-of-the-art. Technical report IRIS-99-375*. Institute of Robotics and Intelligent Systems, School of Engineering, University of Southern California.

Prescott, T. J., Montes-Gonzalez, F., Gurney, K., Humphries, M. D., & Redgrave, P. (2006). A robot model of the basal ganglia: Behavior and intrinsic processing. *Neural Networks*, *19*, 31–61.

Redgrave, P., Prescott, T. J., & Gurney, K. (1999). The basal ganglia: A vertebrate solution to the selection problem? *Neuroscience*, *89*(4), 1009–1023.

Sato, F., Lavallee, P., Lévesque, M., & 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., & Coetsee, J. (1986). Adaptive sliding controller synthesis for nonlinear systems. *International Journal of Control*, *43*(4), 1631–1651.

Slotine, J. J. E., & Lohmiller, W. (2001). Modularity, evolution, and the binding problem: A view from stability theory. *Neural networks*, *14*(2), 137–145.

Staines, W., Atmadja, S., & Fibiger, H. (1981). Demonstration of a pallidostriatal pathway by retrograde transport of HRP-labelled lectin. *Brain Research*, *206*, 446–450.

Tabareau, N., & Slotine, J. (2006). Notes on contraction theory. Arxiv preprint nlin.AO/0601011.

Tepper, J., & Bolam, J. (2004). Functional density and specificity of neostrital interneurons. *Current Opinion in Neurobiology*, *14*, 685–692.

Tepper, J., Koós, T., & Wilson, C. (2004). Gabaergic microcircuits in the neostriatum. *Trends in Neuroscience*, *11*, 662–669.

Tyrrell, T. (1993). The use of hierarchies for action selection. *Adaptive Behavior*, *1*(4), 387–420.

Wu, Y., Richard, S., & Parent, A. (2000). The organization of the striatal output system: A single-cell juxtacellular labeling study in the rat. *Neuroscience Research*, *38*, 49–62.

Zhang, D., & Nagurney, A. (1995). On the stability of projected dynamical systems. *Journal of Optimization Theory and Applications*, *85*(1), 97–124.