



# Feature and structural learning of memory sequences with recurrent and gated spiking neural networks using free-energy: application to speech perception and production II

Alexandre Pitti, Mathias Quoy, Catherine Lavandier, Sofiane Boucenna

## ► To cite this version:

Alexandre Pitti, Mathias Quoy, Catherine Lavandier, Sofiane Boucenna. Feature and structural learning of memory sequences with recurrent and gated spiking neural networks using free-energy: application to speech perception and production II. 2019. hal-02140049

**HAL Id: hal-02140049**

**<https://hal.science/hal-02140049>**

Preprint submitted on 26 May 2019

**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.

# Feature and structural learning of memory sequences with recurrent and gated spiking neural networks using free-energy: application to speech perception and production II

Alexandre Pitti<sup>1\*</sup>, Mathias Quoy<sup>1</sup>, Catherine Lavandier<sup>1</sup> and Sofiane Boucenna<sup>1</sup>

**Abstract**—We present a framework based on iterative free-energy optimization with spiking neural network for modeling the fronto-striatal system (PFC-BG) for the generation and recall of audio memory sequences. In line with neuroimaging studies done in the PFC, we propose a genuine coding strategy using the gain-modulation mechanism to represent abstract sequences based on the rank and location of items within them only. Based on this mechanism, we show that we can construct a repertoire of neurons sensitive to the temporal structure in sequences from which we can represent any novel sequences. The free-energy optimization is used then to explore and to retrieve the missing indices of the items in the correct order for executive control and compositionality. We show that the gain-modulation permits the network to be robust to variabilities and to have long-term dependencies as it implements a gated recurrent neural network. This model, called Inferno Gate, is an extension of the neural architecture INFERNO standing for Iterative Free-Energy Optimization of Recurrent Neural Networks with Gating or Gain-modulation. In experiments done with an audio database of ten thousand MFCC vectors, Inferno Gate is capable to encode efficiently and retrieve chunks of fifty items length. We discuss then about the potential of our network to model the features of the working memory in PFC-BG loop for structural learning, goal-direction and hierarchical reinforcement learning.

## I. INTRODUCTION

During early development, infants are keen on grasping structure in several core domains [1], [2], inferring causal models and hypothesis as little scientists [3], [4]. They develop rapidly knowledge about numerosity, space, physics and psychology but it is only around 8 months that they gain aptitudes to make complex sequences and to retain structural information in environment.

In language acquisition, this skill is central for words segmentation and for detecting grammatical and ungrammatical sentences [5], [6]. For instance, infants are sensitive to the temporal order of events in spoken words and in music so that they can be surprised if one syllable is changed or if one sound is removed, violating their prior expectation [7].

### A. Neural foundations

It is at this period also that the prefrontal cortex (PFC) develops. The prefrontal circuits describe a working memory for executive control and planning that evaluates sequences online based on uncertainty [8] and select/unselect them according to the current context or create new ones if any are satisfying [9], [10], [11].

More than any other brain areas, the PFC can extract abstract rules and parametric information within structured data in order to achieve a plan [12], [13], [14]. This aspect makes it particularly important for problem solving tasks, language and maths [15], [16], [17], [18].

Experiments done on subjects performing hierarchical tasks such as drawing a geometrical figure [19], [20] or detecting temporal patterns within action sequences [21], [22] have permitted to identify some properties of PFC neurons for features binding and higher-order sequence planning. In series of observations done on PFC neurons, a critical finding was that sequences were encoded through a conjunctive code, which crosses items and serial orders [23]. In similar experiments done by Inoue and Mikami, some PFC neurons were found to modulate their amplitude level with respect to the position of items during sequential presentation of two visual shape cues [24]. The PFC neurons displayed graded activity with respect to their ordinal position within the sequence and to the visual shapes; e.g. first-ranked items, or second-ranked items. In more complex tasks, PFC neurons were found to trigger at particular moments within the sequence [13]; eg the beginning, the middle, the end, or even during all the evolution of the sequence.

Despite these findings, the precise role played by conjunctive cells in the PFC and the mechanisms behind are still under investigations. In contrast, the conjunctive cells in the Parietal Cortex are more studied and many neuro-computational models explain how they can serve for spatial representation [25], [26], for coordinate transformation [27], [28] and for numerosity capabilities [29]. In most researches, conjunctive cells or gain-modulation neurons in parieto-motor neurons are seen as a way to bind different informations from each other (eg vision and proprioception) for realizing an action (eg target reaching). In [30], [28], Pouget proposes that gain-modulated conjunctive cells in the Parietal Cortex can serve as basis functions for constructing any spatial metric; eg, a hand-centred relative metric [31], [32], a head-centred relative metric [25]. Similar to the role played by conjunctive cells in the the spatial domain in the Parietal Cortex, we suggest that the conjunctive cells in PFC play a role of basis functions in the temporal domain to decompose and code sequences. Gain-modulation in PFC may serve to extract temporal patterns and to represent them as primitives for encoding any existing sequences or for generating any new ones, see Fig. 1 a-b).

This idea is in line with comparative neuroanatomical

<sup>1</sup> Universit Paris-Seine, Universit de Cergy-Pontoise, ENSEA, CNRS, France surname.name@u-cergy.fr

studies who attributes to the parietal cortex and to the prefrontal cortex similar functions to represent relative metrics or conjunctive representations [33] such as order with relative duration, order with relative distance but only the PFC is in position to generate goal-based on context [34]. It is also suggested by Botvinick and Watanabe in [35] who proposed that these cells in PFC describe a compressive representation of sequences without items. Gain-modulated conjunctive cells can give an insight how the PFC manages to plan sequences and encode them [17]. For instance, they may be seen as a solution to disentangle the features (items) from the sequence (ordinal information) in planning. Following this idea, they may gate information at particular moments. That is, not only predicting which action to perform but also knowing *when* within a sequence [36], [37]. Their role may be in line with other frameworks in which neuronal “pointers” or “fillers” or “timestamp” neurons are proposed for binding or gating information with respect to the current inputs [38], [39], [40], [41].

Furthermore, this mechanism may serve to construct a basis for composing any sequences, to recombine items in different order or to generate any novel sequences with different items, see Fig. 1 b). This capability of combinatorial re-use is particularly robust and specific to human behaviors, which corresponds to the capacity of compositionality, hierarchy and of systematicity found in languages and structured grammars [42], [43], [44], [45], [46], [47], [17].

### B. State of art for prefrontal neural models

In prefrontal models for sentence processing, Dominey proposed earlier versions of echo-state networks to model the associative memory in the corticostriatal system for language processing and sequence learning [48], [49], [50].

These architectures differ from other types of recurrent neural networks for serial recall and online activation, which can show nonetheless complex sequences ordering [51] and error-based predictive coding [41], [52], [53].

Different PFC neural architectures from dynamical systems theory have been proposed to code and retrieve memory sequences based on phase synchronisation for feature binding such as the LISA architecture [54], on chaotic networks as in [55], [56], [57], or on reservoir computing networks [58]. For instance, echo-state networks have been utilized for modeling the learning of structure and the acquisition of a grammar of rules [50], [59], [60], [61]. Despite many advantages as dynamical systems to embed attractors, the learning phase is almost done offline with supervised learning for labelling the patterns. Meanwhile, they have to be initialized properly to be effective and the way how structural information (topology) is embedded within these networks is also not clearly defined as it is often used as a black box. Furthermore, it is not clear also how they can support other coding strategies such as the gating mechanism or other learning mechanisms such as reinforcement learning.

Jun Tani proposed neural architectures to manage dynamics with recurrent neural networks at multiple temporale scales (MTRNN) or with parametric bias (RNNPB) for

learning the attractors (abstract temporal structures) in sequences [43], [62], [46]. In neurobotic experiments, within the framework of dynamical systems and chaos theory, he showed how recurrent neural networks can embed several dynamics as symbolic units (rhythmical and sequential) for robot control, imitation and social interaction, giving rise to compositionality.

In contrast to these algorithms, another family of PFC models is based on a gating or gain-modulation mechanism. In the literature, this corresponds to the Long-Short Term Memory (LSTM) [63], [64], [14], the gated prefrontal networks with stripes by O’Reilly and colleagues [65], [10], the SPAWN architecture with neuronal pointers by Eliasmith and colleagues [39], [66], or the prefrontal architectures that use explicitly gain-modulation as in [35] and in [67].

O’Reilly and colleagues attribute to the PFC neurons the role of variable binders to identify rules in the sentences (subjects, verbs, complements) and to process new ones by filling the holes with current values (eg binding new words) [11], [68]. Current models of the PFC show the importance of gating networks [63], [10], [11], [67], [14], [39]. Gated information is particularly useful to maintain contextual variables for several cycles in order to reuse it later or to process new memories from it. For some models like the long short-term memory (LSTM) networks [63], these algorithms have proved their robustness in spite of their lack of accessibility and of biological plausibility. In literature, the gating mechanism is mostly understood as an on/off switch for maintaining or shunting memories. In comparison, the gain-modulation mechanism is very similar to a gating mechanism expect that it emphasizes more the binding of the signals from each others. For instance, this *analog* gating can serve to bind the relative order of items within temporal sequences and for retrieving them as suggested in [22], [35], [17].

In more general frameworks without close bio-inspiration, we can cite the works by Kemp, Lake and Tenenbaum who proposed several architectures based on bayesian theory for probabilistic encoding and compositional capabilities [44], [4], [69]. Using a bayesian framework for generating probabilistic models, their model could extract primitives from motor sequences to construct new symbols of the same types, differently combined.

### C. Proposal framework for sequence learning

In this paper, we propose to use the neural architecture INFERNO standing for Iterative Free-Energy Optimization in Recurrent Neural Network for the learning of temporal patterns and the serial recall of sequences. We proposed originally this neuronal architecture to model the cortico-basal ganglia loop [70] for retrieving motor and audio primitives using Spike Timing-dependent Plasticity (STDP) under the framework of predictive coding and of free-energy minimization [71], [72], [73]. Here, we propose to implement a similar free-energy minimization network but this time in the prefrontal-basal ganglia loop for the serial recall of

memory sequences and for the learning of temporal pattern primitives, using gain-modulation instead of STDP.

Gain-modulation will serve to model neurons salient to the temporal order of items and their sequential organization. As explained previously, PFC units depend crucially on this type of coding for serial recall. They support a gain-modulated mechanism to encode jointly items and rank-order information in a sequence [35]. This mechanism of gain modulation is also described as a gating or conjunctive function in other researches [67], emphasizing more the properties to filter out or to hold on information.

Since the working memory is using gating cells instead of STDP, we propose to name it Inferno Gate in order to disambiguate this architecture from our original network. We will show that Inferno Gate is capable to learn temporal primitives sensitive to the serial order of items within sequences, to code abstract temporal sequences with information about items and to retrieve and construct accurately sequences with items with respect to the given serial order information.

#### *D. Prefrontal structures for model-based reinforcement learning*

We justify our neural architecture from the works done on several neural structures in PFC identified for serial recall and the temporal organization of behaviors [74], [33], [50], [15], [16].

Functional imaging studies suggest the PFC provides top-down support for organizing the orderly activation of lower stages of the executive hierarchy in sequences of actions (eg, goal representations). Koechlin and colleagues propose that the PFC subserves executive control and decision making in the service of adaptive behavior [75], [15], [16].

In order to sustain such adaptive behaviors, it has been proposed that the working memory in PFC has to embed mechanisms for flexibility [76], [77] for maintaining memory sequences during long time range in a hierarchical manner and for exploring new behavioral strategies. Such mechanisms have been proposed typically within the inferential bayesian theory [9], [8], [78] and within reinforcement learning framework [74], [79], [15]. Such approach has been extended to PFC models based on predictive coding and to free-energy minimization.

In line with these works and the models proposed in [15], [79], we present a detailed neural architecture in Fig. 2 in which we identify the Broadman area B45, the lateral PFC (IPFC), the dorsolateral PFC (dlPFC), the Orbito-Frontal Cortex (OFC) and the Anterior Cingulate Cortex (ACC) to participate to a model-based RL system for the active inference of memory sequences.

In first, we suggest that the group constituted by B45, IPFC and dlPFC are associated to the representation of the temporal organization of sequences. In second, we suggest the group constituted by the dlPFC, ACC and OFC interact for decision-making, executive control and problem-solving tasks.

We base our assumptions principally from the review papers done in [36], [50], [33], [15], [79]. For instance,

in our schema, the rostral lateral PFC in both monkeys and humans (typically Broadman areas 9/46 in [50], [33]) has been identified for grouping actions from each others, for episodic or temporal control. We associate therefore the Broadman area B45 for the chunking of raw memory sequences coming from parietal and striatal areas.

At the second stage, the lateral PFC (IPFC) appears to proactively build actor task sets from long-term memory that match the context in which the individual is acting [36], [33]. Neurons in IPFC represent the long-term memory of behavioral sequences and the plans or schemas of action [36]. These actor task sets correspond to us to temporal patterns or sequence prototypes as presented in Fig. 1.

At the third stage, the dorsal lateral PFC (dlPFC) appears implicated in the temporal integration of information for the attainment of prospective behavioral goals [36], [15]. Reports suggest the involvement of the dlPFC for order memory in term of choosing the correct sequence among several. In our schema, the dlPFC combines the temporal primitives of the IPFC to have an estimation of the most suited sequence.

The group constituted by the dlPFC, ACC and OFC networks appear involved in a model-based RL working memory for which the anterior cingulate cortex (ACC) seems involved in the motivation to perform goal-directed action, the task context units in dlPFC, whereas the orbitofrontal cortex in value-based decision-making implicated in novel choices. According to Fuster, the orbital PFC, which is well connected with the brainstem and limbic formations, plays a major role in the control of basic drives [36]. OFC might realize the downward trend or cascade of the processing of decided goal-directed actions (concrete sequences) and represent option-specific state values [33]. OFC might involve a competition for decision-making among multiple choices. The ensemble is organized for the serial order encoding of sequences in dlPFC and the exploration and recall of sequences in OFC.

The paper is organized as follows. We will explain first how an analog gating can be done using the rank-order coding algorithm and how gain-modulated neurons can represent a compact code for sequences. In comparison to other gain-modulation architectures that require a one-to-one conversion matrix necessary for multiplicative binding –, which consumes neurons for this computation,– we discovered that a rank-order coding algorithm can model gain-modulation in a more efficient manner with spiking neurons.

We apply this network for the learning of temporal primitives from audio sequences. These primitives are used then for representing and recalling these audio sequences of one second length (1000 milliseconds), corresponding to chunks of 50 items length, despite information about the items identity (their index) was lost.

We discuss then the originality of our approach and implications in terms of computation for modeling sequences, extracting temporal tree structure-like patterns, for compressive coding of grammar-like models, recursive representation, compositionality and transfert learning.

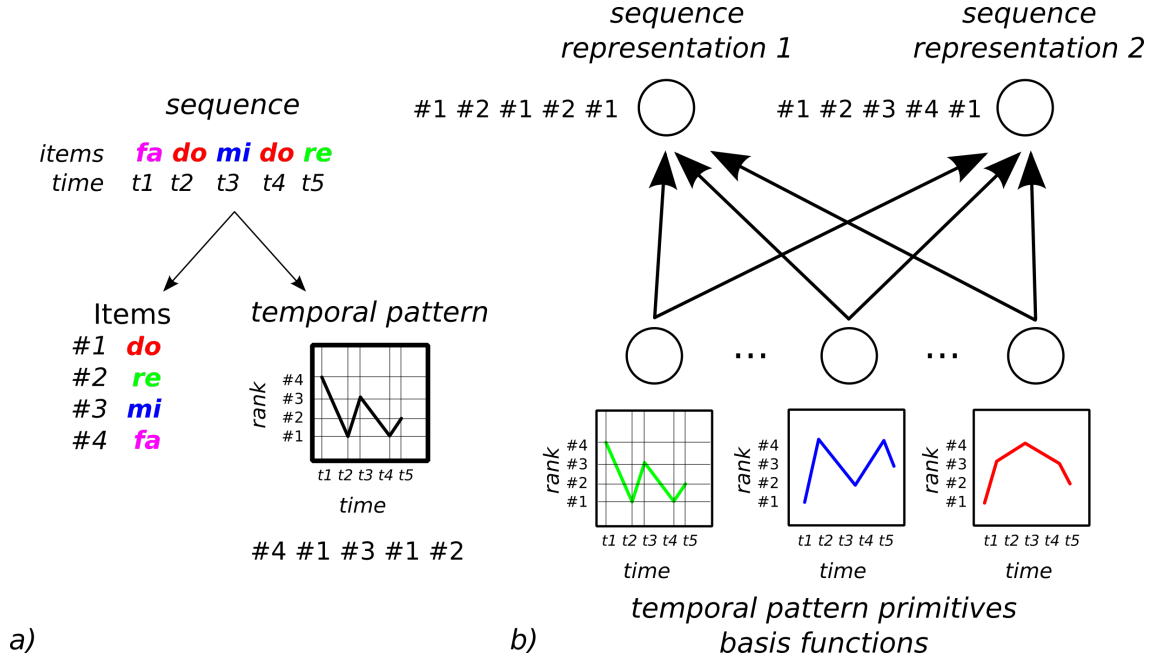


Fig. 1. Gating operation for feature and structural separation in sequence learning. In a), we can discriminate the items' identity (rank) from their position (order) to represent one sequence. By separating the two, we can extract the temporal pattern and arrange items in a different order. Hence, the coding of the temporal pattern can make it robust to variability and can represent many sequences (generalization). This process is operated by a gain-modulation or gating mechanism explained later. In b), the combination of these temporal patterns can serve to compose any novel temporal pattern in the same fashion as radial basis functions would do.

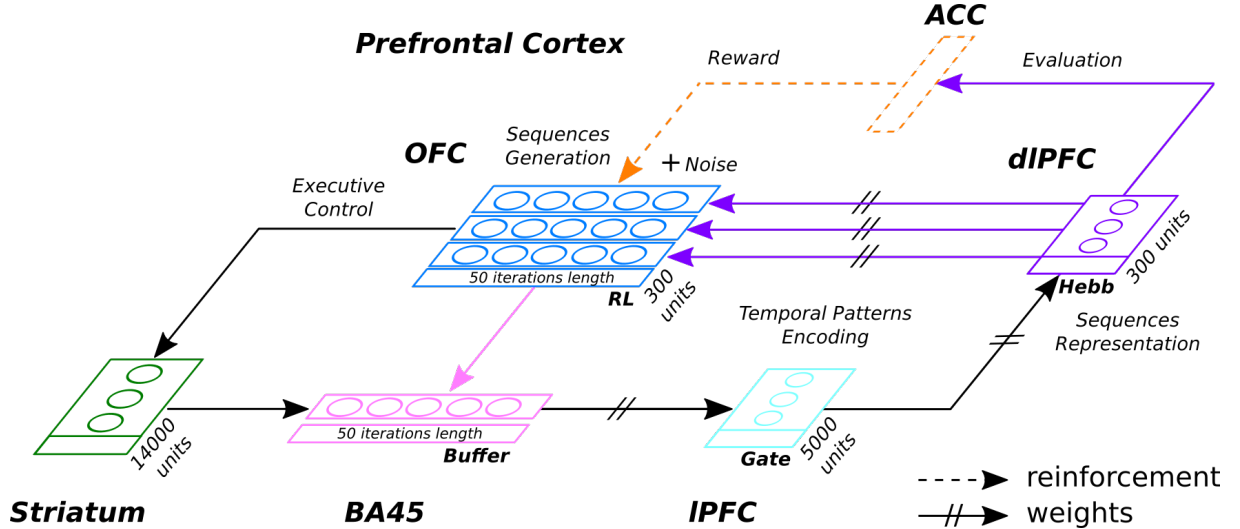


Fig. 2. Framework for sequence learning based on iterative optimization in Fronto-Striatal (PFC-BG) loop. Our putative architecture follows the models proposed by [74], [15]. The Striatum (STR) represents the action/sound units that are assembled into a sequence at the Broadman area B45 in order to form a chunk. This chunk is read by the lateral PFC layer (IPFC), which learns and recognizes the temporal patterns via a gain-modulation/gating mechanism. The different temporal patterns in IPFC are assembled in the dorso-lateral PFC layer (dIPFC), whose units represent each the temporal structure within the sequence at a higher level. Based on the evaluation of the dIPFC, the Orbito-Frontal layer (OFC) and the Anterior Cingulate layer (ACC) generate and select sequences that follow the temporal patterns in dIPFC in order to retrieve the indices lost of STR units for executive control. This reinforcement learning stage corresponds to a free-energy minimization process to reduce error prediction. The framework follows the architecture Inferno proposed for the cortico-basal ganglia (CX-BG) loop.

## II. METHODS

We present in section II-A the neural architecture INFERNO Gate used for serial recall in audio sequences associated with the PFC-Basal Ganglia loop. We describe then in section II-B the coding mechanism used for learning the serial order of items within sequences using the rank-

order coding algorithm for modeling the gain-modulation mechanism with spiking neurons. We define after in section II-C the experimental setup and the parameters used for accurate recall of long-range speech sequence.

### A. the network architecture *Inferno Gate*

The neural architecture *Inferno Gate* reproduces the main configuration of the original *Inferno* network [70] with two coupled learning systems that minimize their mutual prediction error (the free-energy), see Fig. 2. The two networks use both spiking neurons and the difference between the two comes from the types of coding employed to represent temporal events. Namely, the original *Inferno* network employs the spike timing-dependent plasticity mechanism whereas the second one uses gain-modulation.

Considering the global architecture in Fig. 2, the two learning systems (IPFC/dIPFC and dIPFC/OFC) corresponds to two associative networks of spiking neurons (SNNs) similar to radial basis functions. Bidirectionally coupled, the first SNN (IPFC/dIPFC) implements a forward model of the incoming signals while the second SNN (dIPFC/OFC) implements an inverse model aimed at retrieving and controlling those signals. The two learning systems can be viewed as an inverse-forward controller that can be modeled with the function  $Y_{out} = f(I)$  for the first SNN and with the function  $I = g(Y_{out})$  for the second one, in which  $I$  is the input vector and  $Y_{out}$  are the output dynamics.  $I$  is a sequence of Striatal units over time.

In order to minimize error, the second network (dIPFC/OFC) generates intrinsic noise  $I_{noise}$  to control the dynamics of the first one (IPFC/dIPFC) following a reinforcement learning (RL) mechanism. In Fig. 2, this role is devoluted to ACC for error evaluation. The activity level of one unit in dIPFC,  $Y = Y_{out}$ , is compared to its maximum amplitude level  $Y_{max}$  in order to compute the error  $E$  between  $Y_{max}$  and  $Y_{out}$  and the current input  $I(t) = I_{noise}$  is kept for the next step  $I(t+1) = I(t) + I_{noise}$ , if and only if it diminishes the gradient  $\Delta E$ . Over time,  $I$  converges to  $I_{opt}$  its optimum sequence vector, and  $Y_{out}$  converges to  $Y_{max}$  its maximal value. This scheme is in line with predictive coding algorithms and its organization is similar to novel architectures combining two or more competitive neural networks such as auto-encoders or the generative adversarial networks.

We showed in [70] that this variational process is similar to a stochastic descent gradient algorithm performed iteratively. We add here a more sophisticated gradient descent algorithm corresponding to a simulated annealing mechanism in order to account for the neuromodulators involved in decision-making in PFC for uncertainty and surprise [8].

As proposed by [74], adding temporal structure to RL can ease the scaling problem in the exploration process. In the original version of *Inferno*, we found that STDP helped to learn and retrieve temporal chains. Thereinafter, we will show that the gain-modulation can even go beyond for abstracting temporal sequences and to be more robust to variability.

### B. Gain-modulation mechanism based on Rank-Order Coding

The rank-order coding (ROC) algorithm has been proposed by Thorpe and colleagues to model the information process-

ing done in the Visual Cortex by feedforward integrate-and-fire neurons [80]. We have expanded their use to recurrent neural network models in [81], [70] replicating the Spike Timing-Dependent Plasticity learning mechanism.

The main assumption of the ROC algorithm is that spiking neurons performs a quantization of the inputs variable occurring in time discretized with respect to their temporal delays, see Fig. 3 b). The temporal order of the inputs are transcribed into a rank code that is translated into weights value and summed at the neuron's level. The more similar the temporal order of the incoming signals, the higher the amplitude level of the ROC neurons. Reversely, the less similar the sequence order of the incoming signals, the lower the amplitude level of the ROC neurons. Although this mechanism can encode discretized temporal sequences as showed in our previous works, it does not retranscribe a gain-modulation mechanism, a sensitivity to a rank-order within a sequence *independently* to the neurons' identity. We suggest here, as a novel coding strategy using gain-modulation, that we can construct 'compressive codes' of temporal sequences by removing the identity of the neurons (their index) within the temporal sequence and keeping just their ranking order, see Fig. 3 c).

*STDP coding strategy* – If the ROC algorithm computes the neurons activity based on the discretized temporal delays of incoming events, the coding strategy resembles to the STDP learning mechanism with ROC neurons becoming sensitive to the temporal contingency of incoming input; see Fig. 3 b). In this coding scheme, ROC units are considered therefore as contingency detectors inducing phase synchronization [82].

*Gain-modulation coding strategy* – A second coding strategy consists to discretize the serial order of units both in time and space, see Fig. 3 c). Here, the indices of the neurons (or their identities) are no more preserved and it is their rank within the sequence that is taken into account; eg first, second or n-th in the sequence. This strategy reduces drastically the amount of information to process, which makes possible the discovery of an abstract temporal structure disregarding the units indices; eg the sequence becomes a template. This coding mechanism is described as compressive representation by [35]. Hence, since the units identity is not anymore present in the temporal code, it is sensitive to any novel sequences that preserve the global temporal structure.

For instance, in Fig. 3, the temporal encoding of two sequences following the same spatio-temporal pattern is constructed successively by first dismissing the temporal information and then the identity information with the rank-coding algorithm first on the time axis and then on the index axis.

The problem dimensionality for temporal sequences of  $M$  elements is reduced from a continuous time  $\times$  space dimension in  $\mathbb{R}^{2M}$  to an intermediate representation of  $\mathbb{N}^M \times \mathbb{R}^M$  and then to a compressive representation of  $\mathbb{N}^{2M}$ . Although the reduction of complexity does not appear important when looking at the dimensionality of the vector quantization, it permits to represent in a compact way an

infinity of varying spatio-temporal sequences that follow the same structure, which corresponds well to the variable binder property found in PFC neurons.

In comparison to other methods used to code the gain-modulation mechanism as in [35], [53], [83], [84], this one has the advantage to not project the rank code and the position code into an intermediate 2D matrix of complexity  $O(M^2)$ . It does not need also to separate the ranking information into separate modules or stripe codes as proposed in [11], [68].

This property of identity masking appears similar to the idea of variable binding in [68], of timestamp neurons in [85], [52] or of neuronal pointers in [39] as these neurons can fill out any new variables in the correct rank in the sequence. This coding strategy requires therefore two types of units, one for maintaining the input information (variables values) and one for recalling the sequential order. Both are found in the prefrontal cortex for the maintenance of the units activity and for the learning of a task-related activity. The ensemble constitutes the behavior of one Working Memory.

One advantage of the gating strategy compared with the STDP one is that the temporal information is learned separately from the inputs, which enables the network to learn long-range dependencies at an abstract level and to prevent it to loose information less rapidly within a temporal horizon –, which corresponds to the so-called vanishing gradient effect in deep networks. As a remark, feed-forward (deep) networks, standard recurrent neural networks (with/out STDP) or hidden Markov models will easily loose accuracy after several iterations due to the accumulated errors because any errors, noise, delays within a sequence and sensitivity to duration will disrupt the sequence. One explanation why any error introduced in the network makes conventional neural networks brittle is that the state and the temporal information are coded together. This is not the case in neural models with a gating mechanism like PBWMs [68], SPAWN [39] or LSTMs because the temporal information of a sequence can be learned in memory cells separately to the variable values that can be retrieved online or maintained dynamically during an indefinite amount of time.

The equations of the rank-order coding algorithm that we used is as follows. The neurons' output  $Y$  is computed by doing the dot product between the function  $rank()$  sensitive a specific rank ordering within the input signal vector  $I$  and the synaptic weights  $w$ ;  $w \in [0, 1]$ . For a vector signal of dimension  $M$  and for a population of  $N$  neurons ( $M$  afferent synapses), we have

$$Y_n = \sum_m^M rank(I_m) w_{nm}, \forall n \in N \quad (1)$$

We implement the rank function  $rank()$  as a power law of the  $\text{argsort}()$  function normalized between  $[0, 1]$  for modeling the gain-modulation mechanism applied two times on the time axis and on the rank axis. This warranties that the density distribution is borned and that the weight matrix is sparse, which makes the rank-order coding neurons similar

to radial basis functions. This attribute permits to use them as receptive fields so that the more distant the input signal is to the receptive field, the lower is its activity level. The updating rule of the weights is similar to the winner-takes-all strategy in Kohonen networks [86] with an adaptive learning rate  $\alpha_n, \forall n \in N$ . For the best neuron  $Y_b$ , we have:

$$\Delta w_{bm} = \alpha_b(rank(I_m) - w_{bm}), \forall m \in M \quad (2)$$

$$\alpha_b = 0.9 \alpha_b \quad (3)$$

### C. Experiment Setup

We give the implementation details about the striatum-prefrontal working memory modeled by the Inferno Gate architecture. We ascribe to it the role to learn temporal patterns and to represent audio memory sequences, see Fig. 2. The audio database used as input consists on a small audio dataset of 2 minutes length of a native french woman speaker repeating three times five sentences. The audio .wav file is translated into MFCC vectors (dimension 12) sampled at 25ms each and tested with a stride of 10ms. The whole sequence represents 14.000 MFCC vectors, the number of units in the Striatum layer not encoded in the temporal order.

The Inferno Gate architecture is based on the same principle of the Inferno architecture –, the use of noise and reinforcement learning to control a spiking network,– expect that the coding strategy exploits now compressive ranks (the unit's identity is not preserved) and temporal order (sensitivity to the position in the sequence). Here, the B45 area is modeled as a buffer of 50 units length receiving the indices ordered in time of the Striatum layer consisting of 14000 units; the number of coded MFCC in STR. Therefore, each chunk in B45 represents a sequence of 50 MFCCs, corresponding to a chunk of 1250 ms length.

Then, the IPFC layer encodes the ordinal information from the B45 buffer. The IPFC layer consists on 5000 units for which each unit encodes a specific temporal pattern through gain-modulation. Each IPFC unit learns the temporal pattern that follows the serial order within the sequence of 50 units in B45, independent of their true index.

At the next stage, the dIPFC layer combines together the IPFC units to represent abstract sequences. The dIPFC layer consists on 300 contextual neurons. Each unit encodes a compositional representation from the 5000 basis functions in IPFC. The strong dimensionality reduction in IPFC as explained in section II-B and in Fig. 3 permits to learn and generalize rapidly temporal patterns within sequences in dIPFC and to explore rapidly new sequence solutions at high speed.

Finally, the dIPFC units are evaluated by the ACC and a prediction error signal is processed to search and retrieve the optimal sequence in OFC. The OFC layer consists of 300 vectors of 50 iterations length, one vector for each unit in dIPFC. Each vector is generated to retrieve back the corresponding sequence of 50 iterations length with the retrieved STR index values. The OFC vectors are used for the executive control on the Striatum units.

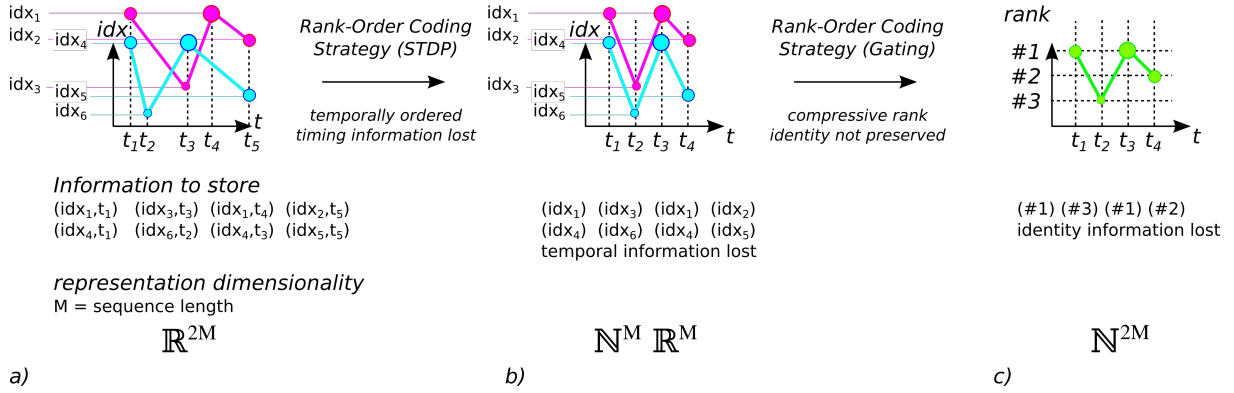


Fig. 3. Rank-order algorithm for compressive rank representation. We describe the two steps process done with the rank-order coding algorithm to model the Spike Timing-dependent Plasticity rule and the Gating mechanism. In a), two sequences in cyan and magenta are represented with different neurons indices, different timing but with same temporal structure (up-down-up-down). In b), the rank-order coding algorithm is used to quantify any sequence in the temporal domain with discrete timing; eg first ranked, second ranked. This is a rough approximation of the STDP rule. The indices of the neurons are kept and only the temporal information is lost. In c), we can use a second time the rank-order coding algorithm to suppress now the neurons identity (their indices) within the sequence in order to keep only their rank within the sequence. This second process permits to have a temporal pattern, a compressive representation of the two sequences in which only the rank order is kept. This second process reduces drastically the amount of information to encode any sequence, irrespective to the neurons' id and to their precise timing. For any sequence of length  $M$ , the problem dimensionality is reduces to  $\mathbb{R}^{2M} \rightarrow \mathbb{N}^{2M}$ .

### III. RESULTS

This section explains the two different experiments done to model the striatum-prefrontal working memory for learning abstract temporal patterns and for retrieving audio sequences through an exploration process.

In the first experiment presented in section III-A, we explain the control done from the striatum on the prefrontal layers to learn temporal patterns and to represent abstract sequences with information flow  $\text{STR} \rightarrow \text{B45} \rightarrow \text{IPFC} \rightarrow \text{dIPFC}$ . In the second experiment presented in section III-B, we explain the exploration process done to generate audio memory sequences using evaluation on prediction error from the encoded abstract sequences previously learned; the information flow goes this time from  $\text{dIPFC} \rightarrow \text{OFC} \rightarrow \text{B45} \rightarrow \text{IPFC} \rightarrow \text{dIPFC}$ .

#### A. Experiment 1 – Learning temporal patterns and encoding abstract sequences

1) *model-based gated control in striatum-prefrontal loops*: Experiment 1 explains the information flow  $\text{STR} \rightarrow \text{B45} \rightarrow \text{IPFC} \rightarrow \text{dIPFC}$  in order to learn respectively the temporal patterns in the IPFC layer and the abstract sequences in the dIPFC layer.

We present in Fig. 4 a) the dynamics of IPFC and dIPFC layers during the learning stage. The activity level of each neuron indicates their sensitivity to the temporal pattern found in the sequence. The neurons in IPFC present a compressive code of the B45 input sequences, which are then combined in the dIPFC layer at a more abstract level. We plot in Fig. 4 b) in the top chart the snapshot of the IPFC population activity taken at one time step and sorted (black line) and in the bottom chart the temporal pattern of the most active neuron in IPFC (red line) and the rank order at the population level (blue line).

The activity level of the IPFC units indicate their saliency to one specific rank order within input sequences. The sorted

activity in black line indicates the coding representation at the population level is not sparse but many neurons are necessary to code the sequences. For instance, because each IPFC unit encodes one temporal pattern at the unit level, the decomposition of one sequence can only be partial as seen in Fig. 4 b) bottom chart, whereas at the population level, the sequence can be represented and discriminated.

2) *Analysis of the sequence encoding in the IPFC network*: After the learning stage of the IPFC and dIPFC units, we can analyze how the ranking information is encoded in the weight matrices. We plot in Fig. 5 a) the weights of the IPFC units (5000 units in Y axis) reordered with respect to their sensitivity to specific positions within B45 sequences (50 iterations length in X axis). The amplitude level of the synaptic weights in the diagonal indicates that each IPFC neuron is sensitive to different positions within the sequences. For instance, neurons with high weight values in the beginning of the sequence will be less sensitive to forthcoming events occurring within the sequence, they will serve as a retrospective template. At reverse, neurons with high weight values in the ending of the sequence will be less sensitive to the past events in the sequence, they will serve as a prospective template. This behavior has been observed in PFC neurons for sequence retrieval [87], [22] as well as in the prefrontal-hippocampal loop in T-Maze tasks [88], [89].

As an example, we plot in Fig. 5 b) the weights distribution of two IPFC units. We select them because these two units have their highest weight value for the position located at the middle of the sequence (position #24), and the other weights located at other positions within the sequence have a lower value. The two circles indicate two positions where the IPFC units have the same weights value: at positions #4 and #24. Based on these weight distributions, we can reconstruct back the temporal patterns for which the IPFC units are the most sensitive to as each weight's value corresponds



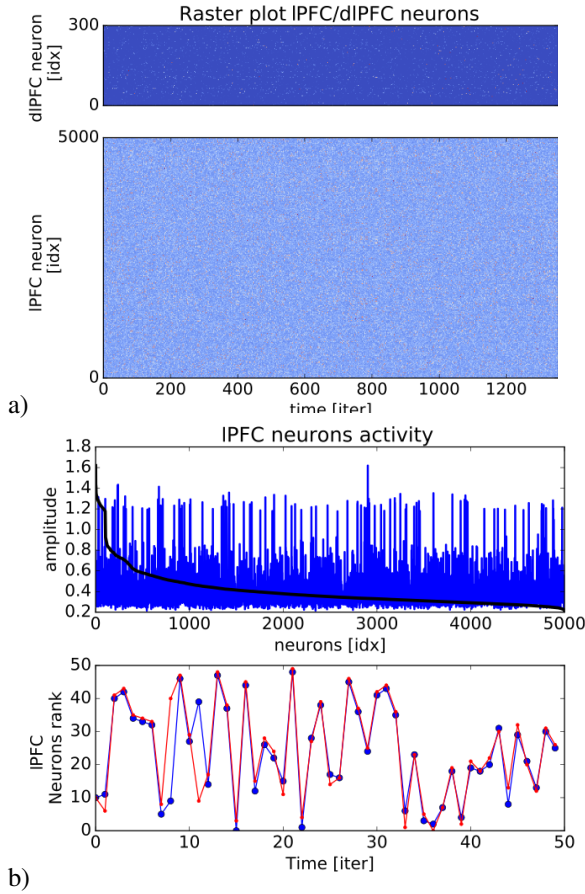


Fig. 4. Encoding of temporal patterns in IPFC/dIPFC layers. In a), each IPFC unit encodes a temporal pattern from the incoming sequences in B45 (bottom chart). The decomposition done at the IPFC layer is a representation of the temporal structure in the sequences that is then learned at the dIPFC layer (top chart). In b) snapshot of the IPFC layer at one particular timing (top chart). The black line indicates the sorted activity of the IPFC units. The red sequence in the bottom chart indicates the temporal pattern for which the most active IPFC neuron is the most sensitive to. The blue sequence indicates the temporal pattern coded at the population level (for all IPFC neurons). This graph indicates that only one unit is not discriminative enough to represent any input sequence and that the coding at the population level is more robust and precise.

to one rank within the sequence. We plot in Fig. 5 c) the two reconstructed temporal patterns. Since the two IPFC units have the same weights at the circle positions, the reconstructed sequences code the same neuron's rank at those positions within the sequence.

This is how in our framework the IPFC neurons retranscribe the gain-modulation mechanism: with respect to the position of one item within a sequence, the IPFC activity level will be modulated with respect to a weight's value depending where is located the corresponding item (neuron's rank) within the sequence.

Depending on the activity level of the different IPFC units for a specific sequence, a decomposition in the IPFC space is represented at the dIPFC level, as in see Fig. 1 b). The decomposition in the IPFC layer permits to represent at a more abstract level, and in a more compact fashion, the compressive rank of the sequence at the dIPFC layer.

## B. dIPFC-OFC Iterative free-energy exploration-optimization

1) *Retrieving memory sequences from incomplete information:* We present in this section the iterative optimization process done at the dIPFC level for retrieving memory sequences at the OFC level with the error rate computed at the ACC. The information flow corresponds to  $dIPFC \rightarrow OFC \rightarrow B45 \rightarrow IPFC \rightarrow dIPFC$ , see Fig. 2.

In order to understand better the global process, we display in Fig. 6 a-c) and in d-f) the iterative optimization process done during 10.000 iterations for two dIPFC units, the dynamics of the IPFC layer and of the ACC unit are showed in a) and d), the final retrieved sequence in OFC with respect to the one represented by the dIPFC units are displayed in b) and e), the raster plot of the iterative search of exact sequences in OFC and B45 are showed resp. c) and f).

In Fig. 6 a) and d) the ACC unit in the top chart represents the error rate, which is the inverse of the activity level of the dIPFC units. The raster plot of the IPFC dynamics is displayed at the bottom chart. The desired sequences we want to reconstruct in OFC are presented at Fig 6 b) and e) in the top chart. The raster plot of the reconstructed OFC/B45 dynamics are plotted in Fig 6 c) and f).

The exploration search is performed after the learning stage done in previous section III-B. Over time, a sequence in OFC is explored iteratively using noise so that the dIPFC activity is maximal activity level and that ACC reaches a minimal value.

We display in Fig 6 b) and e) in red, the retrieved OFC sequence in the top chart and the serial order for which the two different dIPFC neurons are the most sensitive to in the bottom chart.

In the top chart, we can observe that the reconstructed OFC sequences in red follow a similar pattern to the ones in blue although the identity of the neurons is not completely preserved. Nonetheless, we can see that the ordinal information in the bottom chart is matched, which means that the proposed sequence in the top chart follows the temporal pattern encoded in the IPFC and in the dIPFC layers.

Hence, despite the indices in the sequence have been lost in the encoding process, the system is capable to retrieve the memory sequences from incomplete information (due to compressive rank) with small error.

2) *Performance analysis from incomplete information:* In order to analyze the accuracy of the Inferno Gate network, we plot in Fig 7 the euclidean error normalized between  $[0, 1]$  made by the network during recall with respect to the number of items given as input vector resp. a) and with respect to the position within the sequence, resp. b). The exploration stage was limited to 10.000 iterations for each experiment and we plot the retrieved sequences from 0 to 40 items given out of 50 items to retrieve, resp. in Fig. 8 a-e). The grey areas indicate the part of sequence given to the system to reconstitute the missing part.

For 0% of information given, which corresponds to the previous situation in which the system has to retrieve the

whole sequence from scratch, the root error is 0.08 corresponding to 8% error with large variance, see Fig. 8 a).

For 20% of information given, which corresponds to ten items given out of fifty as displayed in Fig. 8 b), the accuracy is not necessarily higher as we would expect and the error rate reaches 0.07 with a higher variance in comparison with the previous case. We analyse this result as the difficulty of the system to go out from local minima with such small constraints added, which is different from the previous situation in which Inferno could freely search for solutions.

For 40% of information given, which corresponds to twenty items given out of fifty as displayed in Fig. 8 c), we can observe a strong decreasing of both the error rate around 0.02 and the variance. This means that the network can retrieve back 60% of the missing information with good accuracy.

Error rate continues to diminish below 0.01 if we provide 60% of the information (30 units), see Fig. 8 d), and serial recall is almost error-free if the network has to retrieve the identity of ten units out of fifty (80% of the information given). We can observe that the order position to which the PFC neurons are sensitive to are all retrieved in the bottom chart although there is some slight errors in the reconstruction.

From additional studies that we did not present here, we observed that it was possible to complete the serial recall for all the cases with an error rate below 0.01 if we continue the exploration search for 50.000 to 100.000 iterations. These results indicate the generalization capabilities of Inferno to separate linearly the input dynamics as we can achieve error-free retrieval.

#### IV. DISCUSSION

We have presented a novel neural architecture based on free-energy minimization using recurrent spiking neural networks for modeling the fronto-striatal (PFC-BG) loop and learning temporal sequences. This network extends our original neural architecture Inferno in [70] aimed at modeling the cortico-basal ganglia (CX-BG) loop for learning motor primitives. Here, we have showed its effectiveness in the more challenging tasks of speech recognition and production. Although the two networks are similar in their functional organization, the encoding type is different. The first network uses the STDP mechanism for learning temporal correlations between spiking events whereas the second one uses a gating mechanism for binding the item's rank and their position within a sequence.

By discriminating content (which sound) and contextual information (when to play it in the sequence), we have showed that the two networks are capable to robustly learn the temporal structure within sequences and to retrieve the items identity in the correct order.

In a complementary paper, we have described the architecture Inferno for modeling the CX-BG structure for sound primitives. The BG network explored and retrieved sound vectors by testing them stochastically through the CX layer. The more the Striatal units recognize and predict

the CX output, the stronger it reinforces its link with the sound vectors encoded in the Globus Pallidus layer, which constitutes at the end of the optimization process one sound repertoire.

Although a stable activity can be retrieved back in a self-organized manner within the CX-BG network, the top-down control of a precise temporal sequence is devoted to another structure, the PFC-BG loop, which selects and influences the first system. In order to model the PFC-BG loop, we reuse the same neural architecture INFERNO but with a different temporal coding to assess the property of the PFC neurons. In contrast to STDP, PFC neurons employ a gain-modulation mechanism to bind multiple information at once; eg the relative position of an item within a sequence for instance. We named this neural architecture INFERNO Gate in consequence.

**Gain modulation**— gain-modulated units learn the order and the rank of one item within a sequence so that any misplacement of it will reduce its activity level. Furthermore, the capacity to encode the items' localization follows a power-law scale. From a computational viewpoint, this power-law scale is one important property in order to construct radial basis functions and to have orthogonal (discriminative) representations coded at the population level. Such coding is apparent to a nonlinear gating as their activity is modulated by the occurrence of multiple informations. The population coding permits in return to have a compressive code that can help the exploration search in a reinforcement learning framework.

This property permits to use the IPFC units as 'fillers' or 'pointers', which adds some variability in the encoding but in a structured way. Thus, we think this property of gain-modulation is interesting in order to learn abstract and temporal information about structure within sequences such as AAB or ABA patterns for which infants are sensitive to very early. Because the activity-level of IPFC units in our system is modulated by multiple information, the item's rank and time order, we suggest that it can be used for representing other relative metrics as suggested in [22], [33] and that gain-modulation can be assimilated to the nonlinear-mixed selectivity mechanism proposed by [90]. For instance, the conjunctive cells in parietal cortex are found to encode relative metrics based on multimodal binding to infer non-trivial information about space and numerosity [91], [35], [92].

In previous neurobotic researches, we have modeled these parietal gain-modulated neurons for visuomotor coordination and for body representation using a more standard coding strategy based on multiplication [93], [53], [83], [84], [94]. However, the number of units necessary to process gain-modulation evolves quadratically with respect to the problem dimensionality. We think spiking neural networks along with the gain-modulation learning mechanism have the potential to represent multimodal information in a compact manner, perhaps even more efficiently as conventional multi-layer feedforward networks do as there is no loss of structural information in the encoding, which is not the case in deep

networks. We can envision some tasks in multimodal integration, which are still difficult to realize with spiking neural networks.

**Retrospective and prospective encoding** – The temporal coding done in Inferno gate extends the STDP mechanism with an extra information making it nonlinear and abstract in the sense that neurons receptive field encode a structural information about the sequence and not the sequence itself. In our experiments, IPFC neurons code for a position in the sequence either in the beginning, middle or end of it. They are nevertheless sensitive to other positions but with less strength.

As some IPFC neurons were sensitive to items in the beginning of a sequence or at the end, this behaviour reflects well the behaviours in PFC to fire to retrospective or to anticipatory events. Retrospective neurons are firing depending on the previous events. Conversely, prospective neurons fire depending on the future events. This properties were shown in hippocampal cells of rats in W-shaped tracks with alternate trajectories [89] or T-Maze with return arms [88]. These results and many others mainly refer to hippocampal activities. However it has been shown that these activities may be modulated by prefrontal information [95], [87]. We make here the prediction that some IPFC neurons exhibit these kind of activities, and that this information is used as a global context for driving neuron firing in the hippocampus [96].

**Compositionality**– the dIPFC units combine the IPFC temporal patterns in a similar way than a radial basis function network: as this layer embeds a variety of temporal primitives, it can encode rapidly any novel sequences. Hence, we suggest that this mechanism of gain-modulation is potentially important in infants for fast inference and for learning abstract patterns with few samples [44], [4]. The learning and error minimization processing of IPFC/dIPFC temporal rules may be similar to logical inference, which attempts to catch up rules in one domain. As these temporal rules are abstracted (no identity information), they may be applied to other domains for grasping other observations; for instance, learning the temporal patterns during motion sequences, like planning, drawing or solving a task. As a beneficial effect, learning higher-order hypothesis may accelerate learning massively in other domains. It may help to learn abstract or contextual words, which is still unachieved challenge by AI architectures.

To summarize, we suggest that this system presents some capabilities suited for learning linguistic systems (eg, a grammar of rules) and timely ordered behaviors. Since Inferno gate encodes temporal patterns in an abstract manner, like AAB or ABA, we may expect that adding another abstracting layer to the model, presumably the Polar Frontal Cortex as proposed in [18], it may be possible to create sequences of sequences, such as ((AAB)BA), mixing two or more temporal patterns in an iterative manner. In this way, our network may be extended to fractal-coding to have a hierarchical representation of sequences at any depth.

**Long-term dependencies** – In experiment 2 in section III-

B, we have showed that it was possible to retrieve accurately in the OFC layer the identity of units and their order in long-range sequences (fifty iterations length) although this information (the neuron’s identity) was not encoded in the dIPFC units. With no external information, the system requires a long period of time above ten thousand iterations to search for the items’ identity as well as their position. All generated sentences present the same temporal pattern as the one we want to retrieve although there is some variability present in them.

This exploration process may be seen as a babbling period in infants development. The explorative search is based on a free-energy minimization process of the OFC vectors based on the evaluation of the dIPFC units computed at the ACC level. We have showed that error minimization is fast and that it requires only several hundred iterations to retrieve the missing items within the sequence. If items are furnished and imposed to the sequence to retrieve (a sequence ‘a trous’), the search is even faster and accurate with respect to the number of items given, below two hundreds iterations.

The generative property of the system show that it incorporates computational capabilities of robustness to noise and to retain long-term dependencies for sequences of fifty iterations length as no information or few is required to retrieve which item has to be performed and when. This property is advantageous in comparison to standard recurrent neural networks. For instance, Inferno Gate behaves differently from classical recurrent networks, including the gated ones, which are directed graphs that attempt to predict the next items depending on the past ones in a Markovian fashion. Our network may extend the idea of vector codes or of vector symbolic architectures by [97] or of neural pointers by [39], [38] or of the “merge” function proposed by Chomsky for encoding nested structures in language [45], [17]. In future research, we should test the performances and computational power of the Inferno Gate network in comparison to other gated networks such as the popular LSTM.

**Developmental learning** – Because the data identity has been removed from the temporal information within the sequence, Inferno Gate can learn temporal patterns. The learning and error minimization processing of these rules may be similar to an inference process that attempts to capture the rules in one domain. As these temporal rules are abstracted, they may be applied to other domains as well or to other modalities. As a beneficial effect, learning higher-order hypothesis may accelerate learning massively in other domains, which is still an unachieved challenge by AI architectures [4], [69].

Moreover, since the PFC is processing a control on other cortical and sub-cortical structures, it may shape also their functional organization and bootstrapping.

**Transfert learning** – As the data identity is not encoded within the PFC neurons, we can expect that they are more robust to variability within the inputs such as distorted voices or voices with different tones (high and low pitch). In this line, we can expect Inferno Gate to find the same temporal patterns between different modalities, visual and auditory for

instance as during lips reading accentuating the pronounced sound with the visual input. This gating mechanism may permit to express another way to solve the binding problem across modalities and to perform transfert learning. For instance, some experiments done with babies found how they bind very early the tactile texture of a protruder and its visual shape although not seen before [98]. Other experiments showed such binding effect done between sound and vision in which babies correlated an ascending sound with a more intense light following the same temopral pattern.

## REFERENCES

- [1] E. Spelke, "What makes us smart? core knowledge and natural language," In: *Language in mind: Advances in the Investigation of language and thought*, ed. D. Gentner & S. Goldin-Meadow. MIT Press., p. 277311, 2003.
- [2] E. Spelke and K. Kinzler, "Core knowledge," *Developmental Science*, vol. 10, no. 1, p. 8996, 2007.
- [3] A. Gopnik, A. Meltzoff, and P. Kuhl, "The scientist in the crib what early learning tells us about the mind," *Publisher William Morrow Paperbacks*, 2000.
- [4] J. Tenenbaum, C. Kemp, T. Griffiths, and N. Goodman, "How to grow a mind statistics, structure, and abstraction," *Science*, vol. 331, no. 6022, pp. 1279–1285, 2011.
- [5] J. Saffran, R. Aslin, and E. Newport, "Statistical learning by 8-month-old infants," *Science*, vol. 274, p. 19261928, 1996.
- [6] J. Saffran and D. Wilson, "From syllables to syntax: Multilevel statistical learning by 12-month-old infants," *Infancy*, vol. 4, p. 273284, 2003.
- [7] A. Basirat, S. Dehaene, and G. Dehaene-Lambertz, "A hierarchy of cortical responses to sequence violations in three-month-old infants," *Cognition*, vol. 132, p. 137150, 2014.
- [8] A. Yu and P. Dayan, "Uncertainty, neuromodulation, and attention," *Neuron*, vol. 46, pp. 681–692, 2005.
- [9] N. Daw, Y. Niv, and P. Dayan, "Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control," *Nat. Neurosci.*, vol. 8, pp. 1704–1711, 2005.
- [10] N. Rougier and R. O'Reilly, "Learning representations in a gated prefrontal cortex model of dynamic task switching," *Cognitive Science*, vol. 26, p. 503520, 2002.
- [11] F. M. O'Reilly, RC, "Making working memory work: A computational model of learning in the prefrontal cortex and basal ganglia," *Neural Comput.*, vol. 18, no. 2, p. 283328, 2006.
- [12] R. Romo, C. Brody, A. Hernandez, and L. Lemus, "Neuronal correlates of parametric working memory in the prefrontal cortex," *Nature*, vol. 399, no. 6735, p. 470, 2018.
- [13] J. Tanji and E. Hoshi, "Behavioral planning in the prefrontal cortex," *Curr. Opin. Neurobiol.*, vol. 11, p. 164170, 2001.
- [14] J. X. Wang, Z. Kurth-Nelson, D. Kumaran, D. Tirumala, H. Soyer, J. Z. Leibo, D. Hassabis, and M. Botvinick, "Prefrontal cortex as a meta-reinforcement learning system," *Nature Neuroscience*, vol. 21, p. 860868, 2018.
- [15] E. Koechlin, "An evolutionary computational theory of prefrontal executive function in decision-making," *Phil. Trans. R. Soc. B*, vol. 369, p. 20130474, 2014.
- [16] —, "Prefrontal executive function and adaptive behavior in complex environments," *Current Opinion in Neurobiology*, vol. 37, pp. 1–6, 2016.
- [17] S. Dehaene, F. Meyniel, C. Wacongne, L. Wang, and C. Pallier, "The neural representation of sequences from transition probabilities to algebraic patterns and linguistic trees," *Neuron*, vol. 88, pp. 2–19, 2015.
- [18] M. Rouault and E. Koechlin, "Prefrontal function and cognitive control: from action to language," *Current Opinion in Behavioral Sciences*, vol. 21, p. 106111, 2018.
- [19] B. Averbeck, M. Chafee, D. Crowe, and G. A.P., "Neural activity in prefrontal cortex during copying geometrical shapes. i. single cells encode shape, sequence, and metric parameters," *Exp Brain Res.*, vol. 150, no. 2, pp. 127–41, 2003.
- [20] B. Averbeck, D. Crowe, M. Chafee, and G. A.P., "Neural activity in prefrontal cortex during copying geometrical shapes. ii. decoding shape segments from neural ensembles," *Exp Brain Res.*, vol. 150, no. 2, p. 143153, 2003.
- [21] K. Shima, M. Isoda, H. Mushiake, and J. Jun Tanji, "Categorization of behavioural sequences in the prefrontal cortex," *Nature*, vol. 445, pp. 315–318, 2007.
- [22] J. Tanji, K. Shima, and H. Mushiake, "Concept-based behavioral planning and the lateral prefrontal cortex," *Trends in Cognitive Sciences*, vol. 11, no. 12, pp. 528–534, 2007.
- [23] P. Barone and J. J.P., "Prefrontal cortex and spatial sequencing in macaque monkey," *Exp Brain Res*, vol. 78, p. 447 464, 2018.
- [24] M. Inoue and A. Mikami, "Prefrontal activity during serial probe reproduction task: encoding, mnemonic and retrieval processes," *J Neurophysiol*, vol. 95, p. 1008 1041, 2018.
- [25] R. Andersen and V. Mountcastle, "The influence of the angle of gaze upon the excitability of the light-sensitive neurons of the posterior parietal cortex," *J. Neuroscience*, vol. 3, pp. 532–548, 1983.
- [26] R. Andersen, G. Essick, and R. Siegel, "Encoding of spatial location by posterior parietal neurons," *Science*, vol. 230, pp. 450–458, 1985.
- [27] R. Andersen, "Multimodal integration for the representation of space in the posterior parietal cortex," *Philos Trans R Soc Lond B Biol Sci*, vol. 353, pp. 1421–1428, 1997.
- [28] A. Pouget and L. Snyder, "Computational approaches to sensorimotor transformations," *Nature Neuroscience*, vol. 3, pp. 1192–1198, 2000.
- [29] E. Hubbard, M. Piazza, P. Pinel, and S. Dehaene, "Interactions between number and space in parietal cortex," *Nat. Rev. Neurosci.*, vol. 6, p. 435438, 2005.
- [30] A. Pouget and L. Snyder, "Spatial transformations in the parietal cortex using basis functions," *J. of Cog. Neuro.*, vol. 3, pp. 1192–1198, 1997.
- [31] G. A.P., H. Merchant, T. Naselaris, and A. B., "Mapping of the preferred direction in the motor cortex," *Proc Natl Acad Sci USA*, vol. 104, no. 26, pp. 11 068–72, 2007.
- [32] S. Kakei, D. Hoffman, and P. Strick, "Sensorimotor transformations in cortical motor areas," *Neuroscience Research*, vol. 46, p. 110, 2003.
- [33] A. Genovesio, S. Wise, and R. Passingham, "Prefrontalparietal function: from foraging to foresight," *Trends in Cognitive Sciences*, vol. 18, no. 2, p. 7281, 2014.
- [34] A. e. a. Genovesio, "Feature- and order-based timing representations in the frontal cortex," *Neuron*, vol. 63, p. 254266, 2009.
- [35] M. Botvinick and T. Watanabe, "From numerosity to ordinal rank a gain-field model of serial order representation in cortical working memory," *The Journal of Neuroscience*, vol. 27, no. 32, pp. 8636–8642, 2007.
- [36] J. Fuster, "The prefrontal cortexan update time is of the essence," *Neuron*, vol. 30, pp. 319–333, 2001.
- [37] J. Paton and D. Buonomano, "The neural basis of timing: Distributed mechanisms for diverse functions," *Neuron*, vol. 98, no. 2, p. 687705, 2018.
- [38] A. Zylberberg, S. Dehaene, P. Roelfsema, and M. Sigman, "The human turing machine a neural framework for mental programs," *Trends in Cognitive Science*, vol. 15, no. 7, pp. 293–300, 2011.
- [39] C. Eliasmith, T. Stewart, X. Choo, T. Bekolay, T. DeWolf, Y. Tang, and D. Rasmussen, "A large-scale model of the functioning brain," *Science*, vol. 338, no. 6111, p. 12021205, 2012.
- [40] A. Zylberberg, L. Paz, P. Roelfsema, S. Dehaene, and M. Sigman, "A neuronal device for the control of multi-step computations," *Papers in Physics*, vol. 5, p. 050006, 2013.
- [41] C. Wacongne, E. Labyt, V. van Wassenhove, T. Bekinschtein, L. Nacache, , and S. Dehaene, "Evidence for a hierarchy of predictions and prediction errors in human cortex," *Proc. Natl. Acad. Sci. USA*, vol. 108, p. 2075420759, 2011.
- [42] N. Chomsky, "Three models for the description of language," *IEEE Trans. Inf. Theory*, vol. 2, p. 113124, 1956.
- [43] Y. Sugita and T. Tani, "Learning semantic combinatoriality from the interaction between linguistic and behavioral processes," *Adapt. Behav.*, vol. 13, no. 1, pp. 33–52, 2005.
- [44] T. Griffiths, N. Chater, C. Kemp, A. Perfors, and J. Tenenbaum, "Probabilistic models of cognition exploring representations and inductive biases," *Trends in Cognitive Sciences*, vol. 14, no. 8, pp. 357–364, 2010.
- [45] N. Chomsky, "Problems of projection," *Lingua*, vol. 130, p. 3349, 2013.

- [46] T. Tani, "Self-organization and compositionality in cognitive brains: A neurobotics study," *Proceedings of the IEEE*, vol. 102, no. 4, pp. 586–605, 2014.
- [47] B. Lake, C. Lee, and J. Tenenbaum, "One-shot learning of generative speech concepts," in *Proceedings of the 36th Annual Conference of the Cognitive Science Society*, 2014.
- [48] P. Dominey, "Complex sensory-motor sequence learning based on recurrent state representation and reinforcement learning," *Biol. Cybern.*, vol. 73, pp. 265–274, 1995.
- [49] P. Dominey, M. Arbib, and J. Joseph, "A model of corticostriatal plasticity for learning oculomotor associations and sequences," *J. Cogn. Neurosci.*, vol. 7, no. 25, 1995.
- [50] P. Dominey, "Recurrent temporal networks and language acquisition—from corticostriatal neurophysiology to reservoir computing," *Front. Psy.*, vol. 4, no. 500, pp. 1–14, 2013.
- [51] M. Botvinick and D. Plaut, "Short-term memory for serial order: A recurrent neural network model," *Psychological Review*, vol. 113, no. 2, p. 201233, 2006.
- [52] C. Wacongne, J. Changeux, and S. Dehaene, "A neuronal model of predictive coding accounting for the mismatch negativity," *J. Neurosci.*, vol. 32, p. 36653678, 2012.
- [53] A. Pitti, R. Braud, S. Mah, M. Quoy, and P. Gaussier, "Neural model for learning-to-learn of novel task sets in the motor domain," *Frontiers in Psychology*, vol. 4, no. 771, 2013.
- [54] B. J. Knowlton, R. G. Morrison, J. E. Hummel, and K. J. Holyoak, "A neurocomputational system for relational reasoning," *Trends in Cognitive Sciences*, vol. 16, no. M, pp. 373–381, 2012. [Online]. Available: <http://dx.doi.org/10.1016/j.tics.2012.06.002>
- [55] R. Laje and D. Buonomano, "Robust timing and motor patterns by taming chaos in recurrent neural networks," *Nature Neuroscience*, vol. 16, no. 7, pp. 925–935, 2013.
- [56] D. Sussillo, "Neural circuits as computational dynamical systems," *Current opinion in neurobiology*, vol. 25, p. 156163, 2014.
- [57] I. Tsuda, "Chaotic itinerancy and its roles in cognitive neurodynamics," *Current Opinion in Neurobiology*, vol. 31, p. 6771, 2015.
- [58] F. Mannella, M. Mirolli, and G. Baldassare, "Goal-directed behavior and instrumental devaluation: A neural system-level computational model," *Front. Behav. Neurosci.*, vol. 10, no. 181, pp. 575–595, 2016.
- [59] X. Hinaut, M. Petit, G. Pointeau, and P. Dominey, "Real-time parallel processing of grammatical structure in the fronto-striatal system: A recurrent network simulation study using reservoir computing," *PLoS ONE*, vol. 8, no. 2, p. e52946, 2013.
- [60] —, "Exploring the acquisition and production of grammatical constructions through human-robot interaction with echo state networks," *Front. Neurobot.*, vol. 8, no. 16, 2014.
- [61] F. Mannella and G. Baldassare, "Selection of cortical dynamics for motor behaviour by the basal ganglia," *Biological Cybernetics*, vol. 109, pp. 575–595, 2015.
- [62] Y. Yamashita and T. Tani, "Emergence of functional hierarchy in a multiple timescale neural network model: a humanoid robot experiment," *PLoS Comput. Biol.*, vol. 4, no. 11, p. e1000220, 2008.
- [63] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, p. 17351780, 1997.
- [64] F. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with lstm," *Neural Comput.*, vol. 12, no. 10, p. 24512471, 2000.
- [65] M. Frank, B. Loughry, and R. O'Reilly, "Interactions between the frontal cortex and basal ganglia in working memory: a computational model," *Cogn. Affect. Behav. Neurosci.*, vol. 1, pp. 137–160, 2001.
- [66] C. Eliasmith, *How to build a brain: A neural architecture for biological cognition*. New York, NY: Oxford University Press, 2013.
- [67] M. Hasselmo and C. Stern, "A network model of behavioural performance in a rule learning task," *Phil. Trans. R. Soc. B*, vol. 373, p. 20170275, 2018.
- [68] T. Kriete, D. C. Noelle, J. D. Cohen, and R. C. O'Reilly, "Indirection and symbol-like processing in the prefrontal cortex and basal ganglia," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 110, no. 41, pp. 16390–5, 2013.
- [69] B. M. Lake, T. D. Ullman, J. B. Tenenbaum, and S. J. Gershman, "Building machines that learn and think like people," *Behavioral and Brain Sciences*, vol. 40, p. e253, 2017.
- [70] A. Pitti, P. Gaussier, and M. Quoy, "Iterative free-energy optimization for recurrent neural networks (inferno)," *PLoS ONE*, vol. 12, no. 3, p. e0173684, 2017.
- [71] K. Friston, "Learning and inference in the brain," *Neural Networks*, vol. 16, no. 9, p. 13251352, 2003.
- [72] K. Friston, J. Kilner, and L. Harrison, "A free energy principle for the brain," *Journal of Physiology-Paris*, vol. 100, no. 1–3, p. 7087, 2006.
- [73] G. Keller and T. M. Flogel, "Predictive processing: A canonical cortical computation," *Neuron*, vol. 100, no. 2, p. 424435, 2018.
- [74] M. Botvinick, Y. Niv, and A. Barto, "Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective," *Cognition*, vol. 113, no. 3, p. 262280, 2009.
- [75] E. Koechlin, C. Ody, and F. Kounieilher, "The architecture of cognitive control in the human prefrontal cortex," *Science*, vol. 302, no. 14, pp. 1181–1185, 2003.
- [76] T. Buschman and E. Miller, "Goal-direction and top-down control," *Phil. Trans. R. Soc. B*, vol. 369, p. 20130471, 2014.
- [77] E. Miller, "The 'working' of working memory," *Dialogues Clin. Neurosci.*, vol. 15, no. 4, pp. 411–418, 2015.
- [78] W. Ma, J. Beck, P. Latham, and A. Pouget, "Bayesian inference with probabilistic population codes," *Nat. Neurosci.*, vol. 9, no. 11, p. 14321438, 2006.
- [79] M. Botvinick and A. Weinstein, "Model-based hierarchical reinforcement learning and human action control," *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, vol. 369, p. 20130480, 2014.
- [80] S. Thorpe, A. Delorme, and R. Van Rullen, "Spike-based strategies for rapid processing," *Neural Networks*, vol. 14, pp. 715–725, 2001.
- [81] A. Pitti, R. Braud, S. Mah, M. Quoy, and P. Gaussier, "Neural model for learning-to-learn of novel task sets in the motor domain," *Frontiers in Psychology*, vol. 4, no. 771, 2013.
- [82] E. Izhikevich, "Polychronization computation with spikes," *Neural Computation*, vol. 18, pp. 245–282, 2006.
- [83] S. Mahe, P. Braud, R. Gaussier, M. Quoy, and A. Pitti, "Exploiting the gain-modulation mechanism in parieto-motor neurons application to visuomotor transformations and embodied simulation," *Neural Networks*, vol. 62, pp. 102–111, 2015.
- [84] J. Abrossimoff, A. Pitti, and P. Gaussier, "Visual learning for reaching and body-schema with gain-field networks," *Joint IEEE International Conferences on Development and Learning and Epigenetic Robotics (ICDL-Epirob)*, pp. 191–197, 2018.
- [85] D. Jin, N. Fujii, and A. Graybiel, "Neural representation of time in cortico-basal ganglia circuits," *Proc. Natl. Acad. Sci. USA*, vol. 106, p. 1915619161, 2009.
- [86] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biological Cybernetics*, vol. 43, pp. 59–69, 1982.
- [87] S. Funahashi, "Prefrontal cortex and working memory processes," *Neuroscience*, vol. 139, no. 1, pp. 251–261, 2006.
- [88] E. Wood, P. Dudchenko, R. J. Robitsek, and H. Eichenbaum, "Hippocampal neurons encode information about different types of memory episodes occurring in the same location," *Neuron*, vol. 27, no. 3, pp. 623–633, 2000.
- [89] L. Frank, B. E.N., and M. Wilson, "Trajectory encoding in the hippocampus and entorhinal cortex," *Neuron*, vol. 27, no. 1, pp. 169 – 178, 2000.
- [90] M. Rigotti, O. Barak, M. Warden, X. Wang, N. Daw, E. Miller, and S. Fusi, "The importance of mixed selectivity in complex cognitive tasks," *Nature*, vol. 497, no. 7451, pp. 585–590, 2013.
- [91] E. Salinas and P. Thier, "Gain modulation a major computational principle of the central nervous system," *Neuron*, vol. 27, pp. 15–21, 2000.
- [92] G. Blohm and J. Crawford, "Fields of gain in the brain," *Neuron*, vol. 64, pp. 598–600, 2009.
- [93] A. Pitti, A. Blanchard, M. Cardinaux, and P. Gaussier, "Gain-field modulation mechanism in multimodal networks for spatial perception," *12th IEEE-RAS International Conference on Humanoid Robots Nov.29-Dec.1, 2012. Business Innovation Center Osaka, Japan*, pp. 297–302, 2012.
- [94] G. Pugach, A. Pitti, O. Tolochko, and P. Gaussier, "Brain-inspired coding of robot body schema through visuo-motor integration of touched events," *Frontiers in NeuroRobotics*, p. accepted, 2019.
- [95] S. Fujisawa, A. Amarasingham, M. Harrison, and G. Buzsaki, "Behavior-dependent short-term assembly dynamics in the medial prefrontal cortex," *Nature Neuroscience*, vol. 11, no. 7, pp. 823–33, 2008.
- [96] J. Ainge, M. van der Meer, R. Langston, and W. E.R., "Exploring the role of context-dependent hippocampal activity in spatial alternation behaviour," *Hippocampus*, vol. 17, p. 9881002, 2007.
- [97] P. Smolensky, "Tensor product variable binding and the representation of symbolic structures in connectionist systems," *Artif. Intell.*, vol. 46, p. 159216, 1990.

[98] A. Meltzoff and R. Borton, "Intermodal matching by human neonates," *Nature*, vol. 282, pp. 403–404, 1979.

## IPFC Analysis

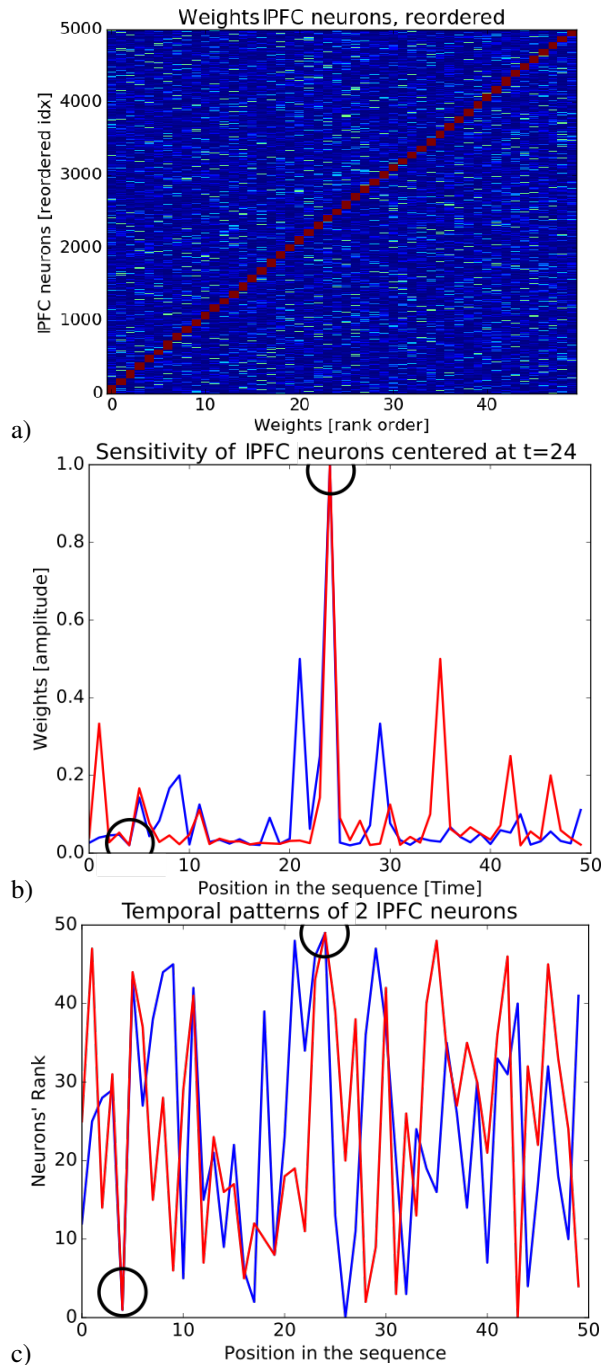


Fig. 5. IPFC coding analysis. In a), raster plot of the weights of the IPFC units (Y axis) reordered with respect to their sensitivity to the location within the sequence (X axis). Each neuron within the IPFC network has learned a weights distribution centered at one position within the sequence; eg the beginning, middle or end of the sequence. Neurons that code well the beginning of the sequence can help to predict the rest of the sequence (prospective memory). Neurons that code well the end of the sequence are salient to the elements during the whole sequence (retrospective memory). In b), weights distribution for two IPFC units centered at location #24 within the sequence. The circles indicate the locations where the weights value overlap. In c), as the weight index indicates the location within the sequence and the weight value indicates the neuron's rank at that position, it is possible to reconstruct the temporal pattern for which the IPFC neurons are the most sensitive to. The circles indicate weights with same index and same value. The two temporal patterns crosses at these locations.



# Reconstructed OFC Dynamics for two dlPFC units

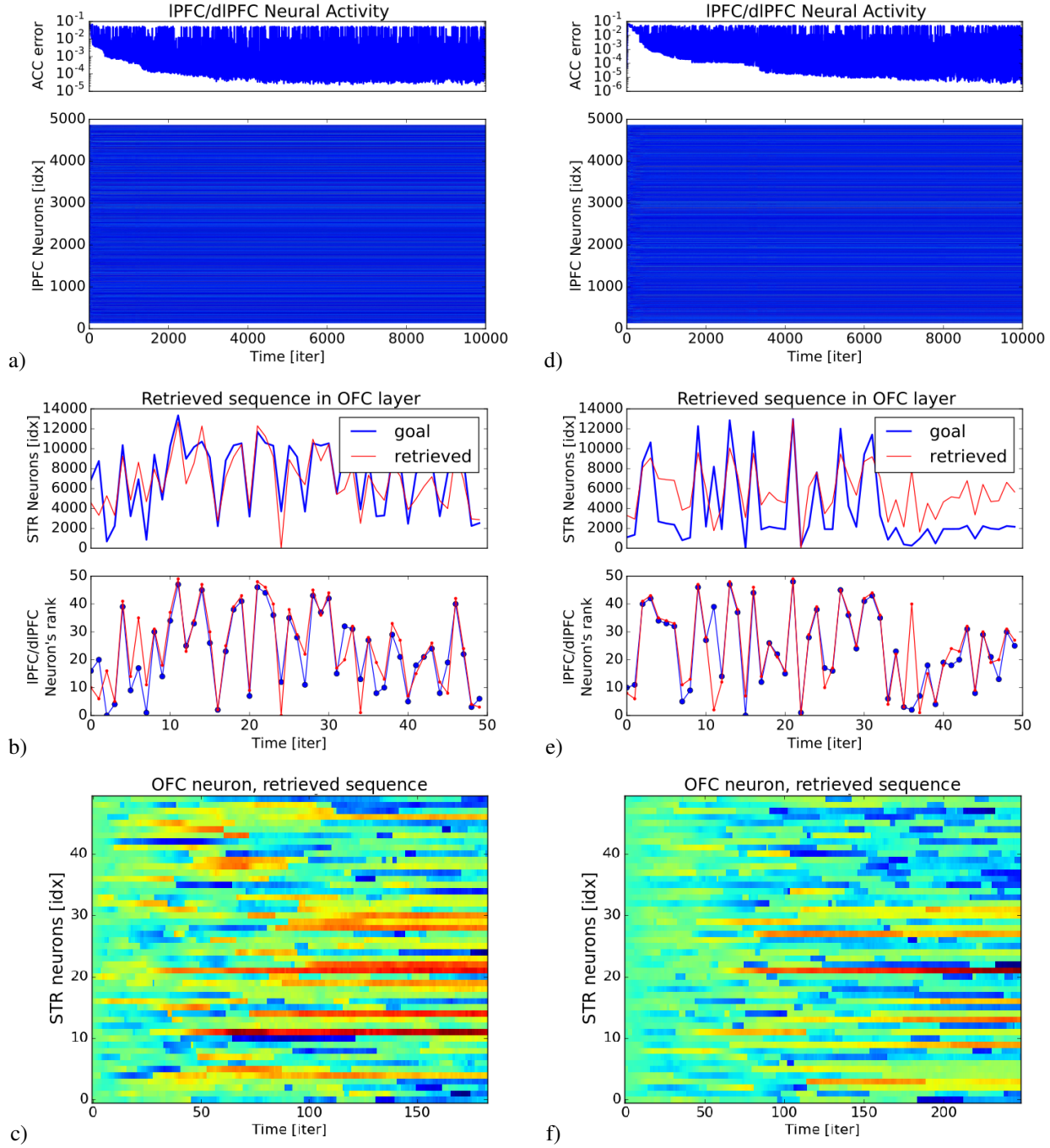


Fig. 6. Free-energy optimization for retrieving sequence in OFC layer. In a), error minimization in ACC unit (top chart) and optimization of dlPFC activity through exploration in OFC and observed in IPFC layer (bottom chart). In b), the final sequence retrieved in red in the OFC layer with neurons identity of the STR layer between [0, 14.000] with respect to a goal sequence in blue (top chart). The temporal pattern retrieved in IPFC/dIPFC layers and the temporal pattern of the goal sequence in blue are displayed in the bottom chart. The OFC layer achieves to retrieve with some uncertainty a temporal sequence but the global structure of the sequence and the rank orders are mostly respected.

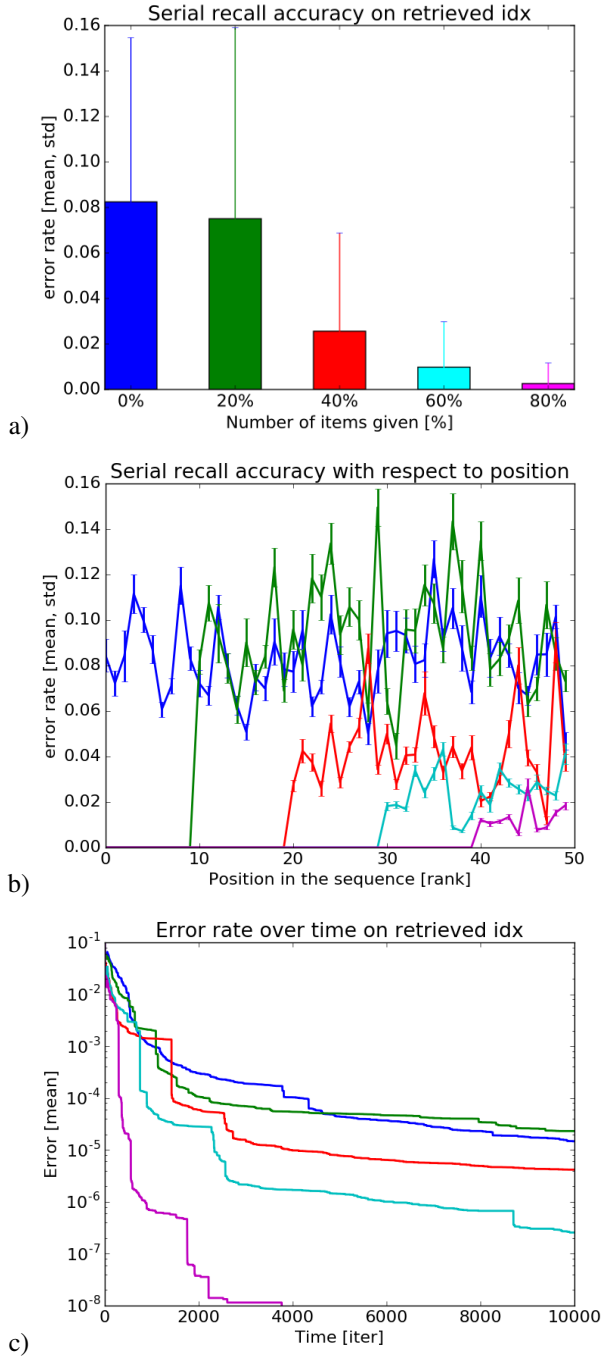


Fig. 7. Performance analysis of the Inferno gate architecture for retrieving sequences with respect to the amount of information given. In a), the error rate computed from goal sequence encoded in IPFC/dIPFC and retrieved sequences in the OFC layer with respect to the amount of items given from 0 to 80% of the sequence given at the B45 level. The more the number of units to search are few, the more accurate is the recall. If we provide 40% of the items of the sequence we want to retrieve, the error on the neurons id is particularly small and almost error free if 80% of the neurons are given. In b), distribution of the error rate with respect to the position within the sequence and the amount of information provided. The error done by the network to retrieve the sequences is not related to the temporal position within the sequence but to the amount of information furnished to it. In c), convergence rate of the network with respect to the amount of information provided. The explorative search for retrieving the sequences follow a power-law curve with respect to the amount of information furnished.



# Serial Recall in OFC from incomplete information

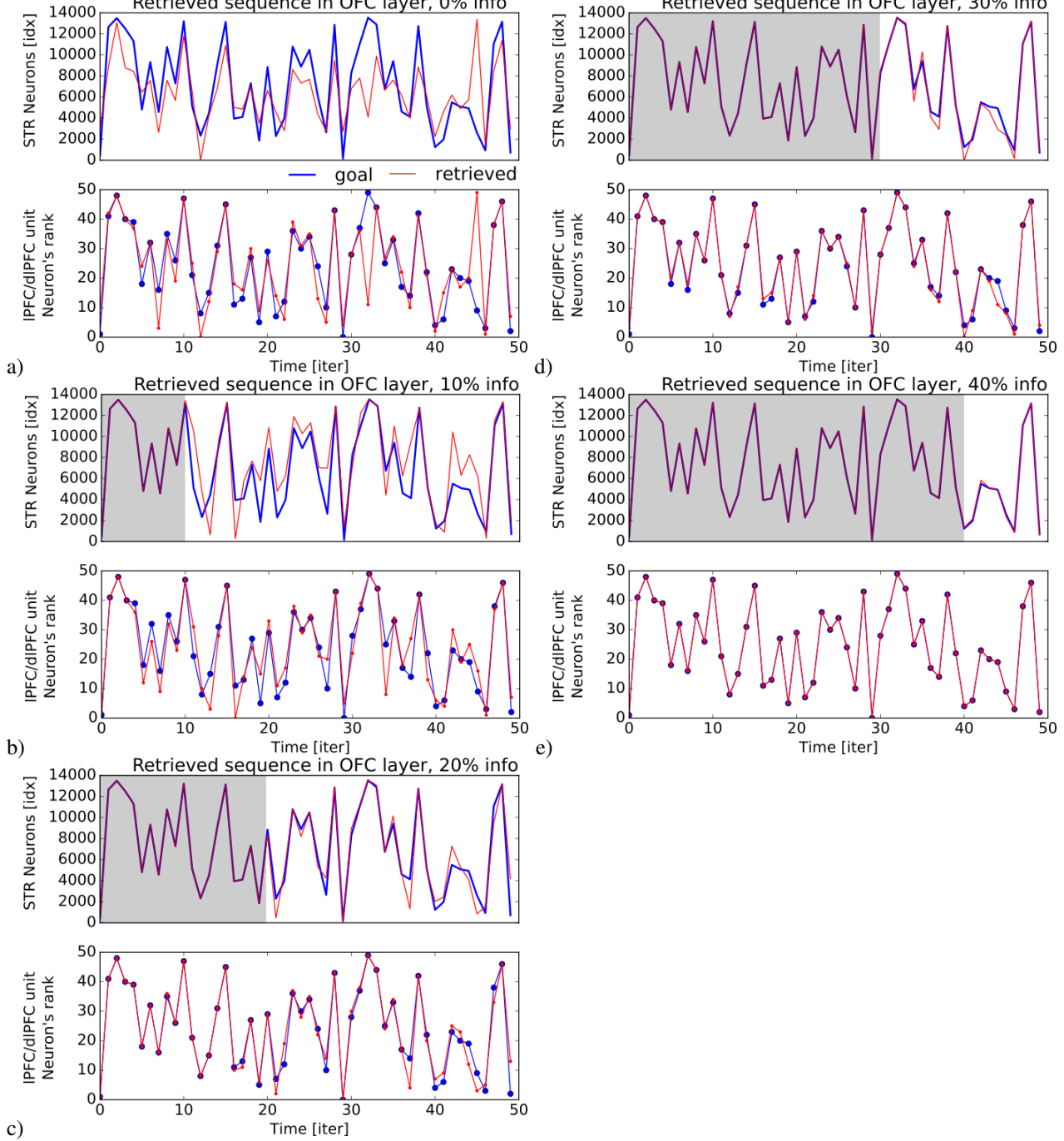


Fig. 8. Serial recall in OFC layer from incomplete information. Retrieved goal sequence when 0%, 20%, 40%, 60%, 80% information are furnished to the system, respectively a-e). In the top charts, the generated sequences in OFC layer with identity for the STR neurons are displayed in red with the goal sequences to retrieve in blue. The more information is given to the system, the easier is the explorative search to retrieve the missing units identity. In the bottom chart, although the rank order in the temporal patterns of the units in IPFC is respected, this does not warranty that the units identity is retrieved correctly in the OFC sequences.