Expressive Power of Non-deterministic Evolving Recurrent Neural Networks in Terms of Their Attractor Dynamics

Jérémie Cabessa[1(✉)] and Jacques Duparc[2]

[1] Laboratory of Mathematical Economics,
University of Paris 2 – Panthéon-Assas, 4 Rue Blaise Desgoffe,
75006 Paris, France
jeremie.cabessa@u-paris2.fr
[2] Department of Information Systems, Faculty of Business and Economics,
University of Lausanne, CH-1015 Lausanne, Switzerland

Abstract. We introduce a model of nondeterministic hybrid recurrent neural networks – made up of Boolean input and output cells as well as internal sigmoid neurons, and equipped with the possibility to have their synaptic weights evolve over time, in a nondeterministic manner. When subjected to some infinite input stream and some specific synaptic evolution, the networks necessarily exhibit some attractor dynamics in their Boolean output cells, and accordingly, recognize some specific neural ω -languages. The expressive power of these networks is measured via the topological complexity of their underlying neural ω-languages. In this context, we prove that the two models of rational-weighted and real-weighted nondeterministic hybrid neural networks are computationally equivalent, and recognize precisely the set of all analytic neural ω-languages. They are therefore strictly more expressive than the nondeterministic Büchi and Muller Turing machines.

Keywords: Recurrent neural networks · Neural computation · Analog computation · Evolving systems · Attractors · Turing machines · Expressive power

1 Introduction

The understanding of the computational and dynamical capabilities of brain-like models of computation represents an issue of central importance. In this context, much attention has been focused on comparing the computational powers of various neural models to those of diverse abstract machines, see for instance [2,4, 13–16,18–20,23]. As a consequence, the computational power of neural networks has been shown to be intimately related to the nature of their synaptic weights and activation functions, and able to range from finite state automata [13–15] up to super-Turing capabilities [2,4,18–20].

Following this global line of thought, the first author initiated the study of the expressive power of recurrent neural networks from the perspective of

C.S. Calude and M.J. Dinneen (Eds.): UCNC 2015, LNCS 9252, pp. 144–156, 2015.
DOI: 10.1007/978-3-319-21819-9_10

their attractor dynamics [7,10]. This approach is motivated by the fact that, in their model, the attractor dynamics of the neural networks are the precise phenomena that underly the arising of spatiotemporal patterns of discharges – a feature considered to be significantly involved in the processing and coding of information in the brain [24,25].

In this context, they proved that Boolean recurrent neural networks provided with some assignment of their attractors into two different kinds are computationally equivalent to Muller automata, and hence recognize precisely the so-called ω-regular neural languages. Consequently, the most refined topological classification of ω-languages [26] can be transposed from the automaton to the neural network context, and yields to some transfinite hierarchical classification of Boolean neural networks according to their attractor dynamics [6], which in turn represents a new attractor-based complexity measurement for Boolean recurrent neural networks [10].

More recently, they considered a model of so-called *hybrid recurrent neural networks* composed with Boolean input and output cells as well as internal sigmoid neurons. They showed that the rational and real-weighted hybrid neural networks are computationally equivalent to and strictly more powerful than deterministic Muller Turing machines, respectively [5]. Furthermore, the evolving hybrid neural nets are equivalent to the real-weighted ones, irrespective of whether their synaptic weights are modelled by rational or real numbers [5]. These results provide a generalization to this specific computational context of those obtained for the cases of classical [2,4] and interactive computation [1,3,9,11].

Here, we provide the nondeterministic counterpart of these results. We consider a model of *nondeterministic hybrid recurrent neural networks*, which consist of hybrid neural nets equipped with the possibility to have their synaptic weights evolve over time – in a nondeterministic manner. When subjected to some infinite input stream as well as to some specific evolution of their synaptic weights, the networks necessarily exhibit some attractor dynamics in their Boolean output cells, which is assumed to be of two possible kinds, either *meaningful* or *spurious*. The neural ω-language of a network corresponds to the set of all input streams which induce a meaningful attractor dynamics, for some possible evolution of its synaptic weights. The expressive power of the networks is then measured via the topological complexity of their underlying neural ω-languages. In this context, we prove that the two models of rational-weighted and real-weighted nondeterministic hybrid neural networks are computationally equivalent, and recognize precisely the set of all analytic neural ω-languages. They are therefore strictly more expressive than the nondeterministic Büchi and Muller Turing machines. These results are discussed in the last section.

2 Preliminaries

A *topological space* is a pair $(S, \mathcal{T})$ where S is a set and $\mathcal{T}$ is a collection of subsets of S such that $\emptyset \in \mathcal{T}$, $S \in \mathcal{T}$, and $\mathcal{T}$ is closed under arbitrary unions and finite intersections. The collection $\mathcal{T}$ is called a *topology* on S, and its members are called *open sets*. Given some topological space $(S, \mathcal{T})$, the class *Borel subsets* of S, denoted by $\mathbf{\Delta^1_1}$, consists of the smallest collection of subsets of S containing all open sets and closed under countable union and complementation. For every ordinal α, one defines by transfinite induction the following *Borel classes*:

- $\mathbf{\Sigma^0_1} = \{X \subseteq S : X \text{ is open}\}$,
- $\mathbf{\Pi^0_\alpha} = \{X \subseteq S : X^{\complement} \in \mathbf{\Sigma^0_\alpha}\}$,
- $\mathbf{\Sigma^0_\alpha} = \{X \subseteq S : X = \bigcup_{n \geq 0} X_n,\ X_n \in \mathbf{\Pi^0_{\alpha_n}},\ \alpha_n < \alpha,\ n \in \mathbb{N}\}$, for $\alpha > 1$,
- $\mathbf{\Delta^0_\alpha} = \mathbf{\Sigma^0_\alpha} \cap \mathbf{\Pi^0_\alpha}$.

The collection of all classes $\mathbf{\Sigma^0_\alpha}$, $\mathbf{\Pi^0_\alpha}$, and $\mathbf{\Delta^0_\alpha}$ provides a stratification of the whole class of Borel sets known as *the Borel hierarchy*. The *rank* of a Borel set $X \subseteq S$ is the smallest ordinal α such that $X \in \mathbf{\Sigma^0_\alpha} \cup \mathbf{\Pi^0_\alpha} \cup \mathbf{\Delta^0_\alpha}$, and represents the minimal number of complementation and countable union operations that are needed in order to obtain X from an initial collection of open sets. It is commonly considered as a relevant measure of the topological complexity of Borel sets.

Besides, given any set A, we let A^*, A^+ and A^ω denote respectively the sets of finite sequences, non-empty finite sequences and infinite sequences of elements of A. For any $x \in A^* \cup A^\omega$, the *length* of x is denoted by $|x|$, the $(i+1)$-th element of x will be denoted by $x(i)$ for any $0 \leq i < |x|$, and the subsequence of the n-th first elements of x is denoted by $x[0{:}n]$, with the convention that $x[0{:}0] = \lambda$, the empty sequence. Hence, any $x \in A^+$ and $y \in A^\omega$ can be written as $x = x(0)x(1)\cdots x(|x|-1)$ and $y = y(0)y(1)y(2)\cdots$, respectively. The fact that x is a *prefix* (resp. *strict prefix*) of y will be denoted by $x \subseteq y$ (resp. $x \subsetneq y$). The concatenation of x and y is denoted $x \cdot y$, and for any $X \subseteq A^*$ and $Y \subseteq A^* \cup A^\omega$, one sets $X \cdot Y = \{z \in A^* \cup A^\omega : z = x \cdot y \text{ for some } x \in X \text{ and } y \in Y\}$. A set of the form $\{x\} \cdot A^\omega$ is generally denoted $x \cdot A^\omega$. Finally, a sequence of $A^* \cup A^\omega$ will also be called a *word*, and a subset of A^ω is generally called an *ω-language*.

In the sequel, the spaces of N-dimensional Boolean, rational and real vectors will be denoted by $\mathbb{B}^N$, $\mathbb{Q}^N$ and $\mathbb{R}^N$, respectively. The space $(\mathbb{B}^N)^\omega$ is naturally assumed to be equipped with the product topology of the discrete topology on $\mathbb{B}^N$. Accordingly, the basic open sets are of the form $p \cdot (\mathbb{B}^N)^\omega$, for some $p \in (\mathbb{B}^N)^*$. The general open sets are countable unions of basic open sets. This space is Polish (i.e., separable and completely metrizable) [12]. The spaces $(\mathbb{Q}^N)^\omega$ and $(\mathbb{R}^N)^\omega$ are assumed to be equipped with the product topologies of the usual topologies on $\mathbb{Q}^N$ and $\mathbb{R}^N$, respectively. Accordingly, the basic open sets are of the form $X_0 \cdot \ldots \cdot X_n \cdot (\mathbb{Q}^N)^\omega$ or $X_0 \cdot \ldots \cdot X_n \cdot (\mathbb{R}^N)^\omega$, for some $n \geq 0$, where each X_i is an open set of $\mathbb{Q}^N$ or $\mathbb{R}^N$ for their usual topologies, respectively. The general open sets are arbitrary unions of basic open sets. These two spaces are also Polish [12].

An ω-language $L \subseteq (\mathbb{B}^N)^\omega$ is *analytic* iff there exists some $\mathbf{\Pi^0_2}$-set $X \subseteq (\mathbb{B}^N)^\omega \times \{0,1\}^\omega$ such that $L = \pi_1(X) = \{s \in (\mathbb{B}^N)^\omega : \exists\, e \in \{0,1\}^\omega \text{ s.t. } (s,e) \in X\}$ [12, Exercise14.3]. This fact will be used in forthcoming Proposition 1. Equivalently, $L \subseteq (\mathbb{B}^N)^\omega$ is *analytic* iff there exists some Polish space E and some Borel set $X \subseteq (\mathbb{B}^N)^\omega \times E$ such that $L = \pi_1(X)$ [12, Exercise14.3]. This fact will be used in forthcoming Proposition 2. The class of analytic sets, denoted by $\mathbf{\Sigma^1_1}$, strictly contains that of Borel sets, namely $\mathbf{\Delta^1_1} \subsetneq \mathbf{\Sigma^1_1}$ [12, Theorem 14.2].

3 Büchi and Muller Turing Machines

The study of the behavior of reactive systems has led to the emergence of a theory of automata working on infinite objects [17,22]. In this context, a *Büchi* (resp. a *Muller*) *Turing machine* can be defined as a pair $(\mathcal{M}, \mathcal{F})$ (resp. a pair $(\mathcal{M}', \mathcal{T})$), where $\mathcal{M}$ (resp. $\mathcal{M}'$) is a classical Turing machine and $\mathcal{F}$ is a subset of the states of $\mathcal{M}$ (resp. $\mathcal{T}$ is a collection of subsets of the states of $\mathcal{M}'$). In the case of $\mathcal{M}$ (resp. $\mathcal{M}'$) being deterministic, an infinite input stream s is said to be recognized by $\mathcal{M}$ (resp. by $\mathcal{M}'$) if the set of states visited infinitely often by $\mathcal{M}$ (resp. by $\mathcal{M}'$) during the processing of s intersects the set $\mathcal{F}$ (resp. belongs to the collection $\mathcal{T}$). In the non-deterministic case, s is said to be *recognized* by each such machine if there exists a computational path which satisfies the required condition. The ω-language associated with each such machine consists of the set of all words that it recognizes.

The deterministic Büchi Turing machines are strictly less powerful than the deterministic Muller ones. Indeed, every ω-language recognized by some deterministic Büchi Turing machine belongs to the topological class $\mathbf{\Pi^0_2}$, whereas the ones recognized by Muller Turing machine belong to the topological class $BC(\mathbf{\Pi^0_2})$, i.e., the finite Boolean combinations of $\mathbf{\Pi^0_2}$-sets [21, Corollaries 3.3 and 3.4]. Moreover, one can easily show the existence of infinitely many ω-languages which are recognizable by some Muller Turing machines but by no Büchi Turing machine. In the non-deterministic case, Büchi and Muller Turing machines are computationally equivalent. They recognize precisely the class of *effectively analytic* ω-languages [21, Theorem 3.5].

The class of effectively analytic sets is usually denoted by Σ^1_1 (lightface), and for the sequel, we recall that the relation $\Sigma^1_1 \subsetneq \mathbf{\Sigma^1_1}$ trivially holds [12].

4 The Model

We introduce a model of so-called *hybrid evolving recurrent neural network*. The term *hybrid* refers to the fact that the network involves both Boolean and sigmoid cells. The term *evolving* refers to the fact that the synaptic weights are able to evolve over time. The expressive power of the networks will be related to the attractor dynamics of their Boolean output cells.

A *hybrid (or Boolean/sigmoid) evolving recurrent neural network* (denoted by Ev-RNN) consists of a synchronous network of neurons related together in a general architecture. The network contains N internal sigmoid neurons $(x_i)_{i=1}^N$, M Boolean input cells $(u_i)_{i=1}^M$, and P Boolean output cells $(y_i)_{i=1}^P$. The dynamics of the network is computed as follows: given the activation values of the internal and input neurons $(x_j)_{j=1}^N$ and $(u_j)_{j=1}^M$ at time t, the activation values of each internal neuron x_i and each output neuron y_i at time $t+1$ are updated by the following equations, respectively:

$$x_i(t+1) = \sigma\left(\sum_{j=1}^{N} a_{ij}(t) \cdot x_j(t) + \sum_{j=1}^{M} b_{ij}(t) \cdot u_j(t) + c_i(t)\right) \text{ for } i = 1, \ldots, N \quad (1)$$

$$y_i(t+1) = \theta\left(\sum_{j=1}^{N} a_{ij}(t)\cdot x_j(t) + \sum_{j=1}^{M} b_{ij}(t)\cdot u_j(t) + c_i(t)\right) \text{ for } i = 1,\ldots,P \quad (2)$$

Here, $a_{ij}(t)$, $b_{ij}(t)$, and $c_i(t)$ are time dependent values describing the evolving weighted synaptic connections and weighted bias of the network, and σ and θ are the classical sigmoid-linear and hard-threshold activation functions respectively defined by:

$$\sigma(x) = \begin{cases} 0 & \text{if } x < 0, \\ x & \text{if } 0 \le x \le 1, \\ 1 & \text{if } x > 1 \end{cases} \quad \text{and} \quad \theta(x) = \begin{cases} 0 & \text{if } x < 1, \\ 1 & \text{if } x \ge 1. \end{cases}$$

We further assume that the synaptic weights $a_{ij}(t)$, $b_{ij}(t)$, $c_i(t)$ might evolve between two designated bounds S and S' imposed by the biological constitution of the synapses.

Throughout this paper, two models of Ev-RNNs are considered according to the nature of their synaptic weights. In fact, an Ev-RNN will be called *rational* (denoted by Ev-RNN[$\mathbb{Q}$]) or *real* (denoted by Ev-RNN[$\mathbb{R}$]) if its synaptic weights $a_{ij}(t), b_{ij}(t), c_i(t)$ are modelled by rational or real numbers at any time step t, respectively. Note that any Ev-RNN[$\mathbb{Q}$] is also an Ev-RNN[$\mathbb{R}$] by definition.

Let $\mathcal{N}$ be some Ev-RNN $\mathcal{N}$. For each time step $t \ge 0$, the Boolean vector

$$\boldsymbol{u}(t) = (u_1(t), \ldots, u_M(t)) \in \mathbb{B}^M$$

describing the activation values of the M input units of $\mathcal{N}$ at time t is the *input* submitted to $\mathcal{N}$ at time t. The pair

$$\langle \boldsymbol{x}(t), \boldsymbol{y}(t)\rangle \in [0,1]^N \times \mathbb{B}^P$$

describing the activation values of the internal and output cells at time t is the *state* of $\mathcal{N}$ at time t. The second element of this pair, namely $\boldsymbol{y}(t)$, is the *Boolean state* of $\mathcal{N}$ at time t.

Assuming the initial state of the network to be $\langle \boldsymbol{x}(0), \boldsymbol{y}(0)\rangle = \langle \boldsymbol{0}, \boldsymbol{0}\rangle$, any infinite input stream

$$s = (\boldsymbol{u}(t))_{t\in\mathbb{N}} = \boldsymbol{u}(0)\boldsymbol{u}(1)\boldsymbol{u}(2)\cdots \in (\mathbb{B}^M)^\omega$$

induces via Eqs. (1) and (2) an infinite sequence of consecutive states

$$c_s = (\langle \boldsymbol{x}(t), \boldsymbol{y}(t)\rangle)_{t\in\mathbb{N}} = \langle \boldsymbol{x}(0), \boldsymbol{y}(0)\rangle\langle \boldsymbol{x}(1), \boldsymbol{y}(1)\rangle\cdots \in ([0,1]^N \times \mathbb{B}^P)^\omega$$

called the *computation* of $\mathcal{N}$ induced by s. The corresponding infinite sequence of Boolean states

$$c'_s = (\boldsymbol{y}(t))_{t\in\mathbb{N}} = \boldsymbol{y}(0)\boldsymbol{y}(1)\boldsymbol{y}(2)\cdots \in (\mathbb{B}^P)^\omega$$

is the *Boolean computation* of $\mathcal{N}$ induced by s.

Note that any Ev-RNN $\mathcal{N}$ (with P Boolean output cells) can only have 2^P – i.e., finitely many – possible distinct Boolean states. Consequently, for any infinite Boolean computation c'_s, there necessarily exists at least one Boolean state that recurs infinitely often in c'_s. In fact, any Boolean computation c'_s necessarily consists of a finite prefix of Boolean states followed by an infinite suffix of Boolean states that repeat infinitely often – yet not necessarily in a periodic manner. The non-empty set of all the Boolean states that repeat infinitely often in c'_s will be denoted by $inf(c'_s)$. According to these considerations, a set of states of the form $inf(c'_s)$ for some computation c'_s will be called an *attractor* for $\mathcal{N}$. A precise definition can be given as follows [10]:

Definition 1. Let $\mathcal{N}$ be some Ev-RNN. A set $A = \{\boldsymbol{y_0}, \ldots, \boldsymbol{y_k}\} \subseteq \mathbb{B}^P$ is an *attractor* for $\mathcal{N}$ if there exists some infinite input stream s such that the corresponding Boolean computation c'_s satisfies $inf(c'_s) = A$.

In words, an attractor of $\mathcal{N}$ is a set of Boolean states into which the computation of the network could become forever trapped – yet not necessarily in a periodic manner –, for some infinite input stream s.

In this work, we suppose that attractors can be of two distinct types, namely either *meaningful* or *spurious*. The type of each attractor could be determined by its neurophysiological significance with respect to measurable observations, e.g. associated with certain behaviors or sensory discriminations. The classification of these attractors into meaningful or spurious types is not the subject of this paper. Hence, from this point onwards, we assume any Ev-RNN to be equipped with a corresponding classification of all of its attractors into meaningful and spurious types.

According to these considerations, an infinite input stream $s \in (\mathbb{B}^M)^\omega$ of $\mathcal{N}$ is called *meaningful* if $inf(c'_s)$ is a meaningful attractor, and it is called *spurious* if $inf(c'_s)$ is a spurious attractor. The set of all meaningful input streams of $\mathcal{N}$ is called the *neural ω-language* of $\mathcal{N}$ and is denoted by $L(\mathcal{N})$. An arbitrary set of input streams $L \subseteq (\mathbb{B}^M)^\omega$ is said to be *recognizable* by some Ev-RNN if there exists a network $\mathcal{N}$ such that $L(\mathcal{N}) = L$.

We now introduce a natural notion of a nondeterministic Ev-RNN, where the nondeterminism is expressed as a set of possible infinite evolving patterns of the synaptic weights. At the beginning of a computation, the network chooses one such possible evolution in a nondeterministic manner, and sticks to it throughout its whole computational process.

A *nondeterministic* Ev-RNN consists of a pair $(\mathcal{N}, E)$, where $\mathcal{N}$ is an Ev-RNN with K evolving synaptic connections, and $E \subseteq ([S, S']^K)^\omega$ is a set of infinite sequences of K-dimensional vectors – describing the possible infinite evolutions for the K synaptic connections of $\mathcal{N}$. Every element e of E is called a possible *evolution* for $\mathcal{N}$, and if the evolution $e = \boldsymbol{e}(0)\boldsymbol{e}(1)\boldsymbol{e}(2)\cdots \in E$ is followed by $\mathcal{N}$, each vector $\boldsymbol{e}(t)$ describes the values of the K synaptic weights of $\mathcal{N}$ at time step t.[1] In this context, the Boolean computation produced by $(\mathcal{N}, E)$ when it

[1] By contrast, a deterministic Ev-RNN has only one possible evolution for its synaptic weights, and hence corresponds to a nondeterministic Ev-RNN where the set E is reduced to a singleton.

receives the input stream $s \in (\mathbb{B}^M)^\omega$ and follows the evolution $e \in E$ is denoted by $c'_{(s,e)}$.

According to these considerations, a *nondeterministic* Ev-RNN[$\mathbb{Q}$] is a pair $(\mathcal{N}, E)$ such that $E \subseteq ((\mathbb{Q} \cap [S, S'])^K)^\omega$, and a *nondeterministic* Ev-RNN[$\mathbb{R}$] is a pair $(\mathcal{N}, E)$ such that $E \subseteq ((\mathbb{R} \cap [S, S'])^K)^\omega$. We assume from now on that $(\mathbb{Q} \cap [S, S'])^K$ and $(\mathbb{R} \cap [S, S'])^K$ are equipped with the induced topologies of $\mathbb{Q}^K$ and $\mathbb{R}^K$, and that $((\mathbb{Q} \cap [S, S'])^K)^\omega$ and $((\mathbb{R} \cap [S, S'])^K)^\omega$ are equipped with the product topologies of these induced topologies, respectively. Moreover, E is always assumed to be a closed subset of these Polish subspaces, and hence is also Polish [12].[2]

Finally, given some nondeterministic Ev-RNN $\mathcal{N}$, an infinite input stream $s \in (\mathbb{B}^M)^\omega$ is called *meaningful* if there exists some evolution $e \in E$ such that $inf(c'_{(s,e)})$ is a meaningful attractor, and it is called *spurious* otherwise, i.e., if for all evolution $e \in E$, the set $inf(c'_{(s,e)})$ is a spurious attractor. The set of all meaningful input streams of $\mathcal{N}$ is called the *neural ω-language* of $\mathcal{N}$ and is denoted by $L(\mathcal{N})$. An arbitrary set of input streams $L \subseteq (\mathbb{B}^M)^\omega$ is said to be *recognizable* by some nondeterministic Ev-RNN if there exists a nondeterministic network $(\mathcal{N}, E)$ such that $L(\mathcal{N}) = L$.

5 Results

Following considerations from ω-languages and automata theory [17], the expressive power of hybrid neural networks is characterized as the topological complexity of their underlying neural ω-language. For the sake of clarity, we first recall the results obtained in the deterministic context [5]. In this case, the static rational-weighted hybrid neural networks are computationally equivalent to deterministic Muller Turing machines, hence recognize neural ω-languages inside the class of finite Boolean combinations of $\mathbf{\Pi^0_2}$-sets ($BC(\mathbf{\Pi^0_2})$). The other models of static real-weighted, evolving rational-weighted, and evolving real-weighted hybrid networks are all computationally equivalent. They recognize precisely all the $BC(\mathbf{\Pi^0_2})$ neural ω-languages and, therefore, are strictly more powerful than deterministic Büchi and Muller Turing machines, since these later cannot recognize the whole class of $BC(\mathbf{\Pi^0_2})$-sets (cf. Sect. 3).

Here, we show that both models of rational- and real-weighted nondeterministic hybrid neural networks are computationally equivalent, and recognize precisely the class of all analytic sets ($\mathbf{\Sigma^1_1}$ boldface). Therefore, their expressive powers strictly encompass those of Büchi and Muller Turing machines, which are restricted to the effectively analytic sets (Σ^1_1 lightface) (cf. Sect. 3).

We first show that any analytic neural ω-language L can be recognized by some nondeterministic rational Ev-RNN $\mathcal{N}$. The idea of the proof is the following. First, we note that the analytic set L can be written as the first projection π_1 of some $\mathbf{\Pi^0_2}$-set $X \subseteq (\mathbb{B}^M)^\omega \times \{0,1\}^\omega$ (cf. Sect. 2). Next, we consider some recursive encoding of X by an infinite word $w_X \in \{0,1\}^\omega$. Afterwards, we consider a

[2] The results of the paper hold equally true even with E taken as $\mathbf{\Pi^0_2}$.

nondeterministic Ev-RNN[$\mathbb{Q}$] $\mathcal{N}$ equipped with only two possible evolving synaptic connections: one which might follow any possible binary evolution $e \in \{0,1\}^\omega$, and the other one which always follows the same binary evolution $w_X \in \{0,1\}^\omega$. We then design the static part of $\mathcal{N}$ such that $\mathcal{N}$ visits a meaningful attractor iff the current input s and evolving synaptic pattern $e \in \{0,1\}^\omega$ are such that (s,e) belongs the set encoded by w_X, namely X. In this way, $L(\mathcal{N}) = \pi_1(X) = L$, and thus L is recognized by $\mathcal{N}$.

Proposition 1. *Let $L \subseteq (\mathbb{B}^M)^\omega$ such that $L \in \mathbf{\Sigma^1_1}$. Then there exists some nondeterministic Ev-RNN[$\mathbb{Q}$] $(\mathcal{N}, E)$ such that $L(\mathcal{N}) = L$.*

Proof. Since $L \in \mathbf{\Sigma^1_1}$, there exists some $X \subseteq (\mathbb{B}^M)^\omega \times \{0,1\}^\omega$ such that $X \in \mathbf{\Pi^0_2}$ and $L = \pi_1(X)$. Since $X \in \mathbf{\Pi^0_2}$, it can be written as $X = \bigcap_{i \geq 0} \bigcup_{j \geq 0} (p_{i,j} \cdot (\mathbb{B}^M)^\omega \times q_{i,j} \cdot \{0,1\}^\omega)$, where each $(p_{i,j}, q_{i,j}) \in (\mathbb{B}^M)^* \times \{0,1\}^*$. Consequently, the set X (and hence also L) is completely determined by the countable sequence of pairs of finite prefixes $((p_{i,j}, q_{i,j}))_{i,j \geq 0}$. We can thus consider some encoding $w_X \in \{0,1\}^\omega$ of the sequence $((p_{i,j}, q_{i,j}))_{i,j \geq 0}$ such that, for any pair of indices $(i,j) \in \mathbb{N} \times \mathbb{N}$, the decoding procedure $(w_X, i, j) \mapsto (p_{i,j}, q_{i,j})$ is actually recursive.

We now consider the infinite procedure given by Algorithm 1 below. This procedure requires as input and auxiliary items the following three infinite sequences delivered step by step: an infinite input stream $s \in (\mathbb{B}^M)^\omega$, an infinite word $e \in \{0,1\}^\omega$ chosen arbitrarily, and the precise infinite word $w_X \in \{0,1\}^\omega$. Note that provided that these three items are correctly supplied by some external source, every instruction of the procedure is actually recursive. Farther note that, by construction, the procedure returns infinitely many 1's iff the pair of infinite sequences (s,e) belongs to X.

Based on the infinite procedure, we provide the description of a nondeterministic Ev-RNN[$\mathbb{Q}$] $(\mathcal{N}, E)$ such that $L(\mathcal{N}) = L$. The network $(\mathcal{N}, E)$ contains only two evolving synaptic weights $w_1(t)$ and $w_2(t)$ which evolve among only two possible values, 0 or 1. All other synaptic weights are static. The weight $w_1(t)$ might follow every possible evolution in $\{0,1\}^\omega$, while $w_2(t)$ always follows the same evolution, which are the successive letters of w_X. Formally, one has the following *closed* set of possible evolutions:

$$E = \{\tilde{e} \in (\{0,1\}^2)^\omega : (\tilde{\boldsymbol{e}}(t))_0 \in \{0,1\} \text{ and } (\tilde{\boldsymbol{e}}(t))_1 = w_X(t), \text{ for any t} \geq 0\}.$$

We then consider a neural circuit which stores the incoming values of the input stream $s \in (\mathbb{B}^M)^\omega$ into M designated neurons, as well as two neural circuits which store the successive bits of $w_1(t)$ and $w_2(t)$ into two designated neurons (see [20] for further technical details). Afterwards, according to the real time computational equivalence between static RNN[$\mathbb{Q}$] and Turing machines [20], we consider a static RNN[$\mathbb{Q}$] which is suitably designed and connected to the above mentioned circuits in order to simulate all the recursive instructions of Algorithm 1. We finally add a single Boolean output neuron y and update the whole construction in order that y takes an activation value of 1 precisely when the simulation of Algorithm 1 by our network enters the instruction "returns 1".

In this way, one has the description of a nondeterministic Ev-RNN[$\mathbb{Q}$] $(\mathcal{N}, E)$ which suitably simulates the behavior of Algorithm 1.

Besides, the single output cell y leads to the existence of only three possible attractors, namely $\{(0)\}, \{(0),(1)\}$, and $\{(1)\}$. We set $\{(0)\}$ as spurious, and $\{(0),(1)\}$ and $\{(1)\}$ as meaningful. This means that $(\mathcal{N}, E)$ visits a meaningful attractor iff the simulation of Algorithm 1 returns infinitely many 1's.

According to all the previous considerations, one has that $s \in L(\mathcal{N})$ iff, by definition, there exists some $\tilde{e} \in E$ such that $inf(c'_{(s,\tilde{e})})$ is meaningful, iff there exists $e \in \{0,1\}^\omega$ such that the simulation of Algorithm 1 returns infinitely many 1's, iff there exists $e \in \{0,1\}^\omega$ such that the pair $(s,e) \in X$, iff, by definition, $s \in \pi_1(X) = L$. In other words, $L(\mathcal{N}) = L$, showing that L is recognized by the nondeterministic Ev-RNN[$\mathbb{Q}$] $(\mathcal{N}, E)$. □

Algorithm 1. Infinite procedure

Require:

1. Input $s = \boldsymbol{s}(0)\boldsymbol{s}(1)\boldsymbol{s}(2)\cdots \in (\mathbb{B}^M)^\omega$ supplied step by step at successive time steps $t = 0,1,2,\ldots$
2. some auxiliary infinite word $e = e(0)e(1)e(2)\cdots \in \{0,1\}^\omega$ supplied step by step at successive time steps $t = 0,1,2,\ldots$
3. the specific auxiliary infinite word $w_X = w_X(0)w_X(1)w_X(2)\cdots \in \{0,1\}^\omega$ supplied step by step at successive time steps $t = 0,1,2,\ldots$

SUBROUTINE 1
$c \leftarrow 0$ // c counts the number of letters provided so far
for all time step $t \geq 0$ **do**
  store each incoming Boolean vector $\boldsymbol{s}(t) \in \mathbb{B}^M$
  store each incoming bit $e(t) \in \{0,1\}$
  store each incoming bit $w_X(t) \in \{0,1\}$
  $c \leftarrow c+1$
end for
END SUBROUTINE 1

SUBROUTINE 2
$i \leftarrow 0$, $j \leftarrow 0$
loop
  wait until $c \geq \max\{|p_{i,j}|, |q_{i,j}|\}$
  wait until $w_X[0{:}c]$ becomes long enough to contain the encoding of $(p_{i,j}, q_{i,j})$
  decode $(p_{i,j}, q_{i,j})$ from $w_X[0{:}c]$ // recursive procedure
  if $p_{i,j} \subseteq s[0{:}c]$ and $q_{i,j} \subseteq e[0{:}c]$ **then** // $(s,e) \in p_{i,j} \cdot (\mathbb{B}^M)^\omega \times q_{i,j} \cdot \{0,1\}^\omega$
    return 1 // $\exists\, j$ s.t. $(s,e) \in p_{i,j} \cdot (\mathbb{B}^M)^\omega \times q_{i,j} \cdot \{0,1\}^\omega$
    $i \leftarrow i+1$, $j \leftarrow 0$ // test if $(s,e) \in p_{i+1,0} \cdot (\mathbb{B}^M)^\omega \times q_{i+1,0} \cdot \{0,1\}^\omega$
  else // $(s,e) \notin p_{i,j} \cdot (\mathbb{B}^M)^\omega \times q_{i,j} \cdot \{0,1\}^\omega$
    return 0 // $\neg\exists j' \leq j$ s.t. $(s,e) \in p_{i,j'} \cdot (\mathbb{B}^M)^\omega \times q_{i,j'} \cdot \{0,1\}^\omega$
    $i \leftarrow i$, $j \leftarrow j+1$ // test if $(s,e) \in p_{i,j+1} \cdot (\mathbb{B}^M)^\omega \times q_{i,j+1} \cdot \{0,1\}^\omega$
  end if
end loop
END SUBROUTINE 2

We now conversely show that every ω-language recognized by some nondeterministic Ev-RNN is analytic.

Proposition 2. *Let $(\mathcal{N}, E)$ be some nondeterministic Ev-RNN[$\mathbb{R}$]. Then $L(\mathcal{N}) \in \mathbf{\Sigma}^1_1$.*

Proof. First of all, note that the dynamics of $(\mathcal{N}, E)$ can naturally be associated with the function $f_{(\mathcal{N},E)} : (\mathbb{B}^M)^\omega \times E \to (\mathbb{B}^P)^\omega$ defined by $f_{(\mathcal{N},E)}(s,e) = c'_{(s,e)}$. The nature of our dynamics ensures that this function is sequential, i.e., for any time step $t \geq 0$, the vectors $\boldsymbol{s}(t)$, $\boldsymbol{e}(t)$ and $\boldsymbol{y}(t)$ are generated simultaneously. Therefore, given any basic open set $w \cdot (\mathbb{B}^P)^\omega$, with $w \in (\mathbb{B}^P)^*$, one has that $f^{-1}_{(\mathcal{N},E)}(w \cdot (\mathbb{B}^P)^\omega)$ is of the form $\Theta_w = \bigcup_{i \in I} \left[u_i \cdot (\mathbb{B}^M)^\omega \times (v_{\mathbb{R},i} \cdot ([S,S']^K)^\omega \cap E)\right]$ with each $u_i \in (\mathbb{B}^M)^{|w|}$ and $v_{\mathbb{R},i} \in ([S,S']^K)^{|w|}$. Notice that for each $i \in I$, $v_{\mathbb{R},i} \cdot ([S,S']^K)^\omega \cap E$ is closed (inside E) and $u_i \cdot (\mathbb{B}^M)^\omega$ is clopen, and hence $(u_i \cdot (\mathbb{B}^M)^\omega) \times (v_{\mathbb{R},i} \cdot ([S,S']^K)^\omega \cap E)$ is closed. By [4, Lemma 9], it follows that given any u_i and $v_{\mathbb{R},i}$ as above, there exists $I_{\mathbb{Q},i} = (\prod_{k=1}^K]a_{j,k}, b_{j,k}[)_{j<|w|}$, where each $a_{j,k}, b_{j,k} \in \mathbb{Q}$ and $v_{\mathbb{R},i} \in I_{\mathbb{Q},i}$, and such that

$$f_{(\mathcal{N},E)} \left[u_i \cdot (\mathbb{B}^M)^\omega \times (I_{\mathbb{Q},i} \cdot ([S,S']^K)^\omega \cap E)\right] \subseteq w \cdot (\mathbb{B}^P)^\omega.$$

One thus has $\Theta_w = \bigcup_{i \in I} \left[u_i \cdot (\mathbb{B}^M)^\omega \times (I_{\mathbb{Q},i} \cdot ([S,S']^K)^\omega \cap E)\right]$. Since there exist only countably many u_i and $I_{\mathbb{Q},i}$, it turns out that Θ_w is a countable union of closed sets, i.e. a $\mathbf{\Sigma}^0_2$ set, which shows that $f_{(\mathcal{N},E)}$ is of Baire class 1, cf. [12].[3]

Furthermore, note that since $\mathcal{N}$ contains finitely many output cells, is also has finitely many possible Boolean states, and therefore also finitely many possible attractors. This feature is independent from the nondeterministic behavior associated with the set of possible evolutions E. Hence, suppose that $\mathcal{N}$ contains the I meaningful attractors $A_i = \{\boldsymbol{b}_{i_1}, \dots, \boldsymbol{b}_{i_{k(i)}}\}$, for $i = 1, \dots, I$, where $1 \leq i_1 < \dots < i_{k(i)} \leq 2^P$, and where $\boldsymbol{b}_n$ denotes the n-th Boolean vector of $\mathbb{B}^P$ according to the lexicographic order.

According to these considerations, the ω-language $L(\mathcal{N})$ can be expressed by the following sequence of equalities:

$$\begin{aligned}
L(\mathcal{N}) &= \left\{s \in (\mathbb{B}^M)^\omega : \text{there exists } e \in E \text{ s.t.} \inf(c'_{(s,e)}) \text{is a meaningful attractor}\right\}\\
&= \left\{s \in (\mathbb{B}^M)^\omega : \text{there exists } e \in E \text{ s.t.} \inf(c'_{(s,e)}) = A_i, \text{ for some } i = 1,\dots,I\right\}\\
&= \pi_1\Big(\left\{(s,e) \in (\mathbb{B}^M)^\omega \times E : \inf(c'_{(s,e)}) = A_i, \text{ for some } i = 1,\dots,I\right\}\Big)\\
&= \pi_1\Big(\bigcup_{i=1}^{I} \left\{(s,e) \in (\mathbb{B}^M)^\omega \times E : \inf(c'_{(s,e)}) = A_i\right\}\Big)\\
&= \pi_1\Big(\bigcup_{i=1}^{I} \big\{(s,e) \in (\mathbb{B}^M)^\omega \times E : \forall j \in \{i_1,\dots,i_{k(i)}\}, f_{(\mathcal{N},E)}(s,e) \text{ has } \infty\text{-many } \boldsymbol{b}'_j\text{s}\\
&\qquad\qquad \text{and } \forall j \in \{1,\dots,2^P\} \setminus \{i_1,\dots,i_{k(i)}\}, f_{(\mathcal{N},E)}(s,e) \text{ has finitely many } \boldsymbol{b}'_j\text{s}\big\}\Big)
\end{aligned}$$

[3] We recall that the preimage by a Baire class 1 function of a set in $\mathbf{\Sigma^0_n}$ (resp. $\mathbf{\Pi^0_n}$) is in $\mathbf{\Sigma^0_{n+1}}$ (resp. $\mathbf{\Pi^0_{n+1}}$).

$$
\begin{aligned}
&= \pi_1\Big(\bigcup_{i=1}^{I}\Big[\bigcap_{j\in\{i_1,\ldots,i_{k(i)}\}}\big\{(s,e)\in(\mathbb{B}^M)^\omega\times E: \\
&\qquad f_{(\mathcal{N},E)}(s,e)\in \underbrace{\bigcap_{n\geq 0}\bigcup_{m\geq 0}(\mathbb{B}^P)^{n+m}\cdot \boldsymbol{b}_j\cdot(\mathbb{B}^P)^\omega}_{\substack{c'_{(s,e)} \text{ contains infinitely many } \boldsymbol{b}'_j \text{ s, i.e.} \\ \forall n\geq 0\ \exists m\geq n\ \boldsymbol{y}(n+m)=\boldsymbol{b}_j,\ \text{thus in } \boldsymbol{\Pi^0_2}}}\big\}\cap \\
&\qquad \bigcap_{j\in\substack{\{1,\ldots,2^P\}\setminus\\ \{i_1,\ldots,i_{k(i)}\}}}\big\{(s,e)\in(\mathbb{B}^M)^\omega\times E: \\
&\qquad f_{(\mathcal{N},E)}(s,e)\in \underbrace{\Big(\bigcap_{n\geq 0}\bigcup_{m\geq 0}(\mathbb{B}^P)^{n+m}\cdot \boldsymbol{b}_j\cdot(\mathbb{B}^P)^\omega\Big)^{\complement}}_{\substack{c'_{(s,e)} \text{ contains only finitely many } \boldsymbol{b}'_j \text{ s, i.e.} \\ \text{complement of a } \boldsymbol{\Pi^0_2}\text{-set, thus in } \boldsymbol{\Sigma^0_2}}}\big\}\Big]\Big) \\
&= \pi_1\Big(\bigcup_{i=1}^{I}\Big[\bigcap_{j\in\{i_1,\ldots,i_{k(i)}\}}\underbrace{f^{-1}_{(\mathcal{N},E)}\Big(\bigcap_{n\geq 0}\bigcup_{m\geq 0}(\mathbb{B}^P)^{n+m}\cdot \boldsymbol{b}_j\cdot(\mathbb{B}^P)^\omega\Big)}_{\text{preimage by a Baire class 1 function of a } \boldsymbol{\Pi^0_2}\text{-set, thus in } \boldsymbol{\Pi^0_3}[12]}\cap \\
&\qquad \bigcap_{j\in\substack{\{1,\ldots,2^P\}\setminus\\ \{i_1,\ldots,i_{k(i)}\}}}\underbrace{f^{-1}_{(\mathcal{N},E)}\Big(\Big(\bigcap_{n\geq 0}\bigcup_{m\geq 0}(\mathbb{B}^P)^{n+m}\cdot \boldsymbol{b}_j\cdot(\mathbb{B}^P)^\omega\Big)^{\complement}\Big)}_{\text{preimage by a Baire class 1 function of a } \boldsymbol{\Sigma^0_2}\text{-set, thus in } \boldsymbol{\Sigma^0_3}[12]}\Big]\Big)
\end{aligned}
$$

It follows that $L(\mathcal{N})$ is a projection of a finite union and intersection of $\boldsymbol{\Pi^0_3}$ and $\boldsymbol{\Sigma^0_3}$ subsets of the Polish space $(\mathbb{B}^M)^\omega \times E$, and therefore, $L(\mathcal{N}) \in \boldsymbol{\Sigma^1_1}$. □

Finally, Propositions 1 and 2 allow to conclude that nondeterministic evolving neural networks recognize precisely the set of all analytic sets, irrespective of whether their synaptic weights are modelled by rational or real numbers.

Theorem 1. *Let $L \subseteq (\mathbb{B}^M)^\omega$. The following conditions are equivalent:*

1. *$L \in \boldsymbol{\Sigma^1_1}$;*
2. *L is recognizable by some nondeterministic Ev-RNN[$\mathbb{Q}$] $(\mathcal{N}, E)$;*
3. *L is recognizable by some nondeterministic Ev-RNN[$\mathbb{R}$] $(\mathcal{N}, E)$.*

Proof. (1) → (2) is provided by Proposition 1. (2) → (3) holds by definition. (3) → (1) is provided by Proposition 2. □

6 Discussion

We have introduced a model of nondeterministic hybrid recurrent neural networks. The nondeterminism is expressed as a set of possible synaptic evolutions associated with each neural network. The network chooses one of these in a nondeterministic manner, and then sticks to it throughout its whole computational process. In this context, we have proven that the two models of rational-weighted and real-weighted nondeterministic hybrid neural networks are computationally equivalent, and recognize precisely the class of all $\boldsymbol{\Sigma^1_1}$ neural ω-languages.

They are therefore strictly more expressive than the nondeterministic Büchi and Muller Turing machines, which recognize the Σ^1_1 (lighface) ω-languages.

These results together with those of [5] show that nondeterminism injects an extensive amount of computational power – from $BC(\mathbf{\Pi^0_2})$ to $\mathbf{\Sigma^1_1}$ – to the hybrid neural systems. Besides, as opposed to the deterministic case, the consideration of real synaptic weights in the present nondeterministic context does actually not add any additional computational power to the neural networks. The added value of the power of the continuum is somehow absorbed by the nondeterminism, and any kind of analog assumption can therefore be dropped without compromizing the achievement of a maximal computational power. More generally, these achievements support the idea that the nondeterminism plays a crucial role in neural information processing. They also support the claim that recurrent neural networks represent a natural model of computation beyond the Turing limits [8].

For future work, the study of the computational capabilities of more biologically-oriented neural models involved in more bio-inspired paradigms of computation is expected to be pursued.

Finally, we hope that such comparative studies between the computational capabilities of neural models and abstract machines might eventually bring further insight to the understanding of the intrinsic natures of both biological as well as artificial intelligences.

References

1. Cabessa, J.: Interactive evolving recurrent neural networks are super-Turing. In: Filipe, J., Fred, A.L.N. (eds.) Proceedings of ICAART, pp. 328–333. SciTePress (2012)
2. Cabessa, J., Siegelmann, H.T.: Evolving recurrent neural networks are super-Turing. In: Proceedings of IJCNN 2011, pp. 3200–3206. IEEE (2011)
3. Cabessa, J., Siegelmann, H.T.: The computational power of interactive recurrent neural networks. Neural Comput. **24**(4), 996–1019 (2012)
4. Cabessa, J., Siegelmann, H.T.: The super-Turing computational power of plastic recurrent neural networks. Int. J. Neural. Syst. **24**(8), 1450029 (2014)
5. Cabessa, J., Villa, A.E.P.: Computational capabilities of recurrent neural networks based on their attractor dynamics. In: Proceedings of IJCNN 2015. IEEE (to appear, 2015) (accepted)
6. Cabessa, J., Villa, A.E.P.: A hierarchical classification of first-order recurrent neural networks. In: Dediu, A.-H., Fernau, H., Martín-Vide, C. (eds.) LATA 2010. LNCS, vol. 6031, pp. 142–153. Springer, Heidelberg (2010)
7. Cabessa, J., Villa, A.E.P.: The expressive power of analog recurrent neural networks on infinite input streams. Theor. Comput. Sci. **436**, 23–34 (2012)
8. Cabessa, J., Villa, A.E.P.: Recurrent neural networks - a natural model of computation beyond the Turing limits. In: Rosa, A.C., et al., (ed.) Proceedings of IJCCI 2012, pp. 594–599. SciTePress (2012)
9. Cabessa, J., Villa, A.E.P.: The super-turing computational power of interactive evolving recurrent neural networks. In: Mladenov, V., Koprinkova-Hristova, P., Palm, G., Villa, A.E.P., Appollini, B., Kasabov, N. (eds.) ICANN 2013. LNCS, vol. 8131, pp. 58–65. Springer, Heidelberg (2013)

10. Cabessa, J., Villa, A.E.P.: An attractor-based complexity measurement for boolean recurrent neural networks. PloS ONE **9**(4), e94204+ (2014)
11. Cabessa, J., Villa, A.E.P.: Interactive evolving recurrent neural networks are superturing universal. In: Wermter, S., Weber, C., Duch, W., Honkela, T., Koprinkova-Hristova, P., Magg, S., Palm, G., Villa, A.E.P. (eds.) ICANN 2014. LNCS, vol. 8681, pp. 57–64. Springer, Heidelberg (2014)
12. Kechris, A.S.: Classical Descriptive Set Theory. Graduate Texts in Mathematics, vol. 156. Springer-Verlag, New York (1995)
13. Kleene, S.C.: Representation of events in nerve nets and finite automata. In: Shannon, C., McCarthy, J. (eds.) Automata Studies, pp. 3–41. Princeton University Press, Princeton, NJ (1956)
14. McCulloch, W.S., Pitts, W.: A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biophys. **5**, 115–133 (1943)
15. Minsky, M.L.: Computation: Finite and Infinite Machines. Prentice-Hall Inc., Englewood Cliffs (1967)
16. Neumann, J.V.: The Computer and the Brain. Yale University Press, New Haven (1958)
17. Perrin, D., Pin, J.-E.: Infinite Words - Automata, Semigroups, Logic and Games. Pure and Applied Mathematics, vol. 141. Elsevier, San Diego (2004)
18. Siegelmann, H.T.: Neural Networks and Analog Computation: Beyond the Turing Limit. Birkhauser Boston Inc., Cambridge (1999)
19. Siegelmann, H.T., Sontag, E.D.: Analog computation via neural networks. Theor. Comput. Sci. **131**(2), 331–360 (1994)
20. Siegelmann, H.T., Sontag, E.D.: On the computational power of neural nets. J. Comput. Syst. Sci. **50**(1), 132–150 (1995)
21. Staiger, L.: ω-languages. In: Rozenberg, G., Salomaa, A. (eds.) Handbook of Formal Languages: Beyond Words, vol. 3, pp. 339–387. Springer-Verlag, New York (1997)
22. Thomas, W.: Automata on infinite objects. In: van Leeuwen, J. (ed.) Handbook of Theoretical Computer Science: Formal Models and Semantics, vol. B, pp. 133–192. Elsevier and MIT Press, Amsterdam (1990)
23. Turing, A.M.: Intelligent machinery. Technical report, National Physical Laboratory, Teddington, UK (1948)
24. Vaadia, E., Haalman, I., Abeles, M., Bergman, H., Prut, Y., Slovin, H., Aertsen, A.: Dynamics of neuronal interactions in monkey cortex in relation to behavioural events. Nature **373**(6514), 515–518 (1995)
25. Villa, A.E.P., Tetko, I.V., Hyland, B., Najem, A.: Spatiotemporal activity patterns of rat cortical neurons predict responses in a conditioned task. Proc. Natl. Acad. Sci. U.S.A. **96**(3), 1106–1111 (1999)
26. Wagner, K.: On ω-regular sets. Inf. Control **43**(2), 123–177 (1979)