

# Estimating the interaction graph of stochastic neural dynamics

ALINE DUARTE<sup>1,\*</sup>, ANTONIO GALVES<sup>1,\*\*</sup>, EVA LÖCHERBACH<sup>2</sup> and GUILHERME OST<sup>1,†</sup>

*Dedicated to Enza Orlandi, in memoriam*

<sup>1</sup>*Instituto de Matemática e Estatística, Universidade de São Paulo, Rua do Matão 1010, 05508-090, São Paulo, Brazil. E-mail: \*alineduarte@usp.br; \*\*galves@usp.br; †guilhermeost@gmail.com*

<sup>2</sup>*AGM, Université de Cergy-Pontoise, CNRS-UMR 8088, 95000 Cergy-Pontoise, France. E-mail: eva.loecherbach@u-cergy.fr*

In this paper, we address the question of statistical model selection for a class of stochastic models of biological neural nets. Models in this class are systems of interacting chains with memory of variable length. Each chain describes the activity of a single neuron, indicating whether it spikes or not at a given time. The spiking probability of a given neuron depends on the time evolution of its *presynaptic neurons* since its last spike time. When a neuron spikes, its potential is reset to a resting level and postsynaptic current pulses are generated, modifying the membrane potential of all its *postsynaptic neurons*. The relationship between a neuron and its pre- and postsynaptic neurons defines an oriented graph, the *interaction graph* of the model. The goal of this paper is to estimate this graph based on the observation of the spike activity of a finite set of neurons over a finite time. We provide explicit exponential upper bounds for the probabilities of under- and overestimating the interaction graph restricted to the observed set and obtain the strong consistency of the estimator. Our result does not require stationarity nor uniqueness of the invariant measure of the process.

*Keywords:* biological neural nets; graph of interactions; interacting chains of variable memory length; statistical model selection

## 1. Introduction

This paper addresses the question of statistical model selection for a class of stochastic processes describing biological neural networks. The activity of the neural net is described by a countable system of interacting chains with memory of variable length representing the spiking activity of the different neurons. The interactions between neurons are defined in terms of their interaction neighborhoods. The interaction neighborhood of a neuron is given by the set of all its presynaptic neurons. We introduce a statistical selection procedure in this class of stochastic models to estimate the interaction neighborhoods.

The stochastic neural net we consider can be described as follows. Each neuron spikes with a probability which is an increasing function of its membrane potential. The membrane potential of a given neuron depends on the accumulated spikes coming from the presynaptic neurons since its last spike time. When a neuron spikes, its potential is reset to a resting level and at the same time postsynaptic current pulses are generated, modifying the membrane potential of all its postsynaptic neurons.

Recently, several papers have been devoted to the probabilistic study of these models, starting with [16] who provided a rigorous mathematical framework to study such systems with an infinite number of interacting components, evolving in discrete time. Its continuous time version has been subsequently studied in [10–12,14,21,29] and [32]. We also refer to [7] and the references therein for a simulation study and mean field analysis. All these papers deal with probabilistic aspects of the model, not with statistical model selection.

Statistical model selection for graphical models has been largely discussed in the literature. Recently much effort has been devoted to estimating the graph of interactions underlying for example, finite volume Ising models ([4,26,27] and [3]), infinite volume Ising models ([17,23] and [24]), Markov random fields ([9] and [31]) and variable-neighborhood random fields ([25]).

Graphical models are very interesting from a mathematical point of view. However, their application to the stochastic modeling of neural data has a major drawback, namely the assumption that the configuration describing the neural activity at a given time follows a Gibbsian distribution. To the best of our knowledge, this Gibbsian assumption is not supported by biological considerations of any kind.

Statistical methods for selecting the graph of interactions in neural nets start probably with [6] and [5]. Recently, new approaches have been proposed by [28] in the framework of multivariate point processes and Hawkes processes. Let us also mention [30] where a new experimental design is introduced and studied from a numerical point of view. These articles have the following drawback. The first mentioned ones only consider finite systems of neurons or processes having a fixed finite memory, the lastly cited does only propose a numerical study.

The present paper provides a rigorous mathematical approach to the problem of inference in neural dynamics. Its main features are the following.

1. The processes we consider are not Markovian. They are systems of interacting chains with memory of variable length.
2. Our approach does not rely on any Gibbsian assumption.
3. We can deal with the case in which the system possesses several invariant measures.
4. Infinite systems of neurons can be also treated under suitable assumptions on the synaptic weights.

Let us briefly comment on Features 3 and 4. Feature 3 makes our model suitable to describe long-term memory in which the asymptotic law of the system depends on its initial configuration. This initial configuration can be seen as the effect of an external stimulus to which the system is exposed at the beginning of the experiment. In this perspective, the asymptotic distribution can be interpreted as the neural encoding of this stimulus by the brain. Feature 4 enables us to deal with arbitrarily high dimensional systems, taking into account the fact that the brain consists of a huge (about  $10^{11}$ ) number of interacting neurons.

The models we consider can be seen as a version of the Integrate and Fire (IF) model with random thresholds, but only in cases in which the postsynaptic current pulses are of the exponential type. Indeed, only in such cases, the time evolution of the family of membrane potentials is a Markov process. For general postsynaptic current pulses, this is not true.

Therefore, we could say that our model is a non-Markovian version of the IF model with random thresholds. This fact places the class of models considered here within a classical and widely accepted framework of modern neuroscience. Indeed, as pointed out by an anonymous

referee, IF models have a long and rich history, going back to the fundamental work [22]. For more insights on IF-models we refer the interested reader to classical textbooks such as [13] and [19].

The inherent randomness of the thresholds in our model leads to random neuronal responses instead of deterministic ones. The idea that the spike activity is intrinsically random and not deterministic can be traced back to [2], see also [1]. Under the name of “escape noise”-models, this question has then been further emphasized by [20] and [18].

We conclude this introduction by briefly describing the statistical selection procedure we propose. We observe the process within a finite sampling region during a finite time interval. For each neuron  $i$  in the sample, we estimate its spiking probability given the spike trains of all other neurons since its last spike time. For each neuron  $j \neq i$ , we then introduce a measure of sensibility of this conditional spiking probability with respect to changes within the spike train of neuron  $j$ . If this measure of sensibility is *statistically small*, we conclude that neuron  $j$  does not belong to the interaction neighborhood of neuron  $i$ .

For this selection procedure, we give precise error bounds for the probabilities of under- and overestimating finite interaction neighborhoods implying the consistency of the procedure in Theorem 1. For interaction neighborhoods which are not contained within the sampling region, a coupling between the process and its locally finite range approximation reduces the estimation problem to the situation of Theorem 1. The coupling result is presented in Proposition 4. In our proofs, we rely on a new conditional Hoeffding-type inequality which is of independent interest, stated in Proposition 1.

We stress that, in our class of models, the probability of a neuron to spike depends only on the history of the process since its last spike time. Therefore, temporal dependencies do not need to be estimated, making our estimation problem different from classical context tree estimation as considered in [9] and [15]. We refer the reader to the companion paper by [8] where our statistical selection procedure is explored, tested and applied to the analysis of simulated and real neural activity data.

The paper is organized as follows. In Section 2, we introduce the model and the selection procedure, present the main assumptions and formulate the main results, Theorems 1 and 2. In Section 3, we derive some exponential inequalities including a new conditional Hoeffding-type inequality, presented in Proposition 1, which is interesting by itself. The proofs of Theorems 1 and 2 are presented in Sections 4 and 5, respectively. In Section 6, we discuss the time complexity of our selection procedure. In the Appendix, we present the coupling result stated in Proposition 4.

## 2. Model definition and main results

### 2.1. A stochastic model for a system of interacting neurons

Throughout this article,  $I$  denotes a countable set,  $W_{j \rightarrow i} \in \mathbb{R}$  with  $i, j \in I$ , a collection of real numbers such that  $W_{j \rightarrow j} = 0$  for all  $j$ , and for  $i \in I$ ,  $\varphi_i : \mathbb{R} \rightarrow [0, 1]$  a non-decreasing measurable function and  $g_i = (g_i(t))_{t \geq 1}$  a sequence of strictly positive real numbers.

In order to be consistent with the neuroscience terminology (see [19]), we call

- $I$  the set of neurons,
- $W_{j \rightarrow i}$  the synaptic weight of neuron  $j$  on neuron  $i$ ,
- $\varphi_i$  the spike rate function of neuron  $i$ ,
- $g_i$  the postsynaptic current pulse of neuron  $i$ .

We consider a stochastic chain  $(X_t)_{t \in \mathbb{Z}}$  taking values in  $\{0, 1\}^I$ , defined on a suitable probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ . For each  $i \in I$  and  $t \in \mathbb{Z}$ ,  $X_t(i) = 1$ , if neuron  $i$  spikes at time  $t$  and  $X_t(i) = 0$ , otherwise. For each neuron  $i \in I$  and each time  $t \in \mathbb{Z}$ , let

$$L_t^i = \sup\{s \leq t : X_s(i) = 1\}, \tag{2.1}$$

be the last spike time of neuron  $i$  before time  $t$ . Here, we adopt the convention that  $\sup\{\emptyset\} = -\infty$ .

For each  $t \in \mathbb{Z}$ , we call  $\mathcal{F}_t$  the sigma algebra generated by the past events up to time  $t$ , that is,

$$\mathcal{F}_t := \sigma(X_s(j), s \leq t, j \in I).$$

The stochastic chain  $(X_t)_{t \in \mathbb{Z}}$  is defined as follows. For each time  $t \in \mathbb{Z}$ , for any finite set  $F \subset I$  and any choice  $a(i) \in \{0, 1\}$ ,  $i \in F$ ,

$$\mathbb{P}(X_{t+1}(i) = a(i), i \in F | \mathcal{F}_t) = \prod_{i \in F} \mathbb{P}(X_{t+1}(i) = a(i) | \mathcal{F}_t) \quad \mathbb{P}\text{-a.s.}, \tag{2.2}$$

where for each  $i \in I$  and  $t \in \mathbb{Z}$ ,

$$\mathbb{P}(X_{t+1}(i) = 1 | \mathcal{F}_t) = \varphi_i \left( \sum_{j \in I} W_{j \rightarrow i} \sum_{s=L_t^j+1}^t g_j(t+1-s) X_s(j) \right) \quad \mathbb{P}\text{-a.s.}, \tag{2.3}$$

if  $L_t^i < t$ , and

$$\mathbb{P}(X_{t+1}(i) = 1 | \mathcal{F}_t) = \varphi_i(0) \quad \mathbb{P}\text{-a.s.},$$

otherwise.

Suitable conditions ensuring the existence of a stochastic chain of this type are presented at the beginning of Section 2.3.

## 2.2. Neighborhood estimation procedure

We write

$$x = (x_t(i))_{-\infty < t \leq 0, i \in I}$$

for any configuration  $x \in \{0, 1\}^{I \times \{\dots, -1, 0\}}$ , and for any  $F \subset I$ , we write

$$x_t(F) = (x_t(i), i \in F).$$

Moreover, for any  $x \in \{0, 1\}^{I \times \{\dots, -1, 0\}}$ ,

$$X_{-\infty}^0 = x, \quad \text{if } X_t(i) = x_t(i) \text{ for all } -\infty < t \leq 0 \text{ and for all } i \in I.$$

Finally, for any  $\ell \geq 1, t \in \mathbb{Z}, F \subset I$  and  $w \in \{0, 1\}^{\{-\ell, \dots, -1\} \times F}$ , we write

$$X_{t-\ell}^{t-1}(F) = w, \quad \text{if } X_{t-s}(j) = w_{-s}(j), \text{ for all } 1 \leq s \leq \ell \text{ and for all } j \in F$$

and

$$X_{t-\ell-1}^{t-1}(i) = 10^\ell, \quad \text{if } X_{t-s}(i) = 0, \text{ for all } 1 \leq s \leq \ell \text{ and } X_{t-\ell-1}(i) = 1.$$

Throughout the article,  $s, t \in \mathbb{Z}$  will be time indices, while  $n \in \mathbb{N}$  will be saved for future use as the length of the time interval during which the neural network is observed.

Let

$$V_i = \{j \in I \setminus \{i\} : W_{j \rightarrow i} \neq 0\} \tag{2.4}$$

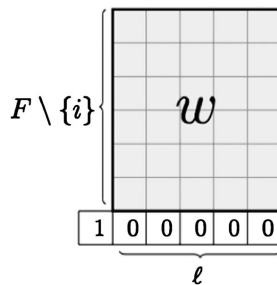
be the set of presynaptic neurons of neuron  $i$ . The set  $V_i$  is called the *interaction neighborhood* of neuron  $i$ . The goal of our statistical selection procedure is to identify the set  $V_i$  from the data in a consistent way.

Let  $X_1(F), \dots, X_n(F)$  be a sample where  $F \subset I$  is a finite sampling region and  $n \geq 3$  is the length of the time interval during which the network has been observed. For any fixed  $i \in F$ , we want to estimate its interaction neighborhood  $V_i$ .

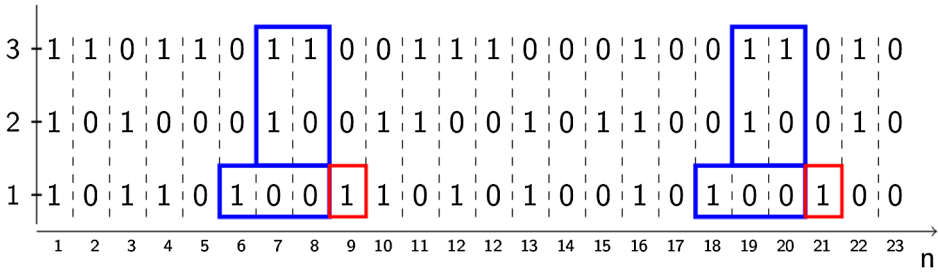
Our procedure is defined as follows. For each  $1 \leq \ell \leq n - 2$ , local past  $w \in \{0, 1\}^{\{-\ell, \dots, -1\} \times F \setminus \{i\}}$  outside of  $i$  (see Figure 1) and symbol  $a \in \{0, 1\}$ , we define

$$N_{(i,n)}(w, a) = \sum_{t=\ell+2}^n \mathbf{1}\{X_{t-\ell-1}^{t-1}(i) = 10^\ell, X_{t-\ell}^{t-1}(F \setminus \{i\}) = w, X_t(i) = a\}.$$

The random variable  $N_{(i,n)}(w, a)$  counts the number of occurrences of  $w$  followed or not by a spike of neuron  $i$  ( $a = 1$  or  $a = 0$ , respectively) in the sample  $X_1(F), \dots, X_n(F)$ , when the last spike of neuron  $i$  has occurred  $\ell + 1$  time steps before in the past, see Figure 2.



**Figure 1.** Local past  $w \in \{0, 1\}^{\{-\ell, \dots, -1\} \times F \setminus \{i\}}$  outside of  $i$  with  $\ell = 5$  and  $|F| = 7$ .



**Figure 2.** Example for  $N_{(i,n)}(w, 1) = 2$ , where  $i = 1$ , for a given word  $w$  (in blue),  $\ell = 2$ ,  $|F| = 3$  and  $n = 23$ .

We define the empirical probability of neuron  $i$  having a spike at the next step given  $w$  by

$$\hat{p}_{(i,n)}(1|w) = \frac{N_{(i,n)}(w, 1)}{N_{(i,n)}(w)}, \tag{2.5}$$

when  $N_{(i,n)}(w) := N_{(i,n)}(w, 0) + N_{(i,n)}(w, 1) > 0$ .

For any fixed parameter  $\xi \in (0, 1/2)$ , we consider the following set

$$\mathcal{T}_{(i,n)} = \left\{ w \in \bigcup_{\ell=1}^{n-2} \{0, 1\}^{\{-\ell, \dots, -1\} \times F \setminus \{i\}} : N_{(i,n)}(w) \geq n^{1/2+\xi} \right\}. \tag{2.6}$$

We use the notation  $|w| = \ell$  whenever  $w \in \{0, 1\}^{\{-\ell, \dots, -1\} \times F \setminus \{i\}}$ . If  $v, w$  both belong to  $\{0, 1\}^{\{-\ell, \dots, -1\} \times F \setminus \{i\}}$  we write

$$v_{\{j\}^c} = w_{\{j\}^c} \quad \text{if and only if} \quad v_{-\ell}^{-1}(F \setminus \{j\}) = w_{-\ell}^{-1}(F \setminus \{j\}).$$

In words, the equality  $v_{\{j\}^c} = w_{\{j\}^c}$  means that  $v$  and  $w$  coincide on all but the  $j$ th coordinate.

Finally, for each  $w \in \mathcal{T}_{(i,n)}$  and for any  $j \in F \setminus \{i\}$  we define the set

$$\mathcal{T}_{(i,n)}^{w,j} = \{v \in \mathcal{T}_{(i,n)} : |v| = |w|, v_{\{j\}^c} = w_{\{j\}^c}\}$$

and introduce the *measure of sensibility*

$$\Delta_{(i,n)}(j) = \max_{w \in \mathcal{T}_{(i,n)}} \max_{v \in \mathcal{T}_{(i,n)}^{w,j}} |\hat{p}_{(i,n)}(1|w) - \hat{p}_{(i,n)}(1|v)|.$$

Our interaction neighborhood estimator is defined as follows.

**Definition 1.** For any positive threshold parameter  $\varepsilon > 0$ , the estimated interaction neighborhood of neuron  $i \in F$ , at accuracy  $\varepsilon$ , given the sample  $X_1(F), \dots, X_n(F)$ , is defined as

$$\hat{V}_{(i,n)}^{(\varepsilon)} = \{j \in F \setminus \{i\} : \Delta_{(i,n)}(j) > \varepsilon\}. \tag{2.7}$$

### 2.3. Consistency of the selection procedure for finite and fully observed interaction neighborhoods

To ensure the existence of our process, we impose the following conditions.

**Assumption 1.** Suppose that

$$r = \sup_{i \in I} \sum_{j \in I} |W_{j \rightarrow i}| < \infty.$$

**Assumption 2.** Suppose that  $g(t) = \sup_{j \in I} g_j(t) < \infty$  for all  $t \geq 1$ .

We define the set  $\Omega^{\text{adm}}$  of *admissible pasts* as follows

$$\Omega^{\text{adm}} = \{x \in \{0, 1\}^{I \times \{\dots, -1, 0\}} : \forall i \in I, \exists \ell_i \leq 0 \text{ with } x_{\ell_i}(i) = 1\}. \quad (2.8)$$

Observe that if  $X_{-\infty}^0 = x \in \Omega^{\text{adm}}$ , then  $L_0^i > -\infty$  for all  $i \in I$ . Therefore, Assumptions 1 and 2 assure that for each  $i \in I$ ,

$$\sum_{j \in I} W_{j \rightarrow i} \sum_{s=L_0^i+1}^0 g_j(1-s) X_s(j) < \infty,$$

which implies that the transition probability  $\mathbb{P}(X_1(i) = 1 | X_{-\infty}^0 = x)$  is well-defined. By induction, for each  $t \geq 0$ , the transition probabilities (2.3) are also well-fined. Thus, the existence of the stochastic chain  $(X_t)_{t \in \mathbb{Z}}$ , starting from  $X_{-\infty}^0 = x \in \Omega^{\text{adm}}$ , follows immediately. Observe that we do not assume stationarity of the chain. To prove the consistency of our estimator we impose also

**Assumption 3.** For all  $i \in I$ ,  $\varphi_i \in C^1(\mathbb{R}, [0, 1])$  is a strictly increasing function. Moreover, there exists a  $p_* \in ]0, 1[$  such that for all  $i \in I$  and  $u \in \mathbb{R}$

$$p_* \leq \varphi_i(u) \leq 1 - p_*.$$

Define for  $i \in I$ ,

$$K_i = \left[ \sum_{j \in V_i^-} W_{j \rightarrow i} g_j(1), \sum_{j \in V_i^+} W_{j \rightarrow i} g_j(1) \right], \quad (2.9)$$

where  $V_i^+ = \{j \in V_i : W_{j \rightarrow i} > 0\}$  and  $V_i^- = \{j \in V_i : W_{j \rightarrow i} < 0\}$ .

Notice that under Assumptions 1 and 2, this interval is always bounded. Finally, we define

$$m_i = \inf_{u \in K_i} \{\varphi_i'(u)\} \inf_{j \in V_i} \{|W_{j \rightarrow i}| g_j(1)\}. \quad (2.10)$$

The following theorem is our first main result. It states the strong consistency of the interaction neighborhood estimator when  $V_i \subset F$ . By strong consistency we mean that the estimated interaction neighborhood of a fixed neuron  $i$  equals  $V_i$  eventually almost surely as  $n \rightarrow \infty$ .

**Theorem 1.** *Let  $F \subset I$  be a finite set and  $X_1(F), \dots, X_n(F)$  be a sample produced by a the stochastic chain  $(X_t)_{t \in \mathbb{Z}}$  compatible with (2.2) and (2.3), starting from  $X_{-\infty}^0 = x$  for some fixed  $x \in \Omega^{\text{adm}}$ . Under Assumptions 1–3, for any  $i \in F$  such that  $V_i \subset F$ , the following holds.*

1. (Overestimation). *For any  $j \in F \setminus V_i$ , we have that for any  $\varepsilon > 0$ ,*

$$\mathbb{P}(j \in \hat{V}_{(i,n)}^{(\varepsilon)}) \leq 4n^{3/2-\xi} \exp\left\{-\frac{\varepsilon^2 n^{2\xi}}{2}\right\}.$$

2. (Underestimation). *The quantity  $m_i$  defined in (2.10) satisfies  $m_i > 0$ , and for any  $j \in V_i$  and  $0 < \varepsilon < m_i$ ,*

$$\mathbb{P}(j \notin \hat{V}_{(i,n)}^{(\varepsilon)}) \leq 4 \exp\left\{-\frac{(m_i - \varepsilon)^2 n^{2\xi}}{2}\right\} + \exp\{-O(n^{1/2+\xi})\}.$$

3. *In particular, if we choose  $\varepsilon_n = O(n^{-\xi/2})$ , where  $\xi$  is the parameter appearing in (2.6), then*

$$\hat{V}_{(i,n)}^{(\varepsilon_n)} = V_i \quad \text{eventually almost surely.}$$

The proof of Theorem 1 is given in Section 4.

### 2.4. Extension to the case of partially observed interaction neighborhoods

We now discuss the case when  $V_i$  is not fully included in the sampling region  $F$ , in particular, the case when  $V_i$  is infinite. In this case, we also impose the following assumptions.

**Assumption 4.**  $\gamma = \sup_{j \in I} \|\varphi'_j\|_\infty < \infty$ .

**Assumption 5.** There exists a constant  $C$  and  $p \geq 1$ , such that  $g(t) = \sup_{j \in I} g_j(t) \leq C(1 + t^p)$  for all  $t \geq 1$ .

Let  $\ell_\infty = \{\xi = (\xi_j)_{j \in I} : \forall j \in I, \xi_j \in \mathbb{R} \text{ and } \|\xi\|_\infty := \sup_{j \in I} |\xi_j| < \infty\}$  be the space of all bounded series of real numbers indexed by  $I$ . Under Assumption 5, we may introduce, for each  $t \geq 1$ , the continuous operator  $H(t) : \ell_\infty \rightarrow \ell_\infty$  defined by  $(H(t)\xi)_j = \sum_{k \in I} H_{j,k}(t)\xi_k$ , for all  $j \in I$ , where

$$H_{j,k}(t) := \gamma |W_{k \rightarrow j}| g_k(t), \quad \text{for } j \neq i, \tag{2.11}$$

and, for  $p_*$  as in Assumption 3,

$$H_{i,i}(t) := (1 - p_*) 1_{\{t=1\}}.$$



By our assumptions, the norm of the operator  $H(t)$  defined by

$$\|H(t)\| = \sup\{\|H(t)\xi\|_\infty : \xi \in \ell_\infty, \|\xi\|_\infty = 1\}$$

satisfies

$$\|H(t)\| \leq C\gamma r(1 + t^p) + (1 - p_*)1_{\{t=1\}}.$$

Then for any  $\alpha > 0$ , the linear operator

$$\Lambda(\alpha) = \sum_{t=1}^\infty e^{-\alpha t} H(t)$$

is well-defined and continuous as well. In particular, there exists  $\alpha_0 \geq 0$  such that

$$\|\Lambda(\alpha_0)\| < 1. \tag{2.12}$$

We are now ready to state our second main result. It gives precise error bounds for the interaction neighborhood estimator when  $V_i$  is not fully observed. These error bounds depend on the tail of the series

$$\Sigma_i(F) := \sum_{j \notin V_i \cap F} |W_{j \rightarrow i}|. \tag{2.13}$$

To state the theorem, we shall also need the definitions

$$K_i^{[F]} = \left[ \sum_{j \in V_i^- \cap F} W_{j \rightarrow i} g_j(1), \sum_{j \in V_i^+ \cap F} W_{j \rightarrow i} g_j(1) \right]$$

and

$$m_i^{[F]} = \inf_{u \in K_i^{[F]}} \{\varphi'_i(u)\} \inf_{j \in V_i \cap F} \{|W_{j \rightarrow i}| g_j(1)\}.$$

**Theorem 2.** *Let  $F \subset I$  be a finite set and  $X_1(F), \dots, X_n(F)$  be a sample produced by a the stochastic chain  $(X_t)_{t \in \mathbb{Z}}$  compatible with (2.2) and (2.3), starting from  $X_{-\infty}^0 = x$  for some fixed  $x \in \Omega^{\text{adm}}$ . Under Assumptions 1–5, for any  $i \in F$  such that  $V_i \cap F \neq \emptyset$ , the following assertions hold true.*

1. (Overestimation). *For any  $j \in F \setminus V_i$ , we have that for any  $\varepsilon > 0$ ,*

$$\mathbb{P}(j \in \hat{V}_{(i,n)}^{(\varepsilon)}) \leq 4n^{3/2-\xi} \exp\left\{-\frac{\varepsilon^2 n^{2\xi}}{2}\right\} + C(e^{\alpha_0 n} \vee n) \Sigma_i(F).$$

2. (Underestimation). *We have that  $m_i^{[F]} > 0$ , and for any  $j \in V_i \cap F$  and  $0 < \varepsilon < m_i^{[F]}$ ,*

$$\mathbb{P}(j \notin \hat{V}_{(i,n)}^{(\varepsilon)}) \leq 4 \exp\left\{-\frac{(m_i^{[F]} - \varepsilon)^2 n^{2\xi}}{2}\right\} + \exp\{-O(n^{1/2+\xi})\} + C(e^{\alpha_0 n} \vee n) \Sigma_i(F).$$

The proof of Theorem 2 is given in Section 5.

### 3. Exponential inequalities

To prove Theorems 1 and 2, we need some exponential inequalities, including a new conditional Hoeffding-type inequality, stated in Proposition 1 below, which is interesting by itself.

For each  $\ell \geq 1$ ,  $F \subset I$  finite and  $w \in \{0, 1\}^{\{-\ell, \dots, -1\} \times F \setminus \{i\}}$ , we write

$$p_i(1|w) = \mathbb{P}(X_{\ell+2}(i) = 1 | X_1^{\ell+1}(i) = 10^\ell, X_2^{\ell+1}(F \setminus \{i\}) = w). \tag{3.1}$$

Notice that  $p_i(1|w) = p_i(1|w(V_i))$  for any set  $F \supset V_i$ , where  $w(V_i)$  is the configuration  $w$  restricted to the set  $V_i$ . Moreover, the homogeneity of the transition probability (2.3) implies that, whenever  $F \supset V_i$ , for any  $t \geq \ell + 2$ ,

$$p_i(1|w) = \mathbb{P}(X_t(i) = 1 | X_{t-\ell-1}^{t-1}(i) = 10^\ell, X_{t-\ell}^{t-1}(F \setminus \{i\}) = w).$$

**Proposition 1.** *Suppose  $V_i$  finite and  $V_i \subset F$ . Then for any  $\ell \geq 1$ ,  $w \in \{0, 1\}^{\{-\ell, \dots, -1\} \times F \setminus \{i\}}$ ,  $\lambda > 0$  and all  $t > \ell + 1$ ,*

$$\mathbb{P}(|M_{(i,t)}(w)| > \lambda) \leq 2 \exp\left\{-\frac{2\lambda^2}{t - \ell + 1}\right\} \mathbb{P}(N_{(i,t)}(w) > 0), \tag{3.2}$$

where  $M_{(i,t)}(w) := N_{(i,t)}(w, 1) - p_i(1|w)N_{(i,t)}(w)$ .

**Proof.** We denote  $p = p_i(1|w)$  and for each  $t \geq \ell + 1$ ,  $N_{(i,t)}(w) = N_t$ ,  $Y_t = \mathbf{1}\{X_t(i) = 1\} - p$ ,  $\chi_t = \mathbf{1}\{X_{t-\ell}^{t-1}(F \setminus \{i\}) = w, X_{t-\ell-1}^{t-1}(i) = 10^\ell\}$  and also  $M_{(i,t)}(w) = M_t$  with the convention that  $M_{\ell+1} = 0$ . Thus for  $t \geq \ell + 2$ ,

$$M_t = M_{t-1} + \chi_t Y_t. \tag{3.3}$$

Since  $\mathbb{P}(M_t > \lambda) = \mathbb{P}(M_t > \lambda, N_t > 0)$ , the Markov inequality implies that

$$\mathbb{P}(M_t > \lambda) \leq e^{-\lambda\sigma} E[e^{\sigma M_t} \mathbf{1}\{N_t > 0\}],$$

for all  $\sigma > 0$ . Notice that  $\{N_t > 0\} = \{N_{t-1} > 0\} \cup \{N_{t-1} = 0, \chi_t = 1\}$ , so that by (3.3), it follows that  $\mathbb{E}[e^{\sigma M_t} \mathbf{1}\{N_t > 0\}]$  can be rewritten as

$$\mathbb{E}[e^{\sigma M_{t-1}} e^{\sigma \chi_t Y_t} \mathbf{1}\{N_{t-1} > 0\}] + \mathbb{E}[e^{\sigma Y_t} \mathbf{1}\{N_{t-1} = 0, \chi_t = 1\}]. \tag{3.4}$$

From the assumption  $V_i \subset F$  it follows that  $p = p_i(1|w) = p_i(1|w(V_i))$  and  $\mathbb{E}[\chi_t Y_t | \mathcal{F}_{t-1}] = 0$ . Since  $-p \leq \chi_t Y_t \leq 1 - p$ , the classical Hoeffding bound implies that  $\mathbb{E}[e^{\sigma \chi_t Y_t} | \mathcal{F}_{t-1}] \leq e^{\sigma^2/8}$  and therefore the expression (3.4) can be bounded above by

$$\mathbb{E}[e^{\sigma M_t} \mathbf{1}\{N_t > 0\}] \leq e^{\sigma^2/8} \mathbb{E}[e^{\sigma M_{t-1}} \mathbf{1}\{N_{t-1} > 0\}] + e^{\sigma^2/8} \mathbb{E}[\mathbf{1}\{N_{t-1} = 0, \chi_t = 1\}].$$

By iterating the inequality above and using the identity

$$\mathbf{1}\{N_t > 0\} = \mathbf{1}\{N_{\ell+1} > 0\} + \sum_{s=\ell+2}^t \mathbf{1}\{N_{s-1} = 0, \chi_s = 1\},$$

we obtain that  $\mathbb{E}[e^{\sigma M_t} \mathbf{1}\{N_t > 0\}] \leq e^{(t-\ell+1)\sigma^2/8} \mathbb{P}(N_t > 0)$ . Thus, collecting all these estimates, we deduce, by taking  $\sigma = 4\lambda(t - \ell + 1)^{-1}$ , that

$$\mathbb{P}(M_t > \lambda) \leq \exp\left\{-\frac{2\lambda^2}{t - \ell + 1}\right\} \mathbb{P}(N_t > 0).$$

The left-tail probability  $\mathbb{P}(M_t < -\lambda)$  is treated likewise. □

As a consequence of Proposition 1, we have the following result.

**Proposition 2.** *Suppose that  $V_i$  is finite and  $V_i \subset F$ . Then for any  $\ell \geq 1$ ,  $t > \ell + 1$ ,  $w \in \{0, 1\}^{\{-\ell, \dots, -1\} \times F \setminus \{i\}}$ ,  $\xi \in (0, 1/2)$  and  $\varepsilon > 0$ , we have*

$$\mathbb{P}(|\hat{p}_{(i,t)}(1|w) - p_i(1|w)| > \varepsilon, N_{(i,t)}(w) \geq t^{1/2+\xi}) \leq 2 \exp\{-2\varepsilon^2 t^{2\xi}\} \mathbb{P}(N_{(i,t)}(w) > 0).$$

**Proof.** For any  $w \in \{0, 1\}^{\{-\ell, \dots, -1\} \times F \setminus \{i\}}$  and  $t > \ell + 1$ ,

$$M_t(w) = (\hat{p}_{(i,t)}(1|w) - p_i(1|w)) N_{(i,t)}(w),$$

so that

$$\mathbb{P}(|\hat{p}_{(i,t)}(1|w) - p_i(1|w)| > \varepsilon, N_{(i,t)}(w) \geq t^{1/2+\xi}) \leq \mathbb{P}(|M_t(w)| > \varepsilon t^{1/2+\xi}).$$

Thus the result follows from Proposition 1 by taking  $\lambda = \varepsilon t^{1/2+\xi}$ . □

The next two results will be used to control the probability of underestimating  $V_i$ . We start with a simple lower bound which follows immediately from Assumption 3.

**Lemma 1.** *For any fixed  $i \in F$ ,  $t > \ell + 1$  and  $w \in \{0, 1\}^{\{-\ell, \dots, -1\} \times F \setminus \{i\}}$ , we define for  $1 \leq s \leq t - \ell$ ,*

$$Z_s = \mathbf{1}\{X_s^{s+\ell}(i) = 10^\ell, X_{s+1}^{s+\ell}(F \setminus \{i\}) = w\}.$$

Under Assumption 3, it follows that

$$\mathbb{P}(Z_s = 1 | \mathcal{F}_{s-1}) \geq p_{\min}^{|F| \ell + 1},$$

where  $p_{\min} = \min\{p_*, (1 - p_*)\} > 0$  with  $p_*$  as in Assumption 3.

**Lemma 2.** *Suppose Assumption 3. For any  $\xi \in (0, 1/2)$ ,  $i \in F$  and  $w \in \{0, 1\}^{\{-1\} \times F \setminus \{i\}}$ , it holds that*

$$\mathbb{P}(N_{(i,t)}(w) < t^{1/2+\xi}) \leq \exp\{-O(t^{1/2+\xi})\}.$$

**Proof.** For each  $1 \leq s \leq t - 1$ , let  $Z_s$  be the random variable defined as in Lemma 1 with  $\ell = 1$ . Now we define  $Y_s = Z_{2(s-1)+1}$  for  $1 \leq s \leq \lfloor t/2 \rfloor$  and observe that  $\mathcal{G}_s := \sigma(Y_1, \dots, Y_s) \subset \mathcal{F}_{2s}$ . Thus, by Lemma 1,

$$\mathbb{P}(Y_s = 1 | \mathcal{G}_{s-1}) = \mathbb{E}[\mathbb{P}(Z_{2(s-1)+1} = 1 | \mathcal{F}_{2(s-1)}) | \mathcal{G}_{s-1}] \geq p_{\min}^{|F|+1}.$$

Define  $q_* = p_{\min}^{|F|+1}$ . Then Lemma A.3 of [9] implies for every  $v \in ]0, 1[$ ,

$$\mathbb{P}\left(\frac{1}{\lfloor t/2 \rfloor} \sum_{s=1}^{\lfloor t/2 \rfloor} Y_s < vq_*\right) \leq \exp\left\{-\lfloor t/2 \rfloor \frac{q_*}{4} (1-v)^2\right\}.$$

Clearly  $N_{(i,t)}(w) = \sum_{s=1}^{t-1} Z_s \geq \sum_{s=1}^{\lfloor t/2 \rfloor} Y_s$ , so that it follows from the inequality above that

$$\mathbb{P}(N_{(i,t)}(w) < vq_* \lfloor t/2 \rfloor) \leq \exp\left\{-\lfloor t/2 \rfloor \frac{q_*}{4} (1-v)^2\right\}.$$

Finally, for any fixed  $v \in (0, 1)$  and all  $t$  large enough,  $\lfloor t/2 \rfloor \frac{q_*}{4} (1-v)^2 > t^{1/2+\xi}$  and  $vq_* \lfloor t/2 \rfloor > t^{1/2+\xi}$ , implying the assertion. □

### 4. Proof of Theorem 1

Suppose that  $V_i \subset F$  and notice that for any  $\ell \geq 1$  and  $w \in \{0, 1\}^{\{-\ell, \dots, -1\} \times V_i}$ , it holds that

$$p_i(1|w) = \varphi_i \left( \sum_{j \in V_i} W_{j \rightarrow i} \sum_{t=-\ell}^{-1} g_j(-t) w_t(j) \right). \tag{4.1}$$

**Proof of Item 1 of Theorem 1.** Using the definition of  $\hat{V}_{(i,n)}^{(\varepsilon)}$  and applying the union bound, we deduce that

$$\begin{aligned} \mathbb{P}(j \in \hat{V}_{(i,n)}^{(\varepsilon)}) &= \mathbb{P}(\Delta_{(i,n)}(j) > \varepsilon) \\ &\leq \mathbb{E} \left[ \sum_{(w,v)} \mathbf{1}\{A_{(i,n)}^{w,v,j}, |\hat{p}_{(i,n)}(1|w) - \hat{p}_{(i,n)}(1|v)| > \varepsilon\} \right], \end{aligned} \tag{4.2}$$

where  $A_{(i,n)}^{w,v,j} := \{(w, v) \in \mathcal{T}_{(i,n)} \times \mathcal{T}_{(i,n)}^{w,j}\}$ . Since  $j \notin V_i$  and  $V_i \subset F$ , the configurations of any pair  $(w, v) \in \mathcal{T}_{(i,n)} \times \mathcal{T}_{(i,n)}^{w,j}$  coincide in restriction to the set  $V_i$ . In other words,  $w(V_i) = v(V_i)$ . In particular, it follows from (4.1) that  $p_i(1|w) = p_i(1|w(V_i)) = p_i(1|v(V_i)) = p_i(1|v)$ .

Therefore, applying the triangle inequality, it follows that on  $A_{(i,n)}^{w,v,j}$ ,

$$\mathbf{1}\{|\hat{p}_{(i,n)}(1|w) - \hat{p}_{(i,n)}(1|v)| > \varepsilon\} \leq \sum_{u \in \{w,v\}} \mathbf{1}\{|\hat{p}_{(i,n)}(1|u) - p_i(1|u)| > \varepsilon/2\},$$

so that the expectation in (4.2) can be bounded above by

$$2\mathbb{E}\left[\sum_w \mathbf{1}_{\{w \in \mathcal{T}_{(i,n)}\}} \sum_v \mathbf{1}_{\{v \in \mathcal{T}_{(i,n)}\}} \mathbf{1}\{|\hat{p}_{(i,n)}(1|v) - p_i(1|v)| > \varepsilon/2\}\right]. \quad (4.3)$$

Now, since  $\sum_w N_{(i,n)}(w) \leq n$ , we have that

$$n \geq \sum_{w: N_{(i,n)}(w) \geq n^{1/2+\xi}} N_{(i,n)}(w) \geq n^{1/2+\xi} |\{w : N_{(i,n)}(w) \geq n^{1/2+\xi}\}|,$$

which implies that  $|\mathcal{T}_{(i,n)}| \leq n^{1/2-\xi}$ . From this last inequality and Proposition 2, which is stated in Section 3 below, we obtain the following upper bound for (4.3),

$$4n^{1/2-\xi} \exp\left\{-\frac{\varepsilon^2 n^{2\xi}}{2}\right\} \mathbb{E}\left[\sum_w \mathbf{1}\{N_{(i,n)}(w) > 0\}\right]. \quad (4.4)$$

Since  $\sum_w \mathbf{1}\{N_{(i,n)}(w) > 0\} \leq n$ , the result follows from inequalities (4.2) and (4.4).  $\square$

Before proving Item 2 of Theorem 1, we will prove the following lemma.

**Lemma 3.** *Suppose that  $V_i$  is finite and define for each  $j \in V_i$ ,*

$$m_{i,j} := \max_{w,v \in \{0,1\}^{\{-1\} \times V_i} : w_{\{j\}^c} = v_{\{j\}^c}} |p_i(1|w) - p_i(1|v)|.$$

*Then, under Assumption 3, we have that*

$$\inf_{j \in V_i} m_{i,j} \geq \inf_{x \in K_i} \{\varphi'_i(x)\} \inf_{j \in V_i} \{|W_{j \rightarrow i}| g_j(1)\} = m_i > 0, \quad (4.5)$$

*where  $K_i$  is defined in (2.9).*

**Proof.** For each  $j \in V_i$  take any pair  $w, v \in \{0,1\}^{\{-1\} \times V_i}$  such that  $w_{\{j\}^c} = v_{\{j\}^c}$  with  $w_{-1}(j) = 1$  and  $v_{-1}(j) = 0$ . By Assumption 3, the function  $\varphi_i$  is differentiable such that, for  $\ell = 1$ ,

$$|p_i(1|w) - p_i(1|v)| \geq \inf_{x \in K_i} \{\varphi'_i(x)\} \left| \sum_{k \in V_i} W_{k \rightarrow i} \sum_{t=\ell}^{-1} g_k(-t)(w_t(k) - v_t(k)) \right|.$$

Since  $|\sum_{k \in V_i} W_{k \rightarrow i} \sum_{t=-\ell}^{-1} g_k(-t)(w_t(k) - v_t(k))| = |W_{j \rightarrow i}| g_j(1)$ , the inequality above implies the first assertion of the lemma.

By Assumption 3, the function  $\varphi_i$  is strictly increasing ensuring that  $\inf_{x \in K_i} \{\varphi'_i(x)\} > 0$ . Thus, since for all  $j \in I$  the sequence  $g_j$  is strictly positive and  $V_i \neq \emptyset$  is finite, we clearly have that  $m_i > 0$ . □

We are now in position to conclude the proof of Theorem 1.

**Proof of Item 2 of Theorem 1.** Lemma 3 implies that  $m_i$  defined in (4.5) is positive. Let  $0 < \varepsilon < m_i$ . If  $j \in V_i$ , Lemma 3 implies the existence of strings  $w^*, v^* \in \{0, 1\}^{\{-1\} \times F \setminus \{i\}}$  such that  $w^*_{[j]^c} = v^*_{[j]^c}$  and

$$|p_i(1|w^*) - p_i(1|v^*)| = |p_i(1|w^*(V_i)) - p_i(1|v^*(V_i))| \geq m_i.$$

Denoting by  $C_n = \{N_{(i,n)}(w^*) \geq n^{\xi+1/2}, N_{(i,n)}(v^*) \geq n^{\xi+1/2}\}$  it follows that

$$\mathbb{P}(j \notin \hat{V}_{(i,n)}^{(\varepsilon)}) \leq \mathbb{P}(|\hat{p}_{(i,n)}(1|w^*) - \hat{p}_{(i,n)}(1|v^*)| < \varepsilon, C_n) + \mathbb{P}(C_n^c). \tag{4.6}$$

Now notice that the first term on the right in (4.6) is upper bounded by

$$\sum_{u \in \{w^*, v^*\}} \mathbb{P}(|\hat{p}_{(i,n)}(1|u) - p_i(1|u)| > (m_i - \varepsilon)/2, N_{(i,n)}(u) \geq n^{\xi+1/2}),$$

and since  $m_i > \varepsilon$ , the result follows from Proposition 2 and Lemma 2, both stated in Section 3 above. □

**Proof of Item 3 of Theorem 1.** Define for  $n \in \mathbb{N}$  the sets

$$O_n = \{j \in F \setminus V_i : j \in V_{(i,n)}^{(\varepsilon_n)}\} \quad \text{and} \quad U_n = \{j \in V_i : j \notin V_{(i,n)}^{(\varepsilon_n)}\}.$$

Applying the union bound and then Item 1, we infer that

$$\mathbb{P}(O_n) \leq 4(|F| - |V_i|)n^{3/2-\xi} \exp\left\{-\frac{\varepsilon^2 n^{2\xi}}{2}\right\}.$$

Applying once more the union bound and then using Item 2, we also infer that

$$\mathbb{P}(U_n) \leq |V_i| \left( 4 \exp\left\{-\frac{(m_i - \varepsilon_n)^2 n^{2\xi}}{2}\right\} + \exp\{-O(n^{1/2+\xi})\} \right).$$

Since  $\{V_{(i,n)}^{(\varepsilon)} \neq V_i\} = O_n \cup U_n$ , we deduce that  $\sum_{n=1}^{\infty} \mathbb{P}(V_{(i,n)}^{(\varepsilon_n)} \neq V_i) < \infty$ , so that the result follows from the Borel–Cantelli lemma. □

## 5. Proof of Theorem 2

To deal with the case  $V_i \not\subset F$ , we couple the process  $X = (X_t)_{t \in \mathbb{Z}}$  with its fixed range approximation  $X^{[F]} = (X_t^{[F]})_{t \in \mathbb{Z}}$ , where  $X^{[F]}$  follows the same dynamics as  $X$ , defined in (2.2) and

(2.3) for all  $j \neq i$ , except that (2.3) is replaced – for the fixed neuron  $i$  – by

$$\mathbb{P}(X_{t+1}^{[F]}(i) = 1 | \mathcal{F}_t) = \varphi_i \left( \sum_{j \in V_i \cap F} W_{j \rightarrow i} \sum_{s=L_t^{i,[F]}+1}^t g_j(t+1-s) X_s^{[F]}(j) \right). \tag{5.1}$$

Also, suppose that  $X$  and  $X^{[F]}$  start from the same initial configuration  $X_{-\infty}^0 = (X^{[F]})_{-\infty}^0 = x$ , where  $x \in \Omega^{\text{adm}}$ . We will show in Proposition 4 in the Appendix that Assumptions 1, 4 and 5 imply the existence of a coupling between  $X$  and  $X^{[F]}$  and of a constant  $C > 0$  such that

$$\sup_{j \in I} \mathbb{P}(\exists t \in [1, n] : X_t(j) \neq X_t^{[F]}(j)) \leq C(e^{\alpha_0 n} \vee n) \sum_{j \notin V_i \cap F} |W_{j \rightarrow i}|. \tag{5.2}$$

Write

$$E_n = \bigcap_{1 \leq s \leq n} \{X_s(i) = X_s^{[F]}(i)\}.$$

On  $E_n$ , instead of working with  $X_s(i)$ ,  $1 \leq s \leq n$ , we can therefore work with its approximation  $X_s^{[F]}(i)$ ,  $1 \leq s \leq n$ , having conditional transition probabilities (for neuron  $i$ ) given by

$$p_i^{[F]}(1|w) = \varphi_i \left( \sum_{j \in V_i \cap F} W_{j \rightarrow i} \sum_{s=-\ell}^{-1} g_j(-s) w_j(s) \right)$$

which only depend on  $w(V_i \cap F)$ . As a consequence, on  $E_n$  the proof of Theorem 2 works as in the preceding section, except that we replace  $m_i$  by

$$m_i^{[F]} = \inf_{x \in K_i^{[F]}} \{\varphi'_i(x)\} \inf_{j \in V_i \cap F} \{|W_{j \rightarrow i}| g_j(1)\} > 0$$

if  $V_i \cap F \neq \emptyset$ . Here  $K_i^{[F]}$  is defined by

$$K_i^{[F]} = \left[ \sum_{j \in V_i^- \cap F} W_{j \rightarrow i} g_j(1), \sum_{j \in V_i^+ \cap F} W_{j \rightarrow i} g_j(1) \right].$$

Finally, writing

$$O_n = \{j \in F \setminus V_i : j \in V_{(i,n)}^{(\varepsilon_n)}\} \quad \text{and} \quad U_n = \{j \in V_i \cap F : j \notin V_{(i,n)}^{(\varepsilon_n)}\},$$

we obtain

$$\mathbb{P}(O_n) \leq \mathbb{P}(O_n \cap E_n) + \mathbb{P}(E_n^c), \quad \mathbb{P}(U_n) \leq \mathbb{P}(U_n \cap E_n) + \mathbb{P}(E_n^c),$$

where as before in Theorem 1,

$$\mathbb{P}(O_n \cap E_n) \leq 4n^{3/2-\xi} \exp \left\{ -\frac{\varepsilon_n^2 n^{2\xi}}{2} \right\}$$

and

$$\mathbb{P}(U_n \cap E_n) \leq |F| \left( 4 \exp \left\{ -\frac{(m_i^{[F]} - \varepsilon_n)^2 n^{2\xi}}{2} \right\} + \exp \{ -O(n^{1/2+\xi}) \} \right).$$

Finally, by inequality (5.2),

$$\mathbb{P}(E_n^c) = \mathbb{P}(\exists t \in [1, n] : X_t(i) \neq X_t^{[F]}(i)) \leq C(e^{\alpha_0 n} \vee n) \sum_{j \notin V_i \cap F} |W_{j \rightarrow i}|,$$

for some constant  $C$ . This concludes the proof.

### 6. Time complexity of the estimation procedure

The time complexity of our selection procedure has quadratic growth with respect to  $n$ , the length of the time interval during which the neural network is observed. This is the content of the following proposition.

**Proposition 3.** *The number of operations needed to compute the set  $\hat{V}_{(i,n)}^{(\varepsilon)}$  is  $O(n^2)$ .*

**Proof.** All the random variables involved in the definition of the random set  $\hat{V}_{(i,n)}^{(\varepsilon)}$  can be written in terms of the counting variables  $N_{(i,n)}(\cdot)$ . Besides, all counting variables  $N_{(i,n)}(w)$  for  $w \in \{0, 1\}^{\{-\ell, \dots, -1\} \times F \setminus \{i\}}$  with fixed length  $\ell$  can be computed simultaneously after  $n - \ell - 1$  operations. Indeed, we set initially  $N_{(i,n)}(w) = 0$  for all pasts  $w \in \{0, 1\}^{\{-\ell, \dots, -1\} \times F \setminus \{i\}}$  and then we increment by 1 the count of the past that has occurred at time  $\ell + 2 \leq t \leq n$ , leaving the counts of all other pasts unchanged. Thus, we compute all the counting variables  $N_{(i,n)}(w)$  with

$$\sum_{l=1}^{\lfloor n-2-n^{1/2+\xi} \rfloor} (n - l - 1) \leq O(n^2 - n^{1+2\xi})$$

operations, for all local pasts  $w \in \{0, 1\}^{\{-\ell, \dots, -1\} \times F \setminus \{i\}}$ , for all  $1 \leq \ell \leq \lfloor n - 2 - n^{1/2+\xi} \rfloor$ , where for each  $x \in \mathbb{R}$ ,  $\lfloor x \rfloor$  is the largest integer less than or equal to  $x$ .

Now, given all counting variables, to compute  $\Delta_{(i,n)}(j)$  we need at most  $O(\sum_{w \in \mathcal{T}_{(i,n)}} |\mathcal{T}_{(i,n)}^{w,j}|) \leq O(n^{1-2\xi})$  computations which in turns implies that, given all counting variables, with at most  $O(|F|n^{1-2\xi})$  operations we compute our estimator  $\hat{V}_{(i,n)}^{(\varepsilon)}$ . Therefore, in the overall, we need to perform at most

$$O(n^2 - n^{1+2\xi}) + O(|F|n^{1-2\xi}) = O(n^2)$$

operations. □



## Appendix: Auxiliary results

In this section, we prove the coupling result (5.2) needed in the proof of Theorem 2. For that sake, let  $F \subset I$  be a finite set, fix  $i \in F$  and let  $U_t(j)$ ,  $j \in I$ ,  $t \geq 1$ , be an i.i.d. family of random variables uniformly distributed on  $[0, 1]$ .

The coupling is defined as follows. For any  $x \in \Omega^{\text{adm}}$ , we define  $X_t(j) = X_t^{[F]}(j) = x_t(j)$  for each  $t \leq 0$  and  $j \in I$ . For each  $t \geq 1$  and  $j \in I$ , we define

$$X_t(j) = \begin{cases} 1, & \text{if } U_t(j) > \varphi_j(\eta_{t-1}(j)), \\ 0, & \text{if } U_t(j) \leq \varphi_j(\eta_{t-1}(j)) \end{cases}$$

and

$$X_t^{[F]}(j) = \begin{cases} 1, & \text{if } U_t(j) > \varphi_j(\eta_{t-1}^{[F]}(j)), \\ 0, & \text{if } U_t(j) \leq \varphi_j(\eta_{t-1}^{[F]}(j)), \end{cases}$$

where for each  $t \geq 0$  and  $j \in I$ ,

$$\eta_t(j) = \sum_{k \in V_j} W_{k \rightarrow j} \sum_{s=L_j^j+1}^t g_k(t+1-s) X_s(k)$$

and, if  $j \neq i$ ,

$$\eta_t^{[F]}(j) = \sum_{k \in V_j} W_{k \rightarrow j} \sum_{s=L_j^{[F]}+1}^t g_k(t+1-s) X_s^{[F]}(k), \quad (\text{A.1})$$

and finally

$$\eta_t^{[F]}(i) = \sum_{k \in V_i \cap F} W_{k \rightarrow i} \sum_{s=L_i^{[F]}+1}^t g_k(t+1-s) X_s^{[F]}(k). \quad (\text{A.2})$$

In other words, the process  $X^{[F]}$  has exactly the same dynamics as the original process  $X$ , except that neuron  $i$  depends only on neurons belonging to  $V_i \cap F$ . Notice that we use the same uniform random variables  $U_t(j)$  to update the values of  $X_t(j)$  and of  $X_t^{[F]}(j)$ . In this way, we achieve a coupling between the two processes. We shall write  $\mathbb{E}_x$  to denote the expectation with respect to this coupling. Then we have the following result.

**Proposition 4.** *Assume Assumptions 1, 4 and 5, and let  $\alpha_0$  be defined as in (2.12).*

1. *If  $\alpha_0 > 0$ , then*

$$\sup_{j \in I} \mathbb{P}_x \left( \bigcup_{s=1}^t \{X_s(j) \neq X_s^{[F]}(j)\} \right) \leq C e^{\alpha_0 t} \sum_{k \notin V_i \cap F} |W_{k \rightarrow i}|. \quad (\text{A.3})$$

2. Suppose now that  $\alpha_0 = 0$  and write for any  $j \in I$ ,  $\varrho_j = \sum_{t=1}^\infty g_j(t)$ ,  $\varrho = \sup_{j \in I} \varrho_j$ . Then

$$\chi = (1 - p_*) + \gamma \sup_{j \in I} \sum_{k \in I} \varrho_k |W_{k \rightarrow j}| < 1, \tag{A.4}$$

and in this case

$$\sup_{j \in I} \mathbb{P}_x \left( \bigcup_{s=1}^t \{X_s(j) \neq X_s^{[F]}(j)\} \right) \leq \frac{\gamma \varrho^t}{1 - \chi} \sum_{k \notin V_i \cap F} |W_{k \rightarrow i}|. \tag{A.5}$$

**Proof.** For notational convenience, we assume that the starting configuration  $x \in \Omega^{\text{adm}}$  satisfies  $x_0(i) = 1$  and extend the definition of  $g_j$  by defining  $g_j(t) = 0$  for all  $t \leq 0$  and  $j \in I$ .

We start proving Item 1. Recall the definition of the continuous operator  $H(t)$  in (2.11). In the sequel, we set also  $H(0) \equiv 0$ .

Let for each  $t \geq 0$ ,

$$D_j(t) = 1 \{L_t^j \neq L_t^{j,[F]}\}, \quad j \in I,$$

and observe that

$$P_x(X_t(j) \neq X_t^{[F]}(j)) \leq E_x[D_j(t)]. \tag{A.6}$$

Given  $\mathcal{F}_t$ , we update  $D_j(t)$  as follows. If neuron  $j$  spikes at time  $t + 1$  in both processes, then  $D_j(t + 1) = 0$  regardless the value of  $D_j(t)$ . By the definition of the coupling, this event occurs with probability  $\varphi_j(\eta_t(j) \wedge \eta_t^{[F]}(j)) \geq p_*$ . When  $D_j(t) = 1$ , then  $D_j(t + 1) = 1$  if and only if neuron  $j$  does not spike in both processes. Clearly, this event has probability  $1 - \varphi_j(\eta_t(j) \wedge \eta_t^{[F]}(j))$ . Finally, if  $D_j(t) = 0$ , then  $D_j(t + 1) = 1$  if and only if neuron  $j$  spikes only in one of the two processes. This event occurs with probability  $|\varphi_j(\eta_t(j)) - \varphi_j(\eta_t^{[F]}(j))|$ . Thus for all  $j \in I$ , we have

$$\begin{aligned} \mathbb{E}_x(D_j(t + 1) | \mathcal{F}_t) &= D_j(t)(1 - \varphi_j(\eta_t(j) \wedge \eta_t^{[F]}(j))) \\ &\quad + |\varphi_j(\eta_t(j)) - \varphi_j(\eta_t^{[F]}(j))|(1 - D_j(t)). \end{aligned} \tag{A.7}$$

Since  $\varphi_i$  is Lipschitz with Lipschitz constant  $\gamma$  and  $L_t^i = L_t^{i,[F]}$  on  $\{D_i(t) = 0\}$ , we have on this event,

$$\begin{aligned} &\frac{1}{\gamma} |\varphi_i(\eta_t(i)) - \varphi_i(\eta_t^{[F]}(i))| \\ &\leq |\eta_t(i) - \eta_t^{[F]}(i)| \\ &\leq \sum_{k \in V_i \cap F} |W_{k \rightarrow i}| \sum_{s=L_t^i+1}^t g_k(t + 1 - s) |X_s(k) - X_s^{[F]}(k)| \end{aligned} \tag{A.8}$$

$$\begin{aligned}
 & + \sum_{k \notin V_i \cap F} |W_{k \rightarrow i}| \sum_{s=1}^t g_k(t+1-s) \\
 & \leq \sum_{k \in V_i \cap F} |W_{k \rightarrow i}| \sum_{s=1}^{t+1} g_k(t+1-s) |X_s(k) - X_s^{[F]}(k)| \\
 & \quad + \sum_{k \notin V_i \cap F} |W_{k \rightarrow i}| \sum_{s=1}^{t+1} g_k(t+1-s),
 \end{aligned}$$

where we have used that  $g_k(0) = 0$  in order to replace the sum  $\sum_{s=1}^t$  by  $\sum_{s=1}^{t+1}$ . Moreover, we have used that  $L_t^i = L_t^{i,[F]} \geq 0$  for all  $t \geq 0$ , by our choice of  $x$ .

Similarly, for all  $j \neq i$ , we have on  $\{D_j(t) = 0\}$ ,

$$\frac{1}{\gamma} |\varphi_j(\eta_t(j)) - \varphi_j(\eta_t^{[F]}(j))| \leq \sum_{k \in V_j} |W_{k \rightarrow j}| \sum_{s=1}^{t+1} g_k(t+1-s) |X_s(k) - X_s^{[F]}(k)|. \tag{A.9}$$

For each  $j \in I$ , let  $\delta_j(t) = \mathbb{E}_x(D_j(t))$  and write  $\delta(t) = (\delta_j(t))_{j \in I}$  for the associated column vector. Taking expectation in (A.7)–(A.9) and using that  $\mathbb{E}_x |X_s(k) - X_s^{[F]}(k)| \leq \delta_k(s)$  (see (A.6)), we obtain

$$\delta(t+1) \leq H * \delta(t+1) + \gamma \Sigma_i(F) g * 1(t+1) e_i, \tag{A.10}$$

where  $e_i$  is the  $i$ -unit vector. In the above formula,

$$(H * \delta(t))_j = \sum_{k \in I} \sum_{s=0}^t H_{j,k}(t-s) \delta_k(s)$$

is the operator convolution product, and the inequality in (A.10) has to be understood coordinate-wise.

Now let  $\alpha_0$  be as in (2.12) and introduce  $\tilde{H}(t) = e^{-\alpha_0 t} H(t)$ ,  $\tilde{\delta}(t) = e^{-\alpha_0 t} \delta(t)$  and  $\tilde{1}(t) = e^{-\alpha_0 t}$ . Multiplying the above inequality with  $e^{-\alpha_0(t+1)}$ , we obtain

$$\tilde{\delta}(t+1) \leq \tilde{H} * \tilde{\delta}(t+1) + \tilde{g} * \tilde{1}(t+1) \gamma \Sigma_i(F) e_i.$$

Let  $\|\tilde{\delta}\|_1 = (\|\tilde{\delta}_i\|_1, i \in I)$  be the column vector where each entry is given by  $\|\tilde{\delta}_i\|_1 = \sum_{t=0}^\infty \delta_i(t)$ . Then we obtain, summing over  $t \geq 0$ ,

$$\|\tilde{\delta}\|_1 \leq \Lambda(\alpha_0) \|\tilde{\delta}\|_1 + \frac{1}{1 - e^{-\alpha_0}} \|\tilde{g}\|_1 \Sigma_i(F) e_i,$$

implying that

$$(Id - \Lambda(\alpha_0)) \|\tilde{\delta}\|_1 \leq \frac{1}{1 - e^{-\alpha_0}} \|\tilde{g}\|_1 \Sigma_i(F) e_i. \tag{A.11}$$

By (2.12),  $Id - \Lambda(\alpha_0)$  is invertible, and it is well-known that the operator norm of the inverse is bounded by

$$\| (Id - \Lambda(\alpha_0))^{-1} \| \leq (1 - \| \Lambda(\alpha_0) \|)^{-1} = C(\alpha_0).$$

Moreover,  $(Id - \Lambda(\alpha_0))^{-1} : \ell_\infty^+ \rightarrow \ell_\infty^+$ , where  $\ell_\infty^+ = \{(\xi_j)_{j \in I} : \xi_j \geq 0\}$ . Therefore, (A.11) implies

$$\sup_{j \in I} \|\tilde{\delta}_j\|_1 \leq \left[ \frac{1}{1 - e^{-\alpha_0}} \|\tilde{g}\|_1 \Sigma_i(F) \right] C(\alpha_0). \tag{A.12}$$

By using the union bound and (A.6), it follows that

$$\sup_{j \in I} \mathbb{P}_x (\exists s \in [1, t] : X_s(j) \neq X_s^{[F]}(j)) \leq \sup_{j \in I} \sum_{s=1}^t \delta_j(s) \leq e^{\alpha_0 t} \sup_{j \in I} \|\tilde{\delta}_j\|_1,$$

which implies the assertion of Item 1.

The proof of Item 2 is similar to the above argument, except that now it is possible to work directly with  $g(t)$  instead of  $\tilde{g}(t)$ . In this case, we write simply  $\bar{\delta}(t) = \sup_{j \in I} \delta_j(t)$ . (A.10) implies that

$$\left( \sup_{0 \leq s \leq t} \bar{\delta}(s) \right) \leq \chi \left( \sup_{0 \leq s \leq t} \bar{\delta}(s) \right) + \gamma \varrho \sum_{k \notin V_i \cap F} |W_{k \rightarrow i}|,$$

which implies the assertion. □

**Proof of (5.2).** The coupling inequality (5.2) follows now directly from (A.3) ((A.5), respectively). □

## Acknowledgments

This work is part of USP project *Mathematics, computation, language and the brain*, FAPESP project *Research, Innovation and Dissemination Center for Neuromathematics* (grant 2013/07699-0), CNPq projects *Stochastic modeling of the brain activity* (grant 480108/2012-9) and *Plasticity in the brain after a brachial plexus lesion* (grant 478537/2012-3), and of the project Labex MME-DII (ANR11-LBX-0023-01).

AD was fully and successively supported by CNPq and FAPESP fellowships (grants 201696/2015-0 and 2016/17791-9). GO was fully and successively supported by CNPq and FAPESP fellowships (grants 201572/2015-0 and 2016/17789-4). AG is partially supported by CNPq fellowship (grant 2016/17791-9.)

We thank the anonymous reviewers for their valuable comments and suggestions which helped us to improve the paper. We warmly thank B. Lindner and A.C. Roque for indicating us important references concerning stochastic models for neuronal activity.

## References

- [1] Adrian, E. and Bronk, D. (1929). The discharge of impulses in motor nerve fibres. Part II. The frequency of discharge in reflex and voluntary contractions. *J. Physiol.* **67** 119–151.
- [2] Adrian, E.D.A. (1928). *The Basis of Sensation: The Action of the Sense Organs*. London: Christophers.
- [3] Bresler, G. (2015). Efficiently learning Ising models on arbitrary graphs. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing* 771–782. New York, NY, USA: ACM.
- [4] Bresler, G., Mossel, E. and Sly, A. (2008). Reconstruction of Markov random fields from samples: Some observations and algorithms. In *Proceedings of the 11th International Workshop, APPROX 2008, and 12th International Workshop, RANDOM 2008 on Approximation, Randomization and Combinatorial Optimization: Algorithms and Techniques* 343–356. Berlin, Heidelberg: Springer.
- [5] Brillinger, D.R. (1988). Maximum likelihood analysis of spike trains of interacting nerve cells. *Biol. Cybernet.* **59** 189–200.
- [6] Brillinger, D.R. and Segundo, J.P. (1979). Empirical examination of the threshold model of neuron firing. *Biol. Cybernet.* **35** 213–220. DOI:10.1007/BF00344204.
- [7] Brochini, L., Costa, A.A., Abadi, M., Roque, A.C., Stolfi, J. and Kinouchi, O. (2016). Phase transitions and self-organized criticality in networks of stochastic spiking neurons. *Sci. Rep.* **6** 35831. DOI:10.1038/srep35831.
- [8] Brochini, L., Hodara, P., Pouzat, C. and Galves, A. (2017). Interaction graph estimation for the first olfactory relay of an insect. Available at [arXiv:1612.05226](https://arxiv.org/abs/1612.05226).
- [9] Csiszár, I. and Talata, Z. (2006). Consistent estimation of the basic neighborhood of Markov random fields. *Ann. Statist.* **34** 123–145. DOI:10.1214/009053605000000912.
- [10] De Masi, A., Galves, A., Löcherbach, E. and Presutti, E. (2015). Hydrodynamic limit for interacting neurons. *J. Stat. Phys.* **158** 866–902. DOI:10.1007/s10955-014-1145-1.
- [11] Duarte, A. and Ost, G. (2016). A model for neural activity in the absence of external stimulus. *Markov Process. Related Fields* **22** 37–52.
- [12] Duarte, A., Ost, G. and Rodríguez, A.A. (2015). Hydrodynamic limit for spatially structured interacting neurons. *J. Stat. Phys.* **161** 1163–1202. DOI:10.1007/s10955-015-1366-y.
- [13] Dyan, P. and Abbott, L.F. (2001). *Theoretical Neuroscience. Computational and Mathematical Modeling of Neural Systems*. MIT Press.
- [14] Fournier, N. and Löcherbach, E. (2016). On a toy model of interacting neurons. *Ann. Inst. Henri Poincaré Probab. Stat.* **52** 1844–1876. MR3573298
- [15] Galves, A. and Leonardi, F. (2008). Exponential inequalities for empirical unbounded context trees. In *In and Out of Equilibrium. 2. Progress in Probability* **60** 257–269. Birkhäuser, Basel. MR2477385
- [16] Galves, A. and Löcherbach, E. (2013). Infinite systems of interacting chains with memory of variable length – A stochastic model for biological neural nets. *J. Stat. Phys.* **151** 896–921. MR3055382
- [17] Galves, A., Orlandi, E. and Takahashi, D.Y. (2015). Identifying interacting pairs of sites in Ising models on a countable set. *Braz. J. Probab. Stat.* **29** 443–459.
- [18] Gerstner, W. (1995). Time structure of the activity in neural network models. *Phys. Rev. E* **51** 738–758. DOI:10.1103/PhysRevE.51.738.
- [19] Gerstner, W. and Kistler, W.M. (2002). *Spiking Neuron Models: Single Neurons, Populations, Plasticity*. Cambridge: Cambridge Univ. Press. MR1923120
- [20] Gerstner, W. and van Hemmen, J.L. (1992). Associative memory in a network of spiking neurons. *Network* **3** 139–164.
- [21] Hodara, P. and Löcherbach, E. (2017). Hawkes processes with variable length memory and an infinite number of components. *Adv. in Appl. Probab.* **49** 84–107. MR3631217
- [22] Hodgkin, A.L. and Huxley, A.F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.* **117** 500–544.

- [23] Lerasle, M. and Takahashi, D.Y. (2011). An oracle approach for interaction neighborhood estimation in random fields. *Electron. J. Stat.* **5** 534–571.
- [24] Lerasle, M. and Takahashi, D.Y. (2016). Sharp oracle inequalities and slope heuristic for specification probabilities estimation in discrete random fields. *Bernoulli* **22** 325–344.
- [25] Löcherbach, E. and Orlandi, E. (2011). Neighborhood radius estimation for variable-neighborhood random fields. *Stochastic Process. Appl.* **121** 2151–2185. [MR2819245](#)
- [26] Montanari, A. and Pereira, J.A. (2009). Which graphical models are difficult to learn? In *Advances in Neural Information Processing Systems* 22 1303–1311. Curran Associates, Inc.
- [27] Ravikumar, P., Wainwright, M.J. and Lafferty, J.D. (2010). High-dimensional Ising model selection using  $\ell_1$ -regularized logistic regression. *Ann. Statist.* **38** 1287–1319.
- [28] Reynaud-Bouret, P., Rivoirard, V. and Tuleau-Malot, C. (2013). Inference of functional connectivity in neurosciences via Hawkes processes. In *1st IEEE Global Conference on Signal and Information Processing*.
- [29] Robert, P. and Touboul, J. (2016). On the dynamics of random neuronal networks. *J. Stat. Phys.* **165** 545–584.
- [30] Soudry, D., Keshri, S., Stinson, P., Oh, M.-H., Iyengar, G. and Paninski, L. (2015). Efficient “shot-gun” inference of neural connectivity from highly sub-sampled activity data. *PLoS Comput. Biol.* **11** e1004464.
- [31] Talata, Z. (2014). Markov neighborhood estimation with linear complexity for random fields. In *Information Theory (ISIT), 2014 IEEE International Symposium on* 3042–3046.
- [32] Yaginuma, K. (2016). A stochastic system with infinite interacting components to model the time evolution of the membrane potentials of a population of neurons. *J. Stat. Phys.* **163** 642–658.

*Received April 2016 and revised June 2017*