

# A System-Level Model of Noradrenergic Function

Maxime Carrere, Frédéric Alexandre

► **To cite this version:**

Maxime Carrere, Frédéric Alexandre. A System-Level Model of Noradrenergic Function. 25th International Conference on Artificial Neural Networks (ICANN), Sep 2016, Barcelona, Spain. pp.214 - 221, 2016, <10.1007/978-3-319-44778-0\_25>. <hal-01401890>

**HAL Id: hal-01401890**

**<https://hal.inria.fr/hal-01401890>**

Submitted on 23 Nov 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A System-level Model of Noradrenergic function

Maxime Carrere<sup>1,2,3</sup> and Frédéric Alexandre<sup>2,1,3</sup>

<sup>1</sup>LaBRI, Université de Bordeaux, Bordeaux INP, CNRS, UMR 5800, Talence, France

<sup>2</sup>Inria Bordeaux Sud-Ouest, 200 Avenue de la Vieille Tour, 33405 Talence, France

<sup>3</sup>IMN, Université de Bordeaux, CNRS, UMR 5293, Bordeaux, France

`frederic.alexandre@inria.fr`

**Abstract.** Neuromodulation is an interesting way to display different modes of functioning in a complex network. The effect of Noradrenaline has often been related to the exploration/exploitation trade-off and implemented in models by modulation of the gain of activation function. In this paper, we show that this mechanism is not sufficient for system-level networks and propose another way to implement it, exploiting reported inhibition of a striatal region by Noradrenaline. We describe here the corresponding model and report its performances in a reversal task.

**Keywords:** Neuromodulation, bio-inspiration, decision making

## 1 Introduction

In neuromodulation, a principle of neural activation already observed in crustacea [5], a small set of neurons projects to most regions of the brain and can modify their functioning and learning modes by acting on the intrinsic properties of neurons and on the synaptic weights. Neuromodulators have been popularized in the modeling domain by a paper by K. Doya [7], proposing how different phases of reinforcement learning might be implemented by global signals representing such neuromodulators, where “dopamine signals the error in reward prediction, serotonin controls the time scale of reward prediction, noradrenaline controls the randomness in action selection, and acetylcholine controls the speed of memory update” (quoted from [7]).

Concerning dopamine [12] and acetylcholine [13], the role of these neuromodulators has been defined more precisely, relying on experimental data and on more precise or more biologically informed models. In this paper, we propose to revisit the role of noradrenaline (or norepinephrine, NE) and particularly of its effects on other brain regions. Whereas an excitatory attentional effect is generally reported for NE, we mention here a new inhibitory effect on a specific striatal region and explain why, in the brain and also in models, this additional effect is important for the global dynamics of the network. In the next sections, we introduce more precisely some information about noradrenaline and the way it is presently integrated in models, including data and results that will be important for our model that is subsequently presented together with simulation results.

## 2 The noradrenergic system

NE originates mainly from the Locus Coeruleus (LC), a brainstem nucleus [4]. One acknowledged role of NE is to modify sensory processing in the thalamus and the cortex [15, 16], depending on the level of arousal and attention required by the external situation, proposed to be encoded by the tonic levels of NE [4]. At low level, the animal is at rest (sleeping or grooming). A highly salient stimulus (for example reliably announcing a reward) is going to increase tonic NE level and trigger phasic NE burst to precisely focus attentional processing on that stimulus and resist to distractors. The highest tonic levels of NE are observed when the conditions are no longer predictable (for example in unknown or changing environments) and require to explore among possible relevant stimuli to extract new contingencies [2].

This general view relating the level of NE to the level of arousal is consistent with its often mentioned implication in choosing between exploitation and exploration (with higher levels) of sensory criteria to select actions [2]. This is also consistent with the reported implication of NE during reversal [1] when a sensory criterion to predict a reward becomes suddenly invalid and requires to look for another predictive sensory criterion. This has been termed unexpected uncertainty in [17], in contrast to expected uncertainty, corresponding to the stochasticity of the criterion and encoded by another neuromodulator, acetylcholine, and requiring only patience and not reconsideration of the criterion.

Going deeper in the description of LC afferents and efferents can allow for a more precise interpretation of the role of NE, based on information available in LC to decide on the release of NE and the nature of NE actions in LC targets. Inputs to LC are of three kinds. Low level signals from peripheral centres give basic information about level of arousal from the sympathetic system and about salient sensory inputs from the oculomotor system [4]. More elaborated elements of information are sent by the central nucleus of the amygdala and the medial prefrontal cortex towards LC [16]. They are generally believed to contribute to evaluate the nature of the present situation, and correspond to information like reward history or response conflicts and errors [2]. Other noteworthy inputs to LC are from other neuromodulatory centres which reciprocally influence LC [16].

LC projects to most brain regions and more heavily to attentional structures like the parietal sensory cortex, where NE can enhance evoked activity [3]. Importantly, the basal ganglia is the only cerebral structure not receiving projections from LC, except for the shell region of the nucleus accumbens, where NE is reported to have an inhibitory effect [11].

## 3 Modeling the role of noradrenaline

In addition to [7, 17] other modeling papers have proposed to implement NE mechanisms. McClure and colleagues [10] propose that the level of NE is estimated by an evaluation function depending on the reward rate (corresponding to

input of the orbitofrontal cortex to LC) and measures of response conflict (corresponding to input of the anterior cingulate cortex, ACC, to LC), computed from two windows of long term and short term history of activity.

The model by Aston Jones and colleagues [2] proposes a mechanism implementing the trade-off between exploration and exploitation, with the Drift Diffusion Model (DDM). This model can be applied for tasks with two choices with two units acting as accumulators, integrating over time possibly noisy signals favoring each choice and responding when the difference of levels exceeds a threshold. Interestingly, DDM has been originally proposed to reproduce reaction times and error rates in decision making processes but appears to explain well neuronal responses recorded during such processes. In the model, the value of the threshold is of course an important parameter, but also the gain of accumulator integration that can be modified to reach the threshold at different speeds. This simple model, which can be equivalently implemented with units in mutual inhibition, has been shown to be a good approximation of the optimal decision [6] but remains limited to two-alternative choices. It has also been shown to maximize the signal-to-noise ratio in the difference between the input signals and, when this ratio changes, [6] shows that the gain of the units (or their mutually inhibiting strength) must be modified accordingly.

It is proposed in [2] that, when tonic NE is released (in a situation identified as unexpected uncertainty), the gain of the sensory units will increase and this will facilitate random activation of sensory neurons and accordingly exploration. Nevertheless, we have observed some limitations, trying to apply DDM to act not only on a sensory layer of units representing candidate stimuli (as it is the case in [2, 6]) but on the sensory part of a neural network learning to associate a representation of the value of stimuli to the best response, presented in [8] as an implementation of decision making in the brain.

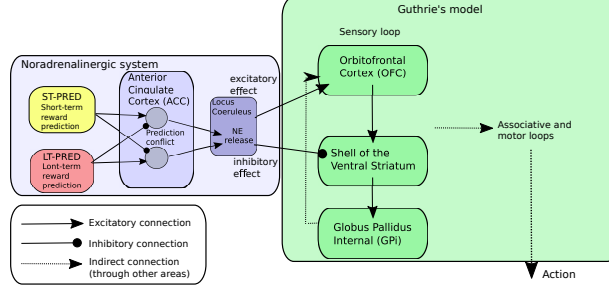
In the [8] network, associations have been learned between sensory neurons and neurons triggering the actions. Even if by NE gain increase a new stimulus is more activated, associative weights might compensate and trigger the habitual action. To tell it differently, an excitatory noise in the sensory layer does not necessarily trigger motor exploration.

## 4 Our model

We have mentioned above a biological fact that has not yet been exploited in models, the inhibitory effect of NE in the shell [11]. In fact, the shell is a striatal region known to participate in the evaluation of the value of stimuli in the orbitofrontal cortex which can in turn activate motor responses toward stimuli. We can consequently explore another alternative of NE neuromodulation effect, where its action in the shell can inhibit previously learned sensorimotor associations.

Our model uses the DANA library for neuronal representation and computation [14]. It extends the model presented in [8] by studying the effect of exploration and tonic NE on it. All the code for the model and parameters are open-

source and available online at <https://github.com/carreremax/basal-ganglia-ne>. We will only describe and discuss here changes made from the Guthrie model.



**Fig. 1.** Main features of our model compared to [8]. ST-PRED and LT-PRED are respectively short-term predictor and long-term predictor which predict reward arrival as the average reward from a long and short amount of trials. Each prediction inhibits the excitatory input of the other in ACC, resulting in ACC activation and NE release in LC only in case of discrepancy between the predictions, ie. in case of unexpected uncertainty. NE release is then used to trigger exploration in [8], by facilitation of cortical excitation and inhibition of striatal inputs.

The level of NE is computed as the difference (or conflict) between a slow and a fast predictor as follows:

The noradrenaline system receives inputs from short and long-term reward predictors, as shown in figure 1. Short and long-term predictors are computed as the average reward on the respectively  $n_{st\_trials}$  and  $n_{lt\_trials}$  last trials.

$$lt\_prediction = (\sum_{k \in n_{lt\_trials}} reward_k) / n_{lt\_trials}$$

$$st\_prediction = (\sum_{k \in n_{st\_trials}} reward_k) / n_{st\_trials}$$

with  $reward_k$  the reward received at trial  $k$ .

These rewards are respectively sent as inputs to two ACC units computing the conflict between the two predictions,  $ne_s$  and  $ne_l$ :

$$\frac{dU_{ne_s}}{dt} = \tau * (-U_{ne_s} + st\_prediction - lt\_prediction)$$

Similarly :

$$\frac{dU_{ne_l}}{dt} = \tau * (-U_{ne_l} + lt\_prediction - st\_prediction)$$

So the long-term prediction is inhibiting the  $ne_s$ , and the short-term one is inhibiting  $ne_l$ . The level of NE release,  $ne$ , is then taken as the sum of  $ne_s$  and  $ne_l$  activities. As a result, if both old and recent predictions are not predicting any

reward, neither  $ne_s$  nor  $ne_l$  activities are strong, and then NE concentration is low. Symmetrically, if both systems are predicting rewards, the two predictions will inhibit the projections of each other, thus resulting in a low NE release. However, if only one system is predicting a reward, ie if the prediction following recent history is different from the prediction based on long history, the corresponding NE population will have a strong, non-inhibited activation, thus triggering a high level of NE release, corresponding to the fact that the reward contingency has recently changed.

Consistent with previous models, NE effect at the cortical level is an excitatory gain :

$$\frac{dV_{ctx}}{dt} = f(U_{ctx} * (1 + ne) * (1 + noise))$$

where  $V_{ctx}$  and  $U_{ctx}$  are respectively the firing rate and membrane potential of cortical neurons,  $f$  and  $noise$  respectively the sigmoid function and activation noise used in [8]. NE inhibitory effect is an original mechanism added in our model, and impacts the output gain of projection from cortex to shell :

$$gain = g_{ctx\_cog\_str\_cog} * ne\_modulation$$

with  $g_{ctx\_cog\_str\_cog}$  the constant gain between cortex and striatum in the sensory loop, and  $ne\_modulation$  the modulatory effect of NE.

$$ne\_modulation = \max(0.5, 1 - ne\_efficiency * ne)$$

NE modulatory effect is limited to halving excitatory projections from cortex to shell, consistent with the effect of NE observed in [11].  $ne\_efficiency$  is a constant set to 0.8, so that only maximum values of  $ne$  will provoke a minimum value of  $ne\_modulation$ .

| Architectural parameters |  |         |
|--------------------------|--|---------|
| Parameter                | Meaning  | Value   |
| $init\_critic$           | Initial values of critic's predictions                           | 0.25    |
| $\alpha\_critic$         | Learning rate of the critic                                      | 0.2     |
| $\alpha\_LTP$            | learning rate for long term potentiation                         | 0.0001  |
| $\alpha\_LTD$            | learning rate for long term depression                           | 0.00005 |
| $g\_ctx\_cog\_str\_cog$  | gain from cognitive cortex to cognitive striatum                 | 1.2     |
| $g\_ctx\_cog\_str\_ass$  | gain from cognitive cortex to associative striatum               | 0.3     |
| $g\_ne\_exc$             | gain of excitatory projections in NE populations                 | 1.0     |
| $g\_ne\_inh$             | gain of inhibitory projections in NE populations                 | -1.0    |
| $n\_st\_trials$          | Number of trials taken into account for the short-term predictor | 3       |
| $n\_lt\_trials$          | Number of trials taken into account for the long-term predictor  | 30      |

**Fig. 2.** Description and values of the parameters added or modified compared to [8].

One of the main problems with unexpected uncertainty-based exploration is the learning rate of the model. If the learning speed of the network is too slow,

the network will perform exploration, but will not be able to learn based on this exploration. At the opposite, if the learning is fast enough to learn from a few trials of exploration, it may converge too quickly, which can lead to sub-optimal choice or stability issues. To address this problem, we modified the learning rate of the critic module in the Guthrie’s model, from previously 0.025 to 0.2, so that the critic could learn based on exploration, and we added critics’ prediction as a sensory input of the network. Consequently, exploration helps the critic to learn insights of the values of alternative strategies, and such values are taken into account in the sensory loop, which in turn helps the network to choose the relevant alternative strategies, and learn from it.

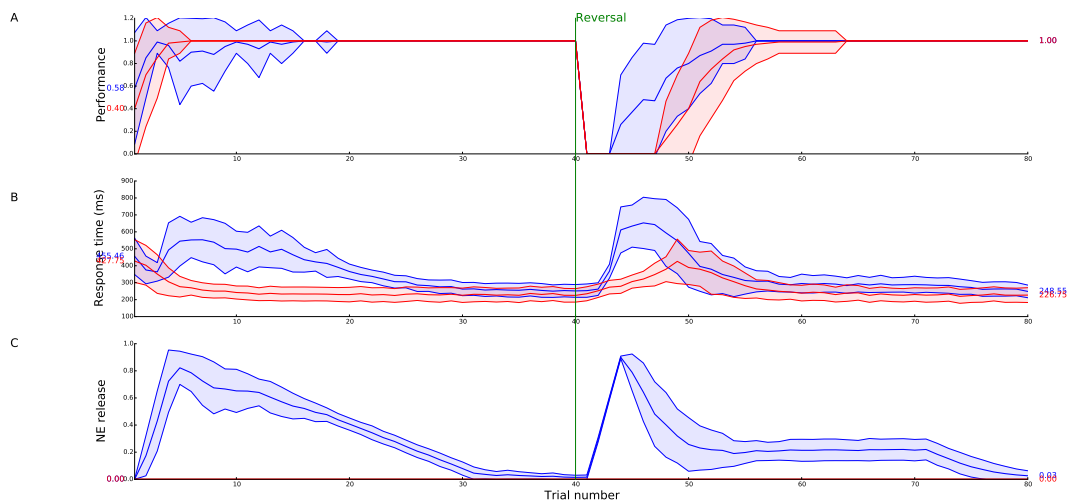
## 5 Experiments

In order to analyze NE effects in the cortex and the shell, together with the switch between exploration and exploitation, we tested our model on reversal learning. At each trial, two sensory CS are simultaneously presented to the network, on two random positions during 2500ms and the network has to perform an action toward one of the CS. As soon as the model performs an action, reward is distributed accordingly to the reward probability of the chosen CS. If no CS is selected after the 2500ms of presentation, the network will not receive any reward. Then neural activities go down to their initial values, and we proceed to the next trial. The acquisition phase consists in 40 trials, in order to perform over-training and to allow habit formation. In each trial, one CS is systematically rewarded while the other is not. During the reversal, which lasts for 40 trials, reward rates for each CS are switched, so the network has to detect the change in reward contingencies and to switch to the other CS.

In fig.3, we report the average performance and decision time on 100 reversal experiments with and without NE release. Each experiment is performed with a “naive” model. The model correctly learns to choose the best rewarded CS during the exploration and reversal phases. However, NE release allows to perform random exploration, and to gradually learn from this exploration, resulting in a faster convergence than networks without exploration (fig.3.A). In addition NE release also increases the decision time of the model during the first trials of reversal (fig.3.B). This is in accordance with [9] results, showing that animals with NE depletion respond with greater rapidity when perseverating. Fig.3.C shows the release of noradrenaline during trials, which is indeed proportional to unexpected uncertainty, with a peak at the reversal onset.

## 6 Conclusion

In this paper, we have reported a model and associated experiments that illustrate the interest of neuromodulation, as a way to modulate existing networks, instead of complexifying their architecture. This is particularly the case for noradrenaline, and its confirmed role in the trade-off between exploration and exploitation. As illustrated in our experiments, NE-based exploration increases



**Fig. 3.** Reversal experiments for our decision making model with NE effects (blue) or without NE effects (red). Each curve is the average of 100 experiments performed each time with a “naive” model. Surrounding shaded areas indicate the standard deviation for each curve. (A) Average performance by trials. Both NE and non-NE models are able to acquire CS values and to learn reversal. Exploration allows NE model to detect quicker the change in reward contingency, and to correctly perform faster than the non-NE one. (B) Average convergence time. During the first trials of reversal, exploration by inhibition of the striatum induces a larger response time for NE model. (C) Average release of NE. NE release is important at the beginning of exploration, and larger during the first trials of reversal. It correlates with unexpected uncertainty.

the convergence speed of a decision network in an unexpected situation, which is a decisive adaptive property for animals and other autonomous systems.

We have pointed out that another solution for NE-based exploration can be the inhibition of learnt rules rather than the excitation of the sensory gain, and have shown a biologically-inspired, neuronal implementation of it, using reported NE effect in the shell [11]. Its excitatory effect in the cortex, widely used in other models, is still present here. Yet, simulation of the model with only striatal NE (not reported here) still exhibits exploration, but with longer decision time. We hypothesize here that because the task does not require exploration of additional representation in the cortex, cortical NE is not necessary for exploration. If the task needs discovery and creation of adequate cortical representation, like for extra-dimensional shift, it would need cortical NE. Another prediction to be tested in both computational and experimental neuroscience is that inhibiting NE release in the shell should both impede exploration and decrease the decision time. These predictions are explored in ongoing work.



## References

1. Aston-Jones, G., Rajkowski, J., Kubiak, P.: Conditioned responses of monkey locus coeruleus neurons anticipate acquisition of discriminative behavior in a vigilance task. *Neuroscience* 80(3), 697–715 (Jul 1997)
2. Aston-Jones, G., Cohen, J.D.: An integrative theory of Locus Coeruleus-Norepinephrine function: Adaptive Gain and Optimal Performance. *Annual Review of Neuroscience* 28(1), 403–450 (2005)
3. Aston-Jones, G., Rajkowski, J., Cohen, J.: Role of locus coeruleus in attention and behavioral flexibility. *Biological Psychiatry* 46(9), 1309–1320 (Nov 1999)
4. Berridge, C.W., Waterhouse, B.D.: The locus coeruleus-noradrenergic system: modulation of behavioral state and state-dependent cognitive processes. *Brain Research Reviews* 42(1), 33–84 (2003)
5. Bouret, S., Sara, S.J.: Network reset: a simplified overarching theory of locus coeruleus noradrenaline function. *Trends in Neurosciences* 28(11), 574–582 (2005)
6. Brown, E., Gao, J., Holmes, P., Bogacz, R., Gilzenrat, M., Cohen, J.D.: Simple Neural Networks that optimize decisions. *Int. J. Bifurcation Chaos* 15(03), 803–826 (Mar 2005), <http://dx.doi.org/10.1142/s0218127405012478>
7. Doya, K.: Metalearning and neuromodulation. *Neural Networks* 15(4-6), 495–506 (Jun 2002), [http://dx.doi.org/10.1016/s0893-6080\(02\)00044-8](http://dx.doi.org/10.1016/s0893-6080(02)00044-8)
8. Guthrie, M., Leblois, A., Garenne, A., Boraud, T.: Interaction between cognitive and motor cortico-basal ganglia loops during decision making: a computational study. *Journal of Neurophysiology* 109(12), 3025–3040 (Jun 2013)
9. Mason, S.T., Iversen, S.D.: An investigation of the role of cortical and cerebellar noradrenaline in associative motor learning in the rat. *Brain Research* 134(3), 513–527 (Oct 1977), [http://dx.doi.org/10.1016/0006-8993\(77\)90826-5](http://dx.doi.org/10.1016/0006-8993(77)90826-5)
10. McClure, S., Gilzenrat, M., Cohen, J.: An exploration-exploitation model based on norepinephrine and dopamine activity. In: Weiss, Y., Schölkopf, B., Platt, J. (eds.) *Advances in Neural Information Processing Systems* 18, pp. 867–874. MIT Press (2006), [http://www.csmb.princeton.edu/~smcclure/pdf/MGC\\_NIPS.pdf](http://www.csmb.princeton.edu/~smcclure/pdf/MGC_NIPS.pdf)
11. Nicola, S.M., Malenka, R.C.: Modulation of synaptic transmission by dopamine and norepinephrine in ventral but not dorsal striatum. *Journal of neurophysiology* 79(4), 1768–1776 (Apr 1998), <http://view.ncbi.nlm.nih.gov/pubmed/9535946>
12. O’Reilly, R.C., Frank, M.J., Hazy, T.E., Watz, B.: PVLV: The primary value and learned value Pavlovian learning algorithm. *Behavioral neuroscience* 121(1), 31–49 (Feb 2007), <http://dx.doi.org/10.1037/0735-7044.121.1.31>
13. Pauli, W.M., O’Reilly, R.C.: Attentional control of associative learning—a possible role of the central cholinergic system. *Brain Research* 1202, 43–53 (Apr 2008)
14. Rougier, N.P., Fix, J.: DANA: Distributed (asynchronous) Numerical and Adaptive modelling framework. *Network: Computation in Neural Systems* 23(4), 237–253 (Dec 2012)
15. Sadacca, B.F., Wikenheiser, A.M., Schoenbaum, G.: Toward a theoretical role for tonic norepinephrine in the orbitofrontal cortex in facilitating flexible learning. *Neuroscience* (Apr 2016)
16. Sara, S.J., Bouret, S.: Orienting and Reorienting: The Locus Coeruleus Mediates Cognition through Arousal. *Neuron* 76(1), 130–141 (Oct 2012)
17. Yu, A.J., Dayan, P.: Uncertainty, Neuromodulation, and Attention. *Neuron* 46(4) (2005)