HETEROGENEOUS MULTISCALE MULTIVARIATE AUTOREGRESSIVE MODEL: EXISTENCE, SPARSE ESTIMATION AND APPLICATION TO FUNCTIONAL CONNECTIVITY IN NEUROSCIENCE

Spaziani Stefano, Girardeau Gabrielle, Bethus Ingrid, Reynaud-Bouret Patricia

HETEROGENEOUS MULTISCALE MULTIVARIATE AUTOREGRESSIVE MODEL: EXISTENCE, SPARSE ESTIMATION AND APPLICATION TO FUNCTIONAL CONNECTIVITY IN NEUROSCIENCE

BY STEFANO SPAZIANI[1,a], GABRIELLE GIRARDEAU[2,b], INGRID BETHUS[3,c] AND PATRICIA REYNAUD-BOURET[4,d]

[1] *Université Côte d'Azur, CNRS, LJAD, France,* [a] *Stefano.Spaziani@univ-cotedazur.fr*

[2] *Institut du Fer-à-Moulin, Inserm U1270, Sorbonne Université, France,* [b] *gabrielle.girardeau@inserm.fr*

[3] *Université Côte d'Azur, CNRS, IPMC, France,* [c] *Ingrid.Bethus@univ-cotedazur.fr*

[4] *Université Côte d'Azur, CNRS, LJAD, France,* [d] *Patricia.Reynaud-Bouret@univ-cotedazur.fr*

In neuroscience, functional connectivity can be seen as a graph of interactions between brain oscillations rhythms and individual neuronal activity. This graph is associated with a cognitive state and helps understand high-cognitive processes such as learning. However, up to our knowledge, there is no model nor method to assess at once directed interactions between all these heterogeneous multiscale data. In this article, we propose a new model called HM-MVAR (Heterogeneous Multiscale Multivariate Autoregressive) to represent linear combinations of classic interaction patterns such as phase-locking or power-triggered phenomena. Because of the multiscale structure, we use a block version of stationarity to exhibit conditions under which the corresponding process exists and is stationary. We also propose a data-driven weighted LASSO estimator based on martingale exponential deviation inequalities that may have an interest per se. We prove that our estimator satisfies an oracle inequality and we show its good performance on realistic simulations. Finally, when applying it on a publicly available multiscale data set from the Buzsaki Lab, we recover interactions described in the literature but also uncover new phenomena of potential interest.

1. Introduction. To understand how the brain processes various external and internal stimuli in order to regulate behavior, neurophysiologists can record simultaneously during an experiment, two types of data: (i) the spiking activity (individual spikes or action potentials) of various neurons and (ii) the local field potentials (LFPs) of brain regions. Even if both spiking activity and LFPs are discretized at the time resolution of the recording device, they are of very different nature. On one hand, the action potentials, reconstructed as "spike trains" are a 0/1 phenomenon: they are time point processes where a point corresponds to a time of emission of an action potential by a neuron. On the other hand, local field potentials (LFPs) are continuous data made of the summation of the synaptic electrical activity of thousands of neurons near a recording electrode. LFPs resemble electroencephalograms (EEGs), except that classic EEGs are recorded on the scalp of the individual, most often a human, whereas LFPs are recorded from within the brain with intracranial electrodes, most often in rodents. The LFP/EEG signals can be decomposed into various frequencies associated with different brain states. These brain rhythms (for instance frequencies in the theta or gamma bands) are believed to play a role in synchronizing the neural activity within and across brain areas to efficiently process information [8].

MSC2020 subject classifications: Primary 62M10, 62P10; secondary 62G05, 62H12, 60G10.

Keywords and phrases: Autoregressive process, Hawkes process, LASSO estimator, Multiscale approach, Functional connectivity, Wavelet.

In the past decade, there have been more and more works in neurobiology showing that there are interactions between spike trains (synchronization, for instance) [23], interactions between rhythms (e.g., specific phases of the theta rhythm are associated with large gamma oscillations) [7] [30], but also that rhythms can influence the spiking activity of neurons (e.g., large gamma oscillations are associated with an increase in the firing rates of neurons) [22] and more recently that spikes reciprocally influence brain rhythms [36]. All these interactions are not constant over time and often appear in association with specific events (stimulus, behavior) or brain states (attention, sleep). This association between an interaction pattern or graph and a cognitive state is called "functional connectivity" in neuroscience. It represents a unique way to understand how the brain works [29].

There are generic tools to estimate interactions between homogeneous data (only spike trains or only LFP/EEG-rhythms), but the spikes-rhythms interplays are much more elusive. In the works cited above, the interactions were discovered by looking specifically at data type pairs (neurons / gamma rhythm for instance) that were already thought to be relevant for the experiment. There is a pressing need in the community for a generic tool that is able to use heterogeneous data (spike trains and LFP) to reconstruct an interaction graph between all of them, especially when the data come from two or more simultaneously recorded brain regions. There are various works providing partial methods to do this, which we discuss below. However, to the best of our knowledge, none is able to produce directed interaction graphs with rhythms and spike trains.

Graphs of dependence between heterogeneous data are very pervasive in recent years in statistics [11], but so far, the models behind do not encompass a time evolution neither a frequency decomposition as we need here.

If we focus on the point processes part, there have been various works to retrieve dependence graph, even when these graphs are fluctuating over time [24]. As for models of spikes trains, there has been a lot of work around Hawkes processes and variants such as Galves-Löcherbach model to estimate a functional connectivity graph [16] [26]. Condition to decide when the model is stationary or not have been found [16] [6], procedures to reconstruct the graph and sparse estimation have been provided [26]. There are also available tools to test the existence of an interaction in the graph [9]. Furthermore, it has been shown that Lasso estimators satisfy an oracle inequality[21] [32]. The present work aims at extending this to the much more general case of heterogeneous data, where point processes are mixed with a continuous signal.

When only LFP or EEG data are considered, there has been a huge amount work spent on multivariate autoregressive regression (MVAR) model [31] [4] [25] [42]. In particular condition for stationarity are known [4] and reconstruction of interaction are often done via Partial Directed Coherence (PDC) and test on the existence of an interaction [3, 14]. In [4], the convergence of the Least Squares estimator is shown, but up to our knowledge, no oracle inequality or sparse estimator of the graph is available. Moreover, when dealing with rhythms in LFP/EEG, the previous PDC approach can be used rhythm per rhythm but cannot directly find interaction between different rhythms. To do so, there has been some attempts to look directly at coherence or partial coherence despite an acknowledged lack of interpretability of the results [2] and even some works use these notions between rhythms and spike trains [38][43]. Note in particular that with these notions, the resulting graph is not directed and one cannot know exactly which data is influencing the other one because there is no temporal model behind.

Some works tried to mix spike trains and LFP without the rhythm approach. In [18], the authors present a model-free framework in which the interactions are detected through non parametric tests using Granger's causalities. In [27] the LFP can influence the spike trains but it is considered only as exogenous data and is not modeled. The authors of [27] prove

consistence of the Least Square estimator of the spiking probabilities, as well as the consistence for the variance estimator of the spike-LFP modulation. In [39] a similar approach is presented, but also the maximal interaction lag is a parameter to be estimated. In [1] with a similar model, the parameters are fitted through an unsupervised learning algorithm. In [35] the model also includes a latent brain state variable.

Let us also mention one work mixing heterogeneous data [12][5], where spike trains - modeled as a Hawkes process - are influencing directly the voltage of one given neuron, the purpose being mainly to adaptively estimate the functions parameterizing the voltage itself and not the whole range of possible interactions between the data at hand.

The model we propose here, called Heterogeneous Multiscale MultiVariate AutoRegressive model (HM-MVAR), mixes the point process signals and the wavelet coefficients of the continuous signals thanks to interaction functions between wavelet coefficients, between spike trains and between wavelets and spikes trains.

Models able to catch dynamics of discrete time wavelets coefficients are numerous in the literature. In particular Markov models have been used, especially for the dependency of one rhythm on another, but only the directed interactions from the coarser wavelet coefficient to the finest wavelet coefficients were considered[40, 15, 41]. In [41] the convergence rate of the Kalman Filter Least Squares estimator is shown, together with the corresponding algorithm complexity. However in all of these works, the relations between the variables is always vertical, generally from coarser to finer scales and time is not taken into account so that the evolution over time of the process is not modeled.

To summarize, there is no work, up to our knowledge, able to estimate directed interactions in time between wavelets coefficients modeling the rhythms and point processes modeling the spike train. The aim of the present work is to:

- Propose a generative probabilistic model, called HM-MVAR, able to take into account various type of interactions between all the LFP rhythms and spike trains at hand. Each time, an interaction corresponds to an interaction function from one type of data (one rhythm, one spike train) to another, with a certain range in the past. Therefore the interaction is directed and has a causal interpretation in Granger sense [19].
- Show that such a process exists and has a block stationary version under some constraints.
- Propose a Lasso estimator of the interaction functions that satisfies an oracle inequality and therefore provide a sparse directed graph to summarize the interaction between the data.
- Show on simulation and on a representative open data set [33, 17] that our method is able to provide a meaningful heterogeneous functional connectivity graph in neuroscience.

In Section 2, we present our model and the variety of typical situations that it can model. In Section 3, we prove the existence of a block stationary solution. In Section 4, we propose a Lasso estimator of the interaction functions, with weights that are issued from martingales concentration results that may have an interest per se. We also prove that this Lasso estimator satisfies an oracle inequality. In Section 5, a simulation study is presented and in Section 6, we present the results on real data and discuss their relevance.

2. The HM-MVAR model .

2.1. *Notation.* We are focusing on a discrete time setting where the time index t lies in $\mathbb{Z}$. We use the notation $a:b$ to represent all integers i such that $a \leq i \leq b$.

We want to model the interactions of P continuous data, $(X_t^p)_{p\in 1:P} \in \mathbb{R}^P$ (e.g. the local field potentials of P different brain regions) and M point data, $(S_t^m)_{m\in 1:M} \in \{0,1\}^M$ (e.g. spike trains of M different neurons).

We assume that we have the following wavelet decomposition of the continuous data:

$$\forall t \in \mathbb{Z}, \quad X_t^p = \sum_{j=0}^{J} \sum_{k \in \mathbb{Z}} W_{j,k}^p \phi_{j,k}(t). \tag{1}$$

The $W_{j,k}^p$'s reflect the intensity of the pth LFP in the frequence band corresponding to 2^j (called the jth rhythm in the sequel) at location $k2^{J+1-j}$ as soon as $j \geq 1$. The $W_{0,k}$'s reflect the slow drift in the data. The finest level J is usually defined in link with the data frequency acquisition and the preprocessing steps performed by the neurophysiologists to separate LFP from spike trains, so that in practice J is fixed and quite small (up to 10).

To write (1), we consider a discrete, eventually biorthogonal wavelet basis $\phi_{j,k}(.)$, where $j \in 0:J$ is the scale parameter and $k \in \mathbb{Z}$ the localisation (see [10], [20]). More precisely, $j = 0$ corresponds to the father wavelet with $\phi_{0,k}(t) = \phi_{0,0}(t - k2^J)$. For $j \geq 1$, we have that $\phi_{(j,k)}(t) = \phi_{j,0}(t - k2^{J+1-j})$. If we were in continuous time, those would have been dilation of the mother wavelet, but in discrete time it is not stricto sensu true, except for the Haar basis [20].

For interpretability of our models, we need that all $\phi_{j,k}$'s have a finite support that we denote $[\ell_{j,k}, r_{j,k}]$ and because of the translation properties above, we have that

$$\forall k \in \mathbb{Z}, \quad \begin{cases} \ell_{j,k} = \ell_{j,0} + k2^{J+1-j} & \text{if } j \geq 1 \\ \ell_{0,k} = \ell_{0,0} + 2^J & \text{if } j = 0. \end{cases}$$

Our model is describing the evolution of $(X_t^p)_{p \in 1:P}$ and $(S_t^m)_{m \in 1:M}$ by describing directly the evolution of the coefficients themselves $(W_{j,k}^p)_{p,j}$ as a function of k, jointly with the evolution of $(S_t^m)_{m=1,\dots,M}$. In particular, since $W_{j,k}^p$ represents the intensity of the jth rhythm of the pth LFP's at location $k2^{J+1-j}$, we want to describe the co-evolution of processes that do not happen at the same time scale. From a modeling point of view, we want to keep X_t^p adapted to the filtration $\mathcal{F}_t$ that we use. This is why we are considering in the sequel that

$$\forall p \in 1:P, j \in 0:J, k \in \mathbb{Z}, \quad \text{the variable } W_{j,k}^p \text{ appears at time } \ell_{j,k},$$

so that with (1), we can compute X_t^p thanks to all the variables that have happen before time t.

More precisely, the filtration is defined by

$$\forall t \in \mathbb{Z}, \quad \mathcal{F}_t = \sigma\left((S_u^m)_{\substack{m \in 1:M \\ u \leq t}}, (W_{j,k}^p)_{\substack{p \in 1:P \\ j \in 0:J \\ \ell_{j,k} \leq t}} \right)$$

2.2. *Evolution of the point process part.* The model of evolution of the point processes part is inspired by discrete Hawkes processes [32], where we incorporate dependency to the rhythms. This leads to this set of equations for the point process part S_t^m:

$$\begin{aligned} \mathbb{P}\left(S_t^m = 1 | \mathcal{F}_{t-1}\right) &= \nu^m + \sum_{\substack{m' \in 1:M \\ u < t}} h_{m'}^m(t-u) S_u^{m'} + \sum_{\substack{p \in 1:P \\ \kappa_{pj}^S \in \mathbb{K}_{pj}^S \\ j,k/\ell_{j,k} < t}} h_{p,j,\kappa_{pj}}^m(t - \ell_{j,k}) \kappa_{pj}^S(W_{j,k}^p) \\ &= \nu^m + \sum_{\substack{m' \in 1:M \\ d \in 1:+\infty}} h_{m'}^m(d) S_{t-d}^{m'} + \sum_{\substack{p \in 1:P \\ \kappa_{pj}^S \in \mathbb{K}_{pj}^S \\ j,k/\ell_{j,k} < t}} h_{p,j,\kappa_{pj}}^m(t - \ell_{j,k}) \kappa_{pj}^S(W_{j,k}^p) \end{aligned} \tag{2}$$

In the previous equation, if the wavelet part on the right is not there, one recognizes the classical Hawkes processes equation (see [32]). Indeed, the spontaneous part, ν^m, corresponds to the firing rate of neuron m if no interaction takes place. The synaptic integration is present as the sum of all $h^m_{m'}(d)$'s if there is a spike ($S^{m'}_u = 1$) at distance d: if $h^m_{m'}(d) > 0$, neuron m' excites neuron m after a delay d, whereas if $h^m_{m'}(d) < 0$, neuron m' inhibits neuron m after a delay d. An example of such interactions is provided in the following example.

EXAMPLE (Excitation of one neuron by another neuron). Figure 1 shows the spike trains of two neurons where Neuron 2 excites Neuron 1. This corresponds to a positive h^1_2 at short range: shortly after a spike generated by Neuron 2 we have more spikes produced by Neuron 1. To pass from the model to a realistic raster plot we have to divide the discrete time indexes by the sampling frequency of our instrument, usually ≥ 1000Hz.

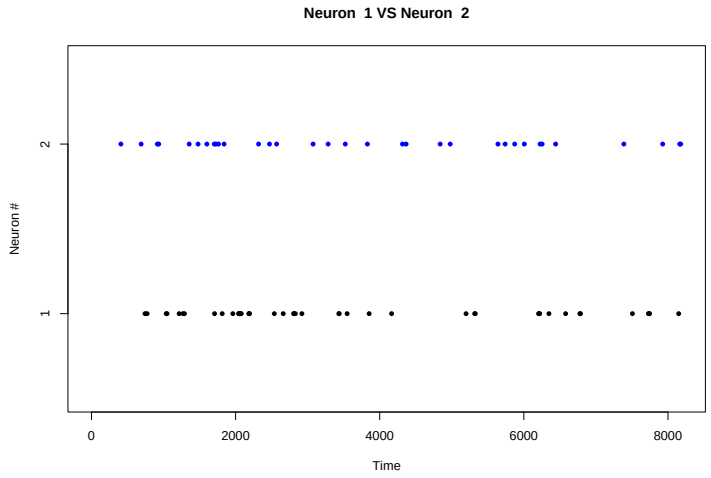


FIG 1. *Excitation of one neuron by another one. The time occurrences of spikes are represented by dots (blue for Neuron 2, black for Neuron 1). The parameters characterizing the two spiking probabilities (see (2)) are:* $\nu^1 = 4e-4, \nu^2 = 2.8e-3, h^1_2(t) = 0.041\!\!1_{t\in[331,360]}$.

The wavelet part in the above equation (2) is more intricate. First, (see below) the $W^p_{j,k}$'s are unbounded, so we cannot use a linear transformation of $W^p_{j,k}$ in (2). This is why we use a mask $\kappa^S_{j,k}(W^p_{j,k})$ with $\kappa^S_{j,k}(0) = 0$, assumed to be 1-Lipschitz and bounded by two absolute constants $\underline{w}_{p,j} \leq 0$ and $\bar{w}_{p,j} \geq 0$ so that

$$\underline{w}_{p,j} \leq \kappa^S_{pj}(W^m_{j,k}) \leq \bar{w}_{p,j}.$$

A typical example is (with fixed $w > 0$)

$$\kappa^S_{pj}(W^p_{j,k}) = \kappa^S_w(W^p_{j,k}) := \begin{cases} w & \text{if } W^p_{j,k} > w \\ W^p_{j,k} & \text{if } -w \leq W^p_{j,k} \leq w \\ -w & \text{if } W^p_{j,k} < -w \end{cases}$$

Because we can envision many of these masks, one might want to incorporate several of them, in the description. This is why we can use a finite dictionary $\mathbb{K}^S_{pj}$ of at most K masks, that might be tuned for each wavelet level (p, j).

For each of these masks $\kappa^S_{pj}(W^p_{j,k})$ and each wavelet level (p,j), one has an interaction function $h^m_{p,j,\kappa^S_{pj}}(t-\ell_{j,k})$. At the difference with the interaction functions $h^m_{m'}$, its impact is not just modulated by a 0/1 phenomenon but by the value of $\kappa^S_{pj}(W^p_{j,k})$. To fix ideas, one can envision the following example.

EXAMPLE (Phase lock of spikes on a given rhythm). Let us define two masks $\kappa_+(W) = |\kappa^S_w(W)|\mathbb{1}_{W>0}$ and $\kappa_-(W) = |\kappa^S_w(W)|\mathbb{1}_{W>0}$. With the Haar basis, one can use for a fixed positive parameter a,

$$h^m_{p,j,\kappa_+} = a\mathbb{1}_{1:2^{J-j}-1} \text{ and } h^m_{p,j,\kappa_-} = a\mathbb{1}_{2^{J-j}:2^{J+1-j}-1}.$$

This models a phase locking phenomenon, where spikes of m are more likely to appear in the upper oscillation of the jth rhythm of the pth LFP (see Figure 2). Indeed if $W^p_{j,k} > 0$, the oscillation of length 2^{J+1-j} presents the up phase followed by the down phase, so in the first half of the period, the neuron is excited. If $W^p_{j,k} < 0$, the oscillation of length 2^{J+1-j} presents the down phase followed by the up phase, so it is in the second half of the period that the neuron is excited. Of course this phenomenon is stronger when $|W^p_{j,k}|$ is large, as one can see on Figure 2.

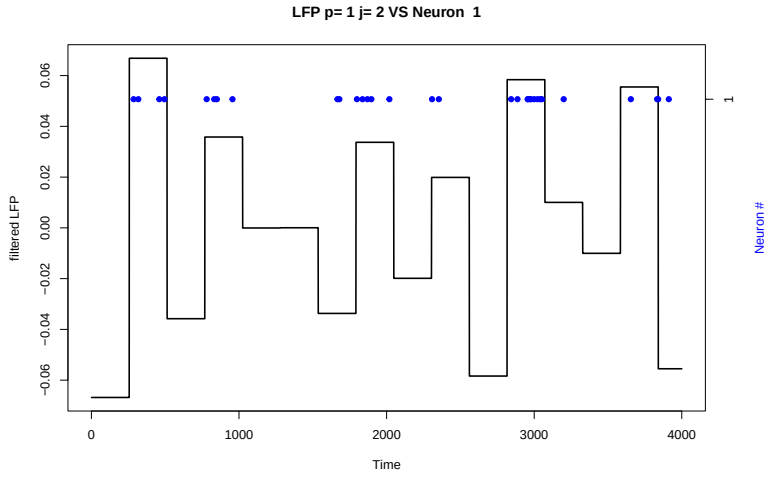


FIG 2. *Phase lock of spikes on a given rhythm. The LFP filtered at level $j = 2$ is represented in black and the spikes of Neuron 1 are represented by blue dots. The parameters characterizing the equations (2) and (4) are: $\nu^1 = 0.0024, h^1_{1,2,\kappa_+} = 0.011\mathbb{1}_{1:2^8-1}, h^1_{1,2,\kappa_-} = 0.011\mathbb{1}_{2^8:2^9-1}, \mu^{1,2} = 0.5, \sigma^{1,2} = 1$. The LFP signal is always computed starting from the wavelets coefficient through the wavelet antitransform (1).*

Note that (2) can only hold if a condition of the following type is satisfied, to ensure that the full quantity can indeed be interpreted as a probability:

$$(3)\quad \begin{cases} 1 \geq \nu^m + \displaystyle\sum_{\substack{m' \in 1:M \\ u<t}} \max(h^m_{m'}(t-u),0) + \sum_{\substack{p\in 1:P \\ \kappa^S_{pj}\in \mathbb{K}^S_{pj} \\ j,k/\ell_{j,k}<t}} \max(h^m_{p,j,\kappa^S_{pj}}(t-\ell_{j,k}),0)\bar{w}_{p,j} \\ 0 \leq \nu^m + \displaystyle\sum_{\substack{m' \in 1:M \\ u<t}} \min(h^m_{m'}(t-u),0) + \sum_{\substack{p\in 1:P \\ \kappa^S_{pj}\in \mathbb{K}^S_{pj} \\ j,k/\ell_{j,k}<t}} \min(h^m_{p,j,\kappa^S_{pj}}(t-\ell_{j,k}),0)\underline{w}_{p,j} \end{cases}$$

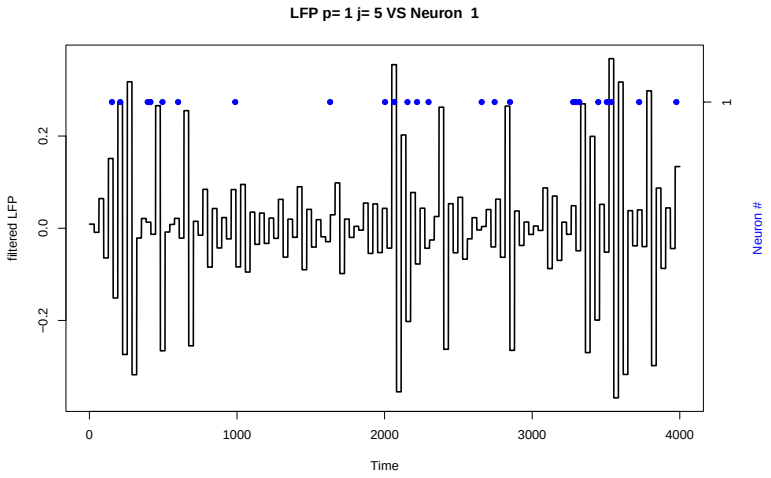


FIG 3. *Spikes triggering a given rhythm. The LFP filtered at $j = 5$ is represented in black and the spike train of Neuron 1 in blue. The parameters characterizing equations (2) and (4) are: $\nu^1 = 0.0064, \mu^{1,5} = -0.3, \sigma^{1,5} = .1, h_1^{1,5}(t) = 21\!\!1_{t\in[31,60]}$.*

2.3. *Evolution of the wavelet coefficients.* The model of evolution for the continuous part is inspired by Multivariate Autoregressive models for the continuous processes [14], except that it is written on the rhythms directly, keeping in mind that a given $W^p_{j,k}$ appears at $\ell_{j,k}$. This leads to the following equation.

(4)
$$W^p_{j,k} = \mu^{p,j} + \sum_{\substack{m\in 1:M \\ u<\ell_{j,k}}} h^{p,j}_m(\ell_{j,k}-u)S^m_u + \sum_{\substack{p'\in 1:P \\ \kappa^W_{p'j'}\in \mathbb{K}^W_{p'j'} \\ j',k'/\ell_{j',k'}<\ell_{j,k}}} h^{p,j}_{p',j',\kappa^W_{p'j'}}(\ell_{j,k}-\ell_{j',k'})\kappa^W_{p'j'}(W^{p'}_{j',k'}) + \varepsilon^p_{j,k},$$

In the previous equation, all the $\varepsilon^p_{j,k}$ are independent Gaussian variables with mean 0 and variance $\sigma^{p,j}$. The parameter $\mu^{p,j}$ gives the main trend of the wavelet coefficient $W^p_{j,k}$. In classical MVAR settings, variables are usually centered before hand [14], but it is not possible here because of the interaction term with the point process part given by the $h^{p,j}_m$'s. However if we use the Haar basis, set $J = 0$ and discard interaction with spikes, we recover a classical MVAR model where $W^{p,0}_t = X^p_t$.

The interaction with the point process part can be easily understood in the light of the next example.

EXAMPLE (Spikes triggering a given rhythm). If the function $h^{p,j}_m$ is large over $1 : 2^{J+1-j} - 1$ and if neuron m spikes a lot, the next wavelet coefficient increases. This is what we see in Figure 3.

Interaction between rhythms can be interpreted in the same way. Again one might want to use masks $\kappa^W_{p'j'}$ such that $\kappa^W_{p'j'}(0) = 0$. Here these masks are not forced to be bounded, 1-Lipschitz is a sufficient constraint that is assumed in the sequel. A typical mask is the identity $\kappa^W_{p'j'}(W^{p'}_{j',k'}) = W^{p'}_{j',k'}$, but one might want to use the following set of masks to describe a phase locked phenomenon between rhythms, as in the following example.

EXAMPLE (A fast rhythm power that is in phase-power modulation with slower oscillations). Imagine $j > j'$ and use $\kappa^1(w) = \max(w, 0)$, $\kappa^2(w) = \min(w, 0)$, then with

$$h^{p,j}_{p',j',\kappa^1} = a\mathbb{1}_{1:2^{J-j'}-1} \text{ and } h^{p,j}_{p',j',\kappa^2} = a\mathbb{1}_{2^{J-j'}:2^{J+1-j'}-1},$$

we obtain that

$$h^{p,j}_{p',j',\kappa^1}(\ell_{j,k} - \ell_{j',k'})\kappa^1(W^{p'}_{j',k'}) + h^{p,j}_{p',j',\kappa^2}(\ell_{j,k} - \ell_{j',k'})\kappa^2(W^{p'}_{j',k'}),$$

is large when $|W^{p'}_{j',k'}|$ is large and when $\ell_{j,k}$ is in the upper part of the oscillation corresponding to $W^{p'}_{j',k'}$, whatever the sign of $W^{p'}_{j',k'}$. A fast rhythm phase locked on the low part of a slower oscillation is given in Figure 4.

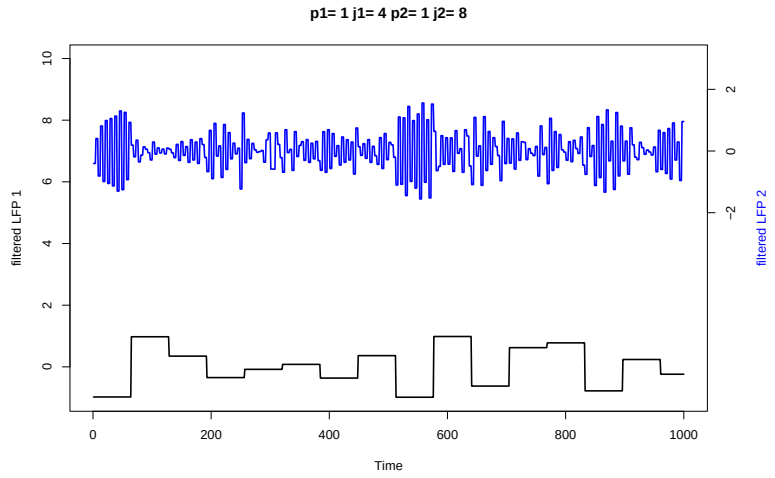


FIG 4. *A fast rhythm power that is in phase-power modulation with slower oscillations. The slow rhythm (LFP filtered at $j = 4$) is presented in black whereas the fast rhythm (LFP filtered at $j = 8$) is in blue. The parameters of the models (4) are: $\mu^{1,4} = \mu^{1,8} = -0.3, \sigma^{1,8} = \sigma^{1,4} = 1, h^{1,8}_{1,4,\kappa_1} = -2\mathbb{1}_{2^6:2^7-1} h^{1,8}_{1,4,\kappa_2} = -2\mathbb{1}_{1:2^6-1}$.*

DEFINITION 2.1 (Connectivity graph). For a HM-MVAR model given by (2) and (4), the corresponding connectivity graph is defined as a graph where vertices correspond to either neurons (from 1 to M) or couples (p, j) standing for the jth rhythm of the pth LFP. A directed arrow from one vertex i' to another i means that there is non null interaction function $h^i_{i'}$. Colors of the edges can stand for instance for excitation or inhibition, style for the type of mask that is used and width for the strength of the interaction function. Some examples are given in Sections 5 and 6.

From now on, we are interested in a model where both (2), (3) and (4) hold. The next section focus on giving condition for existence of such processes on $\mathbb{Z}$, whereas Section 4 focus on the estimation of the unknown parameters of the model.

3. Existence of a block stationary solution .

3.1. *Block notation.* Let us vectorize the previous equations. For all time t, indices p and j, we denote $W_t^{p,j}$ the process defined by $W_{j,k}^p$ when t is an $\ell_{j,k}$ and 0 otherwise.

We denote by Y_t the vector of $\mathbb{R}^{M+P(J+1)}$ such that its first M coordinates are the S_t^m and the next are the $W_t^{p,j}$ (using the lexicographic order on (p,j)).

We also denote by $\kappa^S(Y_t)$ (resp. $\kappa^W(Y_t)$) the finite vector formed as Y_t but where a given coefficient $W_t^{p,j}$ is replaced by a family of coefficients $\kappa_{pj}^S(W_t^{p,j})$ (resp. $\kappa_{pj}^W(W_t^{p,j})$), for $\kappa_{pj}^S \in \mathbb{K}_{pj}^S$ (resp. $\kappa_{pj}^W \in \mathbb{K}_{pj}^W$). Note that $\kappa^S(Y_t)$ (resp. $\kappa^W(Y_t)$) is a vector with at most $M + KP(J+1)$, where K is the maximal cardinality for the dictionnaries. Note that if $K = 0$, meaning that no interaction from wavelets are considered, then $\kappa^S(Y_t) = \kappa^W(Y_t) = (S_t^m)_{m\in 1:M}$.

The matrix formed of columns Y_t for $a \le t \le c$, is denoted Y_a^c. In this notation, a might be $-\infty$. In the same spirit, $\kappa^S(Y_a^c)$ (resp. $\kappa^W(Y_a^c)$) is a matrix formed of the columns defined by $\kappa^S(Y_t)$ (resp. $\kappa^W(Y_t)$).

For a given matrix Z that has a potentially infinite number of columns labeled from $-\infty$ to -1 and Q rows and a given matrix H with Q columns and D rows of functions with support in $1:+\infty$, we denote $H * Z$ the vector of dimension D given by

$$H * Z = \sum_{k=1}^{+\infty} H(k) Z_{-k},$$

where the product $H(k)Z_{-k}$ is the classical matrix product between the matrix H evaluated at time k and the column of Z with label $-k$ if it exists - one multiplies by 0 if Z_{-k} does not exist in Z.

In particular, let H^S be the matrix formed at row m by the $h_{m'}^m(.)$'s and $h_{p,j,\kappa_{p,j}^S}^m$'s and let H^W be the matrix formed at row (p,j) (in the lexicographical order) by the $h_m^{p,j}(.)$'s and $h_{p',j',\kappa_{p',j'}^W}^{p,j}$'s. We have that for all indices m and p,j,

$$[H^S * \kappa^S(Y_{-\infty}^{t-1})]_m = \sum_{\substack{m' \in 1:M \\ u<t}} h_{m'}^m(t-u) S_u^{m'} + \sum_{\substack{p \in 1:P \\ \kappa_{pj}^S \in \mathbb{K}_{pj}^S \\ j,k/\ell_{j,k}<t}} h_{p,j,\kappa_{pj}}^m (t-\ell_{j,k}) \kappa_{pj}^S(W_{j,k}^p)$$

and

$$[H^W * \kappa^W(Y_{-\infty}^{t-1})]_{p,j} = \sum_{\substack{m \in 1:M \\ u<\ell_{j,k}}} h_m^{p,j}(\ell_{j,k}-u) S_u^m + \sum_{\substack{p' \in 1:P \\ \kappa_{p'j'}^W \in \mathbb{K}_{p'j'}^W \\ j',k'/\ell_{j',k'}<\ell_{j,k}}} h_{p',j',\kappa_{p'j'}^W}^{p,j}(\ell_{j,k}-\ell_{j',k'}) \kappa_{p'j'}^W(W_{j',k'}^{p'}).$$

Note that it is straightforward to prove the following control of the ℓ_1 norm:

$$\forall H, Z, \quad \|H * Z\|_1 \le \sum_{k=1}^{+\infty} \|H(k)\|_{1,\infty} \|Z_{-k}\|_1, \tag{5}$$

where for any matrix A, $\|A\|_{1,\infty} = \sum_i \max_j |A_{i,j}|$.

Finally let us denote for all $b \in \mathbb{Z}$, $i_b = b2^J$, $j_b = (b+1)2^J - 1$ and $X_b = Y_{i_b}^{j_b}$.

Let us formalize a trivial statement that is useful to define properly block stationarity.

LEMMA 3.1 (Definition of Block Stationarity). *If the sequence of blocks $(X_b)_{b\in\mathbb{Z}}$ is stationary, then for all $s,t \in \mathbb{Z}$ such that $s = t \mod (2^J)$, $Y_s^{s+2^J-1}$ has the same distribution as $Y_t^{t+2^J-1}$. This is what we call block stationarity.*

In particular, it means that

$$\left((S_u^m)_{\substack{m\in 1:M \\ u\in s:s+2^J-1}}, (W_{j,k}^p)_{\substack{p\in 1:P \\ j\in 0:J \\ \ell_{j,k}\in s:s+2^J-1}} \right) \quad \textit{and} \quad \left((S_u^m)_{\substack{m\in 1:M \\ u\in t:t+2^J-1}}, (W_{j,k}^p)_{\substack{p\in 1:P \\ j\in 0:J \\ \ell_{j,k}\in t:t+2^J-1}} \right)$$

have the same distribution if $s = t \mod (2^J)$.

Therefore, in the sequel, we want to derive under which condition the blocks X_b of length 2^J may form a stationary sequence. Note that we cannot hope for full stationarity of the vector Y_t, because even the number of non zero coefficients in it depends on the pattern of the $\ell_{j,k}$'s and is therefore 2^J periodic. To prove block stationarity, we want to apply Theorem 3.1 of Doukhan and Wintenberger [13] and to do so we need to express X_b as a function of the previous blocks. This is what we do in the following section.

3.2. *Recursion formulas* . Let us introduce the vector of innovation $\xi_t = \begin{bmatrix} \mathbf{U}_t \\ \varepsilon_t \end{bmatrix}$ of size $M + P(J+1)$ such that the first M coordinates, given by $\mathbf{U}_t$, are i.i.d. Uniform variables on $[0,1]$ and the other $P(J+1)$ coordinates, ε_t, represent the independent noises on the wavelet coefficients, the coordinate (p,j) (in the lexicographical order) being $\varepsilon_t^{p,j}$, a centered Gaussian variable of variance $\sigma^{p,j}$.

At time s, it is quite clear that if (2), (3) and (4) are satisfied, then one can represent Y_s as a function g_s of the past $Y_{-\infty}^{s-1}$ and of the innovations:

$$Y_s = g_s(Y_{-\infty}^{s-1}, \xi_s) = \begin{bmatrix} \mathbb{1}_{\mathbf{U}_s \le \nu + H^S * \kappa^S(Y_{-\infty}^{s-1})} \\ (\mu + H^W * \kappa^W(Y_{-\infty}^{s-1}) + \varepsilon_s).\delta(s) \end{bmatrix},$$

where the first indicator function has to be understood coordinate by coordinate. The . product is the product coordinate by coordinate and $\delta(s)$ is the vector of size $P(J+1)$ which at coordinate (p,j) is $\mathbb{1}_{\exists k, \ell_{j,k}=s}$.

Note that $\delta(s)$ is 2^J periodic, so that $g_s(.,.) = g_{s+2^J}(.,.)$. This means that the dependence of a block X_b on its past only depends on its past and the innovation vectors. To find the exact function, let us now focus on the first block X_0, with index $b = 0$, for simplicity. The second column of the block is depending on $Y_{-\infty}^0 = \left[Y_{-\infty}^{-1}, g_0(Y_{-\infty}^{-1}, \xi_0)\right]$ etc. So we can recursively define for a given past $y = Y_{-\infty}^{-1}$ of the block X_0

$$\begin{cases} \text{if } s = 0, \mathbb{G}_0(y, \xi_0) = g_0(y, \xi_0) & \in \mathbb{R}^{M+P(J+1)} \\ \text{if } s > 0, \mathbb{G}_s(y, \xi_0^s) = [\mathbb{G}_s(y, \xi_0^{s-1}), g_s([y, \mathbb{G}_s(y, \xi_0^{s-1})], \xi_s)] & \in \mathbb{R}^{(M+P(J+1))\times(s+1)} \end{cases} \tag{6}$$

In the end, if (2), (3) and (4) hold, we always have that for all $b \in \mathbb{Z}$,

$$X_b = \mathbb{G}_{2^J-1}(X_{-\infty}^{b-1}, \zeta_b), \tag{7}$$

with the following notations:

- $X_{-\infty}^{b-1} = Y_{-\infty}^{j_b}$ is the past of block b and consists of all the blocks that have happened before block X_b,
- $\zeta_b = \xi_{i_b}^{j_b}$ is the matrix of innovations for the block X_b.

3.3. *Block stationarity.* Now we are ready to apply Doukhan and Wintenberger's result [13].

THEOREM 3.2. *Recall that K is the maximal size of the mask dictionaries $\mathbb{K}^S_{pj}$ and $\mathbb{K}^W_{pj}$.*

$$e_b = (K \vee 1)\left(\max_{k\in i_{b-1}+1:j_b} \|H^S(k)\|_{1,\infty} + \max_{k\in i_{b-1}+1:j_b} \|H^W(k)\|_{1,\infty}\right)$$

and

$$d = (K \vee 1)\left(\max_{k\in 1:2^J-1} \|H^S(k)\|_{1,\infty} + \max_{k\in 1:2^J-1} \|H^W(k)\|_{1,\infty}\right).$$

Assume that (3) holds and that

$$\begin{cases} \sum_{b=1}^{+\infty} e_b < \dfrac{d}{(1+d)^{2^J}-1} & \text{if } d>0 \\ \sum_{b=1}^{+\infty} e_b < 2^{-J} & \text{if } d=0 \end{cases}$$

then there exists a block τ-weakly dependent stationary solution to (2) and (4), which is integrable.

If in addition all the masks κ^W_{pj}'s are bounded, then this block stationary solution has moment of every order.

Note that the condition on the interaction functions is satisfied as soon they are summable and small enough. In particular the condition does not depend on $\sigma_{p,j}$.

When $J=0$, meaning that we do not use the wavelet expansion and just use a spike/LFP interaction model, we have $d=0$. In this case, the term e_b is just

$$e_b = (K\vee 1)\left(\|H^S(b)\|_{1,\infty} + \|H^W(b)\|_{1,\infty}\right)$$

In particular, if there is only spikes and no LFP in the system (and therefore $K=0$), we recover the condition

$$\sum_{b=1}^{\infty}\sum_{m\in 1:M}\max_{m'\in 1:M}|h^m_{m'}(b)|<1.$$

This condition is therefore slightly more reductive than the classical Hawkes condition saying that the spectral radius of $(\sum_b |h^m_{m'}(b)|)_{m,m'\in 1:M}$ is strictly smaller than 1 [6] (see also [32] for more restrictive conditions).

Therefore our condition is not necessary but only sufficient in this case. We do not know if it is necessary when considering the whole multiscale approach.

To apply Doukhan and Wintenberger to our problem, we need to rely on a norm on $\mathbb{R}^{M+P(J+1)}$ that we chose as the ℓ_1-norm and an Orlicz norm linked to a convex function Φ, that we chose equal to identity. This is why we end up proving the existence of only a L^1 block stationary solution. Passing directly to an L^2 Orlicz norm would require Lipschitizianity of the point process component w.r.t. the L^2 norm and it is not possible. To realize it, let us consider $a<c\in[0,1]$ and U a uniform random variable on $[0,1]$: $\mathbb{E}((\mathbb{1}_{U<a}-\mathbb{1}_{U<c})^2) = \mathbb{E}(|\mathbb{1}_{U<a}-\mathbb{1}_{U<c}|) = \mathbb{E}(\mathbb{1}_{a<U<c}) = |c-a|$ so $\mathbb{1}_{U<a}$ is Lipschitz w.r.t. an L^1 Orlicz norm but not an L^2 Orlicz norm.

However, the point process components being always bounded, they have moments of every order and as soon as the masks $\kappa^W_{p,j}$'s are bounded as well, the fact that the noise is Gaussian lead to existence of moment of every order. We have not been able to prove it without the boundedness assumptions on the masks.

We need to prove two conditions to apply Theorem 3.1 of Doukhan and Wintenberger [13]. This is detailed in the two following lemmas. We inform that all the proofs can be found in Section 1 of the Supplementary Material (Spaziani, Girardeau, Reynaud-Bouret, Bethus (2023)).

LEMMA 3.3. *Let $m > 0$ be a fixed integer. The function $\mathbb{G}_{2^J-1}$ defined in Section 3.2 by (7) and the associated innovation matrix ζ_0 for block 0, satisfy, for all $s \geq 0$*

$$\mathbb{E}(\|\mathbb{G}_{2^J-1}(\mathbf{0}, \zeta_0)\|_1^m) < \infty,$$

where $\mathbf{0}$ is the matrix corresponding to null past before 0.

LEMMA 3.4. *Let $X_{-\infty}^{-1}$ and $Z_{-\infty}^{-1}$ be two different (block) infinite pasts that are possible for the block X_0, then*

$$\mathbb{E}(\|\mathbb{G}_{2^J-1}(X_{-\infty}^{-1}, \zeta_0) - \mathbb{G}_{2^J-1}(Z_{-\infty}^{-1}, \zeta_0)\|_1) \leq \sum_{b=1}^{\infty} a_b \mathbb{E}(\|X_{-b} - Z_{-b}\|_1),$$

where

$$a_b = \begin{cases} ((1+d)^{2^J} - 1)\dfrac{e_b}{d} & \text{if } d > 0 \\ 2^J e_b & \text{if } d = 0 \end{cases}$$

4. Sparse estimation by Lasso criterion . In this section, we want to estimate μ, ν and the interaction functions, to reconstruct a sparse connectivity graph between neurons and LFP rhythms. Because formula (2) and (4) are linear in these unknown functions and parameters, and since these interaction functions only depend on a certain lag with respect to the data at hand (S_t^m or $W_{j,k}^p$), we are going to approximate formulas (2) and (4) by a linear combination of cylindrical functions of the past, i.e. functions only depending on a small neighborhood of the data that we want to explain. Let us present in more details this dictionary of cylindrical functions.

4.1. *Dictionary of cylindrical functions.* Let us denote

$$\mathcal{I} = 1 : M \ \cup \ (1 : P \ \times \ 0 : J),$$

so that an index $i \in \mathcal{I}$ can be either referring to neuron m or to a couple (p, j), standing for the jth rhythm of the pth LFP.

DEFINITION 4.1 (Neighborhood). A neighborhood v is a finite subset of $\mathcal{I} \times 1 : +\infty$. It represents a collection of indices of data and lags.

We consider real valued functions ϕ of the past, that is $\phi : Y_{-\infty}^{t-1} \mapsto \mathbb{R}$, that only depends on the past configuration and not on the precise value of t. Note however that because of the pattern of 0 in Y_{t-2^J}, which is the same as the one at time t, this function ϕ can use the fact that t is an $\ell_{j,k}$ or not.

DEFINITION 4.2 (Cylindrical function). Such a function ϕ of the past is cylindrical on the neighborhood v if $\phi(Y_{-\infty}^{t-1})$ only depends on any ith coefficient of Y_{t-s}, if $(i, s) \in v$. Informally, it means that $\phi(Y_{-\infty}^{t-1})$ only depends on the data that are in the neighborhood v of Y_t. Its maximal lag is given by $\max\{d / \exists i \in \mathcal{I}, (i, d) \in v\}$.

Let us give some illustrative examples of such functions. Let us fix some $t \in \mathbb{Z}$. Since we are approximating (2) and/or (4) by a linear combination of these functions, we also give an interpretation of the multiplicative coefficient in the expansion.

1. $\phi(.) = 1$ is cylindrical on the empty neighborhood $v = \emptyset$. Its maximal lag is therefore 0. The coefficient in front of this ϕ is typically a candidate to be a ν^m (if we approximate (2)) or $\mu^{p,j}$ (if we approximate (4)).

2. $\phi(Y_{-\infty}^{t-1}) = S_{t-d}^m$, for some index $m \in 1:M$ and lag $d > 0$. This ϕ is cylindrical on $v = \{(m,d)\}$. The coefficient in front of this ϕ is typically a candidate to be a $h_m^{m'}(d)$ (if we approximate (2)) or $h_m^{p,j}(d)$ (if we approximate (4)).
3. For $j > 0$, one can also represent the last coefficient of the jth rhythm on the pth LFP before time t by
$$\phi(Y_{-\infty}^{t-1}) = \sum_{k\in\mathbb{Z}} W_{j,k}^p \mathbb{1}_{t-2^{J+1-j} \le \ell_{j,k} < t} = \sum_{s=1}^{2^{J+1-j}} W_{t-s}^{p,j},$$
since in both sums, only one term is not null. This ϕ is cylindrical on $v = \{(p,j)\} \times 1 : 2^{J+1-j}$. The coefficient in front of this ϕ is typically a candidate to be the constant value of $h_{p,j}^{m'}(d)$ for $d \in 1:2^{J+1-j}$(if we approximate (2)) or $h_{p,j}^{p',j'}(d)$ (if we approximate (4)). This coefficient corresponds to the dependency w.r.t. to the previous rhythm.
4. One can also want to look at the first half-period, which is especially useful for phase lock phenomena. In this case,
$$\phi(Y_{-\infty}^{t-1}) = \sum_{s=1}^{2^{J-j}} W_{t-s}^{p,j},$$
is the last non null $W_{j,k}^p$ if it appears less than half a period away and 0 otherwise. It is cylindrical on $v = \{(p,j)\} \times 1 : 2^{J-j}$.
5. One can also consider the second half period with
$$\phi(Y_{-\infty}^{t-1}) = \sum_{s=2^{J-j}+1}^{2^{J+1-j}} W_{t-s}^{p,j},$$
which is the last non null $W_{j,k}^p$ if it appears in the second half period and 0 otherwise.
6. In Cases 3-5, one can replace $W_{t-s}^{p,j}$ by masked version, for instance $\kappa_w(W_{t-s}^{p,j})$, $\kappa^1(W_{t-s}^{p,j})$ or $\kappa^2(W_{t-s}^{p,j})$, depending if the function ϕ is there to explain spikes or wavelet coefficients. The use of masks might make the linear approximation closer to (2) or (4). Also the oracle inequality proved in the sequel, only works for bounded dictionaries, which means that we need to use at least mask of the type κ_w even for the wavelet part. One can also take $|\kappa_w(W_{t-s}^{p,j})|$ to have access to (an approximation of) the power of the rhythm itself.
7. To avoid having too many coefficients to estimate, one can also want to group several times or cycles in one single ϕ. For instance, with spikes, one might take, for fixed positive lags $L < R$
$$\phi(Y_{-\infty}^{t-1}) = \sum_{s=L}^{R} S_{t-s}^m,$$
which is in fact the number of spikes in the window $t-R : t-L$. This ϕ is cylindrical on the neighborhood $v = \{m\} \times L : R$. Its corresponding coefficient a_ϕ would correspond to an approximation of $h_m^{m'}(d)\mathbb{1}_{L\le d\le R}$ in (2) or of $h_m^{p,j}(d)\mathbb{1}_{L\le d\le R}$ in (4) by a constant value $a_\phi \mathbb{1}_{L\le d\le R}$.
8. The same thing can be done on rhythms, except that it makes more sense to take multiple of the periods. So let $L > R$ be fixed positive integers and define for instance
$$\phi(Y_{-\infty}^{t-1}) = \sum_{s=R2^{J-j+1}+1}^{L2^{J-j+1}} W_{t-s}^{p,j},$$

which is the sum of all the $L-R$ coefficients of the jth rhythm of the pth LFP before time $t-R2^{J-j+1}$. This is cylindrical on $\{(p,j)\}\times R2^{J-j+1}+1:L2^{J-j+1}$. This might be especially useful if we think that large oscillations in a fast rhythm for a certain time might trigger a large oscillation at a slower rhythm a few cycles later. This could make also sense if we replace $W^{p,j}_{t-s}$ by $|W^{p,j}_{t-s}|$, to only count power and discard phase cancellations between the previous cycles.

9. Finally, and especially for influence of spikes on LFP, one might want to aggregate over all neurons, so that one could consider

$$\phi(Y_{-\infty}^{t-1})=\sum_{m=1}^{M}\sum_{s=L}^{R}S^m_{t-s},$$

which the total number of spikes produced on all neurons that have been registered, in the window $t-R:t-L$. The corresponding coefficient measures the excitation effect of a population of neurons on a particular rhythm. In the real data application (see Section 6), we will in particular aggregate over all the neurons in a given brain region.

With the previous examples, it is easy to see that formula (2) and (4) can be seen as fixed linear combination of elementary cylindrical functions (Cases 1 to 6 typically). Moreover, if we want to approximate the interaction functions by piecewise constant functions over a coarser grid than the one given by t, we might want to use the grouped version (Case 7 and 8 for instance). One might also want to group in particular neurons (Case 9). These dimension reductions (at least the ones on time) are common in the Hawkes setting [32], much less in the MVAR one.

In the sequel, we are going to fix the index i we are looking at (either a neuron m or a rhythm j of a LFP p), and fix a finite dictionary Φ^i of cylindrical functions to explain the conditional expectation $f^i_t(Y_{-\infty}^{t-1})=\mathbb{E}(Y_{t,i}|\mathcal{F}_{t-1})$, where $Y_{t,i}$ is the ith index of Y_t.

Note that f^i_t is as g_t in Section 3, a periodic function of the past $Y_{-\infty}^{t-1}$ and that in the most general case $f^i_{t+2^J}=f^i_t$. In particular, f^i_t is null when $\delta^i(t)$ is null, where $\delta^i(t)$ is defined by

$$\delta^i(t)=\begin{cases}1 & \text{if } i=m\in 1:M\\ \mathbb{1}_{\exists \ell_{j,k},\ell_{j,k}=t} & \text{if } i=(p,j)\in 1:P\times 0:J.\end{cases}$$

The approximation of $f^i=(f^i_t)_{t\in 0:T}$ on the dictionary Φ^i is therefore given, for a $a\in\mathbb{R}^{\Phi^i}$ by $f^i_a=(f^i_{a,t})_{t\in 0:T}$, with

$$f^i_{a,t}(Y_{-\infty}^{t-1})=\sum_{\phi\in\Phi^i}a_\phi\phi(Y_{-\infty}^{t-1})\delta^i(t).$$

We work with the following random norm on the vectors of functions of the past, $q=(q_t)_{t\in 0:T}$

$$\|q\|^2=\sum_{t\in 0:T}\left[q_t(Y_{-\infty}^{t-1})\right]^2,$$

and associate scalar product $<q,q'>=\sum_{t\in 0:T}q_t(Y_{-\infty}^{t-1})q'_t(Y_{-\infty}^{t-1})$. For intuition, in the sequel note that if q is 2^J periodic, which is the case for f^i and f^i_a, these quantities should be growing with N the number of blocks. If the sequence Y is block stationary, we can even hope by ergodicity that this grows linearly with N.

On the other hand, we also use the classical euclidian norm on the vector $a\in\mathbb{R}^{\Phi^i}$ given by

$$\|a\|^2=\sum_{\phi\in\Phi^i}a_\phi^2$$

and associated scalar product also denoted $<.,.>$. We use the same notation for both norms but note that depending on the objects it has different meaning.

4.2. *Contrast and Lasso penalization.* Let A be the maximal lag of all the cylindrical functions ϕ used in dictionary Φ. Let B be the largest integer such that $B2^J \geq A$. Let us also set N and T two positive integers such that $T = N2^J - 1$

Up to a shift in time of the data, let us assume that we observe the P LFP's and the M spike trains on $-(B+1)2^J : (N+1)2^J - 1$. Thanks to the wavelet transformation [20], it means we that we have access to the $W^p_{j,k}$ for all $\ell_{j,k} \in -A : T$, which means that all $f^i_a(Y^{t-1}_{-\infty})$ are computable for all $i \in \mathcal{I}$ and $t \in 0 : T$.

We consider the following least-square contrast.

DEFINITION 4.3 (Least-square contrast). The least-square contrast is defined by:

$$C(f^i_a) = -2\sum_{t=0}^{T} f^i_{a,t}(Y^{t-1}_{-\infty})Y_{t,i} + \sum_{t=0}^{T}\left[f^i_a(Y^{t-1}_{-\infty})\right]^2 = -2 < f^i_a, Y^i > + \|f^i_a\|^2,$$

with $Y^i = (Y_{t,i})_{t\in 0:T}$. This can be seen as $C(f^i_a) = -2a^\top b_i + a^\top G_i a$, where $a = (a_\phi)_{\phi\in\Phi^i}$ is the vector of unknown coefficients,

$$b_i = \left(\sum_{t=0}^{T}\phi(Y^{t-1}_{-\infty})Y_{t,i}\right)_{\phi\in\Phi^i} \quad \text{and} \quad G_i = \left(\sum_{t=0}^{T}\phi(Y^{t-1}_{-\infty})\phi'(Y^{t-1}_{-\infty})\right)_{\phi,\phi'\in\Phi^i}.$$

We are looking at sparse solutions, this is why we want to solve this weighted Lasso problem:

$$\hat{a}_i = \arg\min_{a\in\mathbb{R}^{\Phi^i}} -2a^\top b_i + a^\top G_i a + \gamma w_i^\top |a|, \tag{8}$$

where $.^\top$ denotes the transpose, w_i is a weight vector, which adapts the penalty to each ϕ (Weighted Lasso), $\gamma > 0$ is a tuning parameter and $|a| = (|a_\phi|)_{\phi\in\Phi^i}$. This will lead us to a sparse reconstructed connectivity graph.

DEFINITION 4.4 (Reconstructed connectivity graph). This connectivity graph has vertices given by $\mathcal{I}$. A directed edge from $i' \in \mathcal{I}$ to $i \in \mathcal{I}$ means that

- there is a non zero $\hat{a}_{i,\phi}$ in the Lasso problem of i,
- for a ϕ cylindrical on a neighborhood v such that there exists some positive lag d for which $(i', d) \in v$.

To evaluate the full connectivity graph between spike trains and rhythms, we are going solve successively all the Lasso problems for each $i \in \mathcal{I}$. For each of these Lasso problems, we might want to tune the dictionary Φ^i and the weights w_i: in particular they depend on the nature of i (neurons or rhythm).

4.3. *Choice of minimal weights d.* In this section, we derive a set of minimal weights for which our method works. All the $w_{i,\phi}$ used in the sequel will be such that $w_{i,\phi} \geq d_{i,\phi}$, where the values of $d_{i,\phi}$ are derived in this section. This choice of minimal weights d_i is based on exponential deviation inequalities, derived from the martingale structure of our model. The first result holds for the spike data and is inspired by similar results in continuous time [21].

LEMMA 4.5. *For any function bounded $\phi : Y^{t-1}_{-\infty} \mapsto \mathbb{R}$, and any index m, let*

$$M^\phi_T = \sum_{t=0}^{T}\phi(Y^{t-1}_{-\infty})\left[S^m_t - \mathbb{E}(S^m_t|\mathcal{F}_{t-1})\right] \quad \textit{and} \quad \hat{V}^\phi_T = \sum_{t=0}^{T}\left[\phi(Y^{t-1}_{-\infty})\right]^2 S^m_t.$$

Then for any $\epsilon, x > 0$,

$$\mathbb{P}\left(M_T^\phi \geq \sqrt{3(1+\epsilon)\hat{V}_T^\phi x} + \left(\frac{1}{3} + \sqrt{6(1+\epsilon)}\right)\|\phi\|_\infty x\right) \leq 2\left(\frac{\log\left(\frac{T+1}{2x}+1\right)}{\log(1+\varepsilon)}+1\right)e^{-x}.$$

As a consequence, for any family Φ of such functions ϕ, with finite cardinal $\#\Phi$, if $T > 1$ and $\alpha \in (0,1)$, the event

$$\Omega_\Phi^\alpha = \{\forall \phi \in \Phi, |M_T^\phi| \leq d_\phi\},$$

with

$$d_\phi = \sqrt{3(1+\epsilon)\hat{V}_T^\phi \log\left(\frac{4\#\Phi\left(\frac{\log(T)}{\log(1+\epsilon)}+1\right)}{\alpha}\right)} + \left[\frac{1}{3}+\sqrt{6(1+\epsilon)}\right]\log\left(\frac{4\#\Phi\left(\frac{\log(T)}{\log(1+\epsilon)}+1\right)}{\alpha}\right)\|\phi\|_\infty,$$

is of probability at least $1-\alpha$. In particular, one can take

$$d_\phi = \sqrt{\frac{9}{2}\hat{V}_T^\phi \log\left(\frac{16\#\Phi\log(T)}{\alpha}\right)} + \frac{10}{3}\log\left(\frac{16\#\Phi\log(T)}{\alpha}\right)\|\phi\|_\infty, \tag{9}$$

LEMMA 4.6. *For any function bounded $\phi : Y_{-\infty}^{t-1} \mapsto \mathbb{R}$, and any index p, j, let*

$$M_T^\phi = \sum_{k/0\leq \ell_{j,k}\leq T} \phi(Y_{-\infty}^{t-1})\left[W_{j,k}^p - \mathbb{E}(W_{j,k}^p|\mathcal{F}_{\ell_{j,k}-1})\right] \quad \textit{and} \quad V_T^\phi = \sigma_{p,j}^2 \sum_{k/0\leq \ell_{j,k}\leq T}\left[\phi(Y_{-\infty}^{t-1})\right]^2.$$

Then for any $x, \epsilon, \eta > 0$,

$$\mathbb{P}\left(M_T^\phi \geq \sqrt{2(1+\epsilon)\left[V_T^\phi + \eta\sigma_{p,j}^2\|\phi\|_\infty^2\right]x}\right) \leq \left(\frac{\log\left(\frac{(T+1)}{\eta}+1\right)}{\log(1+\varepsilon)}+1\right)e^{-x}.$$

As a consequence, for any family Φ of such functions ϕ, with finite cardinal $\#\Phi$, if $T > 1$ and $\alpha \in (0,1)$, the event

$$\Omega_\Phi^\alpha = \{\forall \phi \in \Phi, |M_T^\phi| \leq d_\phi\},$$

with

$$d_\phi = \sqrt{2(1+\epsilon)[V_T^\phi + \eta\sigma_{p,j}^2\|\phi\|_\infty^2]\log\left(\frac{\#\Phi\left(\frac{\log((T+1)\eta^{-1}+1)}{\log(1+\epsilon)}+1\right)}{\alpha}\right)}$$

is of probability at least $1-\alpha$. In particular one can take

$$d_\phi = \sqrt{3[V_T^\phi + 0.5\sigma_{p,j}^2\|\phi\|_\infty^2]\log\left(\frac{36\#\Phi\log(T)}{\alpha}\right)} \tag{10}$$

Note that to use this d_ϕ, we need to know $\sigma_{p,j}^2$, but an upper bound can work as well and we could take in practice an overall estimation of the variance of $W_{j,k}^p$.

4.4. *Oracle inequality.* Before proving oracle inequalities, let us state the property we assume on the matrices G_i's. In the following for a subset $J \subset \Phi^i$ and $x \in \mathbb{R}^{\Phi^i}$, we denote $x_J = (x_\phi)_{\phi \in J}$, $|x_J|_1 = \sum_{\phi \in J} |x_\phi|$, $\|x_J\|^2 = \sum_{\phi \in J} |x_\phi|^2$ and $J^c = \Phi^i \setminus J$. We also note $S(x)$ the support of the vector x that is $S(x) = \{\phi \in \Phi^i / a_\phi \neq 0\}$.

DEFINITION 4.7 (Property $\mathbf{RE}(\kappa, c, s)$ for G_i/N). For κ, c real positive numbers and s positive integer, G_i/N has the Restricted Eigenvalue property $\mathbf{RE}(\kappa, c, s)$, if for all $x \in \mathbb{R}^{\Phi^i}$ and J of cardinal less than s, such that $|x_{J^c}|_1 \leq c|x_J|_1$, we have

$$x^\top G_i x \geq N\kappa \|x_J\|^2.$$

First of all, we divided G_i by N, the number of blocks because in the stationary regime, we can hope that by ergodicity, this quantity converges. This has been proved for instance in the pure Hawkes case [32]. Moreover, by controlling both the distance to this limit and the limit itself, it has been proved (see [32]) that if the models have some fast decaying Kalikow decomposition, with large probability, the G_i/N matrix satisfies $\mathbf{RE}(\kappa, c, s)$ for some parameters κ, c and s. The same argument is not possible here because Kalikow decomposition do not exist for Gaussian processes (up to our knowledge). On the other hand, there has been many works on Gaussian matrices to show under which property they are satisfying $\mathbf{RE}(\kappa, c, s)$ (see [34] and the references therein). We do not know how to prove such properties when the process is a mix of both cases. This is why the following theorem is stated on the event where $\mathbf{RE}(\kappa, c, s)$ holds.

THEOREM 4.8. *Let $i \in \mathcal{I}$ be a fixed index, with associated finite dictionary of cylindrical functions Φ^i and let γ, the tuning parameter, be such that $\gamma > 2$.*

Let us fix some weights w_i such that $w_{i,\phi} \geq d_{i,\phi}$, for all $\phi \in \Phi^i$ where the d_i's and the corresponding $\Omega^\alpha_{\Phi^i}$ are given in Lemma 4.5 or 4.6, depending on the type of i (spike train or rhythm). Let also $\gamma > 2$.

On the event

- *where $\Omega^\alpha_{\Phi^i}$ of probability larger than $1 - \alpha$ holds*
- *and where G_i/N satisfies $\mathbf{RE}(\kappa, c, s)$ with*

$$c = \left(\frac{\gamma + 2}{\gamma - 2}\right) \frac{\max_{\phi \in \Phi^i} w_{i,\phi}}{\min_{\phi \in \Phi^i} w_{i,\phi}},$$

one has that $\hat{a}_i$ defined by (8) satisfies

$$\|f^i_{\hat{a}_i} - f^i\|^2 \leq \inf_{a/\#S(a) \leq s} \left\{ \|f^i_a - f^i\|^2 + \frac{(\gamma + 2)^2}{4\kappa} \sum_{\phi \in S(a)} \frac{w^2_{i,\phi}}{N} \right\}.$$

Moreover if there exists a^*_i such that $f^i = f^i_{a^*_i}$ and $\#S(a^*_i) \leq s$, then

$$|\hat{a}_i - a^*_i|_1 \leq \frac{2\gamma(\gamma + 2)}{2\kappa(\gamma - 2)\sqrt{N}} \sum_{\phi \in S(a^*)} \frac{w^2_{i,\phi}}{\sqrt{N} \min_{\phi \in \Phi^i} w_{i,\phi}}.$$

Note first that if we take $w_{i,\phi} = d_{i,\phi}$, the leading term in $w^2_{i,\phi}$ is of order $\hat{V}^\phi_T$ in Lemma 4.5 and V^ϕ_T in Lemma 4.6, times a logarithmic factor of the form $\log(\#\Phi^i) + \log\log(T)$. In particular, it increases with the number of blocks N and we can hope by ergocidicity that $\hat{V}^\phi_T/N$ and V^ϕ_T/N converges to constant. Also in both cases these terms are estimating a variance. Therefore one can see roughly $w^2_{i,\phi}/N$ as a constant with respect to N up to

logarithmic term. Hence the first equation is an oracle inequality stated on a norm to the square that should increase linearly in N and for which the bias term is counterbalanced by a variance term that should remain roughly constant and proportional to the number of non zero coefficients in a.

The convergence when N increases is more clear in the second equation. If there is indeed a sparse representation of f^i on the dictionary, then one is able by the previous procedure to retrieve this set of coefficients and the error on the coefficients decrease like $[\log(\#\Phi^i) + \log\log(T)]^{1/2}/\sqrt{N}$, which is what one would expect for such convergence results in Lasso setting, except the factor $\log(\log T)$, which comes from a very sharp control of the martingales in our dependent setting.

Note however that we do not need for the previous theorem to assume that the sequence of observations Y is block stationary to prove the theorem. The stationary case just gives us the intuition of the order of magnitude of the several terms.

The constant c in $\mathbf{RE}(\kappa, c, s)$ depends both on γ and the weights. In particular if γ is large or close to 2 or if the weights have a very large amplitude, c becomes very large and $\mathbf{RE}(\kappa, c, s)$ is much less likely to be fulfilled. One could in particular have a smaller c by taking constant weights $w_{i,\phi} = \max_{\phi \in \Phi^i} d_{i,\phi}$, so that $c = (\gamma + 2)/(\gamma - 2)$.

However, in practice the choice of varying weights makes the estimation much more robust to heterogeneity of the inputs and of the dictionary. This kind of comparison on simulation has already been made in [21, 37].

5. Simulations . In this section we test the performance of the previous LASSO estimation procedure on simulated data.

5.1. *Overall simulation plan.* All the following experiments follow the same schedule:

1. A point of the discrete time domain corresponds to a time (in seconds) thanks to the sampling frequency, that we chose to resemble the acquisition of real data. We choose a sampling frequency $f_S = f_{LFP}$ and Υ is the observation time in seconds.
2. We performed 5 different numerical settings, where each time a different sparse true connectivity graph is used. Each setting corresponds therefore to the choice of the number of neurons M and LFP signals, P, as well as the finer scale J. Then each neuron m and LFP rhythm (p, j) is characterized by its spontaneous value ν^m ($\mu^{p,j}$) and by its interactions with the other variables $h^m_\cdot(t)$ ($h^{p,j}_\cdot(t)$). To simplify, we chose each time to be in the case where we can decompose our process on a specific dictionary, presented in the next section. The dictionary has been chosen to mimic patterns of cerebral activity that are known in the literature (see also the examples of Section 2). The parameters' significations are given in Table 1. The values of the parameters which remain constant throughout all simulations are also written in the table, they are of the order of magnitude of classical experiments in neuroscience. (Note: when we set the ν^m in the tables of parameters, we are writing the biology-friendly $\nu^m \cdot f_S$ - in Hz). The remaining values will be made explicit for each experiment.
3. Thanks to equations (2) and (4) we can simulate the evolution of the spiking activity of the neurons and of the wavelet coefficients of the LFPs.
4. The synthetic data is then employed for the estimation of the original connectivity graph. The choice of a dictionary Φ permits the computation of the vectors b_i, the Gram matrices G_i, and of the weights w_i. The latter are chosen for the neuron part as (9) and for the wavelet part as (10), with the slight difference that $\sigma^{p,j}$ is over-estimated by the empirical variance of each of the wavelet coefficients.
5. Finally we studied 3 different methods: Least Squares without penalization, denoted LS, Lasso (8), and Lasso followed Least Square on the support found by the Lasso, denoted L+LS.

Parameter name	Definition
Simu$\# = 100$	number of simulations for a given set of parameters.
$f_S = 1250$	sampling frequency for the spikes (in Hz).
$f_{LFP} = 1250$	sampling frequency for the LFPs (in Hz).
M	number of neurons.
$P = 1$	number of LFP signals.
$J = 10$	finest scale.
$N = 70$	number of blocks of 2^J LFP points (see Section 4.2). The real duration of each block is $2^J/f_{LFP}$ seconds.
$\Upsilon = 57.344$	observation duration, equal to $2^J/f_{LFP} \cdot N$ in seconds.
γ_S	LASSO penalization parameter when estimating the parameters for a spiking probability.
γ_W	LASSO penalization parameter when estimating the parameters for a wavelet coefficient.
ν^m	spontaneous firing rate for neuron m (in Hz).
$\mu^{p,j} = 0$	spontaneous value for LFP p, rhythm j (in mV).
$\sigma^{p,j} = 0.25$	variance of the Gaussian noise $\epsilon^p_{j,k}$ (see 4).
$\gamma_S = 0.6$	penalization parameter (see 8) when estimating the parameters of 2.
$\gamma_W = 0.3$	penalization parameter (see 8) when estimating the parameters of 4.

TABLE 1
Signification of the simulation parameters.

6. The estimation performance is evaluated with the help of various measures. The differences between the two connectivity graphs can be analyzed by counting missing or extra edges, and in general how many times the graph is perfectly reconstructed. Then, the precision of the estimation is assessed by measuring the errors on the spontaneous values and on the coefficients. The signification of all the performance measures we used is given in Table 2 and more precise mathematical definitions of the errors are given in Section 2 of the Supplementary Materials. Note that in the sequel, we give the average values of all the performance measures over the 100 simulations that are made.

Names	Meaning	Names	Meaning
ν_e	total error for the spontaneous firing rates	$c+_{sS}$	# false positive spike-spike parameters
μ_e	total error for the LFP spontaneous values	$c-_{sS}$	# false negative spike-spike parameters
e_{sS}	total error for spike-spike interactions	$c+_{sW}$	# false positive spike-LFP parameters
e_{sW}	total error for spike-LFP interactions	$c-_{sW}$	# false negative spike-LFP parameters
e_{wS}	total error for LFP-spikes interactions	$c+_{wS}$	# false positive LFP-spike parameters
e_{wW}	total error for LFP-LFP interactions	$c-_{wS}$	# false negative LFP-spike parameters
DG	=1 if graph is perfectly reconstructed	$c+_{wW}$	# false positive LFP-LFP parameters
S	# of spontaneous values detected	$c-_{wW}$	# false negative LFP-LFP parameters

TABLE 2
Definition of the various measures of performance.

5.2. *Precise tuning of the dictionary for neuroscience purpose.* Throughout all the simulations we used a dictionary that is specially tuned to detect interaction that are well-known in the biological literature. In particular, we focused on the interactions that can appear among and between the hippocampus and the amygdala in a mouse during a specific sleep phase called "slow wave sleep" (see Section 6 on experimental data). Most of the rhythms j have a special meaning in the neuroscientific literature (see Table 3). Hence we did not look for all possible patterns but only specific interactions that are relevant from a neuroscientific point of view.

5.2.1. *Dictionary for a neuron.* Let us first describe the dictionary to reconstruct the interactions targeting a neuron:

Rhythm name	Slow 1	δ	Δ	θ	Θ	β	γ	Γ	Ripples	N/D
Scaling (j)	1	2	3	4	5	6	7	8	9	10
Frequency (Hz)	0.82	1.64	3.28	6.55	13.11	26.21	52.43	104.86	209.71	419.43

TABLE 3

Discrete set of frequencies employed in the simulations and in the analysis of experimental data, corresponding to a maximum scale $J = 10$ and to an LFP sampling frequency of 1250Hz. As the names of the frequencies generally try to follow the known literature [8], we stress that we are forced to use a discrete set of frequencies and so the names have to be considered as merely indicative.

1. The *Spontaneous firing rate* ν^m corresponds to the coefficient of
$$\phi_1(Y_{-\infty}^{t-1}) = 1 \tag{11}$$
2. *Effect of other neurons.* This function is needed to detect the increase or decrease of the spiking probability due to the firing of other neurons. Given a time $t \in \mathcal{T}$, a lag $r \in \mathbb{N}$ and a bin size $\delta \in \mathbb{R}$, the function counts the spikes of other neurons in the interval $1 + (r-1)\delta f_S : r\delta f_S$.
$$\phi_{r,\delta}^m(Y_{\infty}^{t-1}) = \sum_{1+(r-1)\delta f_S \le s \le r\delta f_S} S_{t-s}^m \tag{12}$$
In the experiments, we took $\{\phi_{r,\delta}^m, m \in 1:M, r \in 1:2\}$ with $\delta = 0.03$ seconds.
3. *LFP Power influence.* Powerful LFP rhythms can increase or decrease the spiking probability. In the slow wave sleep phase theta waves ($j = 4, 5$) are negligible [28], so we avoided the theta power influences on the neurons. For each of the other j ($0:9 \setminus \{4,5\}$), we have added to the dictionary:
$$\phi_{Pow}^{p,j}(Y_{\infty}^{t-1}) = \sum_{s=1}^{\cdot 2^{J-j+1}} \max(|W_{t-s}^{p,j}|, w), \tag{13}$$
which in fact represents the last wavelet coefficient in absolute value, truncated at the fixed positive constant w. In practice we took $w = 2$ and this value is so large that the wavelet coefficients have never been truncated in our experiments.
4. *Phase lock.* Phase lock phenomena are common in the recordings of cerebral activity. For the same j indexes as before, we need to add 2 functions in the dictionary, one for the up phase and one for the down phase:
(14)
$$\phi_{\uparrow}^{p',j'}(Y_{\infty}^{t-1}) = \sum_{s=1}^{2^{J-j}-1} \max(|W_{t-s}^{p,j}|, w)\mathbb{1}_{W_{t-s}^{p,j}>0} + \sum_{s=2^{J-j}}^{2^{J-j+1}-1} \max(|W_{t-s}^{p,j}|, w)\mathbb{1}_{W_{t-s}^{p,j}<0},$$
which in fact is the absolute value of the last wavelet coefficient if t is in the corresponding up phase, and
(15)
$$\phi_{\downarrow}^{p',j'}(Y_{\infty}^{t-1}) = \sum_{s=1}^{2^{J-j}-1} \max(|W_{t-s}^{p,j}|, w)\mathbb{1}_{W_{t-s}^{p,j}<0} + \sum_{s=2^{J-j}}^{2^{J-j+1}-1} \max(|W_{t-s}^{p,j}|, w)\mathbb{1}_{W_{t-s}^{p,j}>0},$$
which in fact is the absolute value of the last wavelet coefficient if t is in the corresponding down phase.

5.2.2. *Dictionary for a wavelet coefficient..* In addition to the previous functions, we want here to model the *global effect of a neuronal population*, so that instead of using $\phi_{r,\delta}^m$ for

specified neurons m, we grouped them in regions $R_i = \{m \in 1:M | \text{neuron } m \text{ is in region } i\}$ and considered the cumulative effect of all the spikes in a region:

$$\phi^{R_i}_{r,\delta}(Y^{t-1/f_S}_{1/f_S}) = \sum_{m \in R_i} \sum_{1+(r-1)\delta f_S \leq s \leq r\delta f_S} S^m_{t-s}, \tag{16}$$

still with $r \in 1:2$ and $\delta = 0.03$s.

Concerning LFP-LFP interactions, the choice of the dictionary depends on j, because of its neurobiological interpretation.

- For all of the rhythms we use ϕ_1 so that the corresponding coefficient is estimating $\mu^{p,j}$.
- Again for all of them we added $\phi^{R_i}_{r,\delta}$ for $r = 1, 2$ and $\delta = 0.03$ seconds, to leave open the possibility of estimating the effect of the spikes on each LFP frequency.
- To study the effect of the power of LFP frequencies on other frequencies, we employ the dictionary function $\phi^{p',j'}_{Pow}(Y^{t-1}_{\infty})$. We employed tactically this fraction of the dictionary by adding the effect of one cycle of gamma power ($j' = 7, 8$) on Ripples ($j = 9$). Indeed, ripples are bursts of oscillatory activity ranging from 140 to 200 Hz. Therefore they should be mainly caught by the rhythm $j = 9$ but it is also possible that its effect leaks to lower, close frequencies.
- Finally, we left the highest number of degrees of freedom for the power-phase modulations. Since by definition of this phenomenon slower frequencies target faster frequencies [8] for all $j \in \{0:9\}/\{4,5\}$ and all $j' < j$ we added the two dictionary functions $\phi^{p',j'}_{\uparrow}(Y^{t-1}_{\infty})$ and $\phi^{p',j'}_{\downarrow}(Y^{t-1}_{\infty})$.

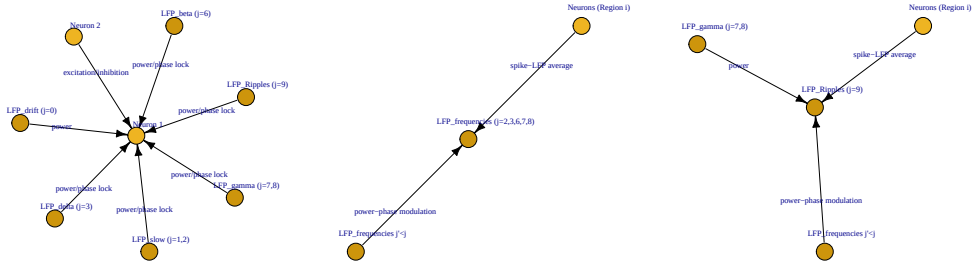


FIG 5. *Possible interactions between variables. Each interaction corresponds to the coefficient of some dictionary function. Left: possible interactions targeting a neuron. "excitation/inhibition": see* (12)*; "power": see* (13)*; "phase lock": see* (14)(15)*. Center/right: possible interactions targeting an LFP coefficient. "Spike-LFP average": see* (16)*; "power": see* (13)*; "power-phase modulation": see* (14)(15)*.*

5.3. *Results.* The description of the parameters of each experiment, as well as the corresponding connectivity graph and the average performance measures are reported in Figures 6 and 7. The first 4 experiments correspond to the classic examples reported in Section 2, except that the parameters are smaller to make the problem more difficult to see with bare eyes. The last experiment has been inspired by our results on real data (see Section 6).

1. *Excitation*: The spiking of Neuron 2 prompts the spiking of Neuron 1.
2. *Phase Lock:* The gamma ($j = 7$) up phase of the LFP signal excites Neuron 1.
3. *LFP power-phase modulation* Gamma ($j = 7$) coefficients are diminished in the down phase of the Delta ($j = 3$) frequency.

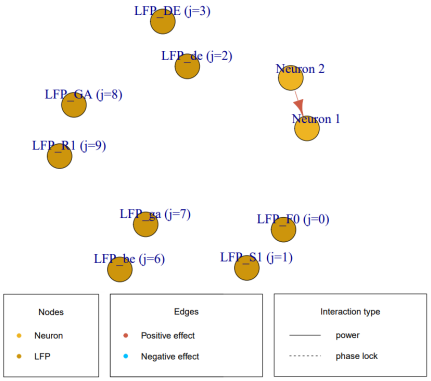
4. *Spike-LFP average* In this experiment we mimic a negative effect of neuronal spiking onto a specific LFP rhythm.
5. *Close to real data.* In the final simulation we estimate a complex connectivity graph, which is the fitting result of the cerebral sleep activity of a rat (see Section 6). The difference is that for simplicity the spontaneous values and interaction strengths are restricted to few possible values.

M	ν^1	ν^2
2	3	3

coefficients of $\phi^2_{r,\delta}, r = 1:2$, target is Neuron 1
(0, 0.03)

Exp1	LS	LASSO	L + LS
ν_e	6.370	2.132	**0.292**
μ_e	0.315	**0**	**0**
e_{sS}	0.033	0.003	**0.001**
e_{sW}	0.317	**0**	**0**
e_{wS}	0.356	**0**	**0**
e_{wW}	1.349	**0**	**0**
DG	0/1	**1**/1	**1**/1
S	**12**/12	**12**/12	**12**/12

Exp1	LS	LASSO	L + LS
$c+_{sS}$	7/7	**0**/7	**0**/7
$c-_{sS}$	**0**/1	**0**/1	**0**/1
$c+_{sW}$	20/20	**0**/20	**0**/20
$c-_{sW}$	**0**/0	**0**/0	**0**/0
$c+_{wS}$	44/44	**0**/44	**0**/44
$c-_{wS}$	**0**/0	**0**/0	**0**/0
$c+_{wW}$	44/44	**0**/44	**0**/44
$c-_{wW}$	**0**/0	**0**/0	**0**/0

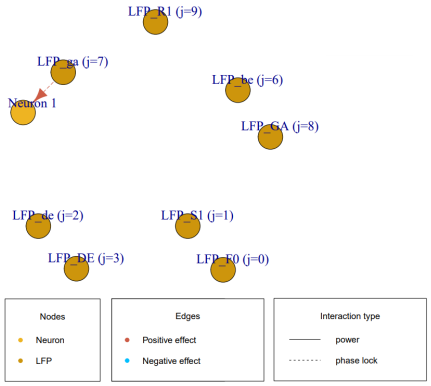


M	ν^1
1	2

coefficient of $\phi^{1,7}_{\uparrow}$, target is Neuron 1
0.03

Exp2	LS	LASSO	L + LS
ν_e	10.742	4.053	**0.248**
μ_e	0.336	**0**	**0**
e_{sS}	0.005	**0**	**0**
e_{sW}	0.417	**0**	**0**
e_{wS}	0.155	0.022	**0.001**
e_{wW}	1.415	**0**	**0**
DG	0/1	**1**/1	**1**/1
S	**11**/11	11/11	11/11

Exp2	LS	LASSO	L + LS
$c+_{sS}$	2/2	**0**/2	**0**/2
$c-_{sS}$	**0**/0	**0**/0	**0**/0
$c+_{sW}$	20/20	**0**/20	**0**/20
$c-_{sW}$	**0**/0	**0**/0	**0**/0
$c+_{wS}$	21/21	**0**/21	**0**/21
$c-_{wS}$	**0**/1	**0**/1	**0**/1
$c+_{wW}$	44/44	**0**/44	**0**/44
$c-_{wW}$	**0**/0	**0**/0	**0**/0



M	ν^1
1	5

coefficient of $\phi^{1,3}_{\downarrow}$, target is LFP 1 frequency 7
-1

Exp3	LS	LASSO	L + LS
ν_e	6.598	1.088	**0.287**
μ_e	0.289	0.071	**0.006**
e_{sS}	0.011	**0**	**0**
e_{sW}	0.554	**0**	**0**
e_{wS}	0.150	**0**	**0**
e_{wW}	1.582	0.523	**0.022**
DG	0/1	**1**/1	**1**/1
S	**11**/11	11/11	11/11

Exp3	LS	LASSO	L + LS
$c+_{sS}$	2/2	**0**/2	**0**/2
$c-_{sS}$	**0**/0	**0**/0	**0**/0
$c+_{sW}$	20/20	**0**/20	**0**/20
$c-_{sW}$	**0**/0	**0**/0	**0**/0
$c+_{wS}$	22/22	**0**/22	**0**/22
$c-_{wS}$	**0**/0	**0**/0	**0**/0
$c+_{wW}$	43/43	**0**/43	**0**/43
$c-_{wW}$	**0**/1	**0**/1	**0**/1

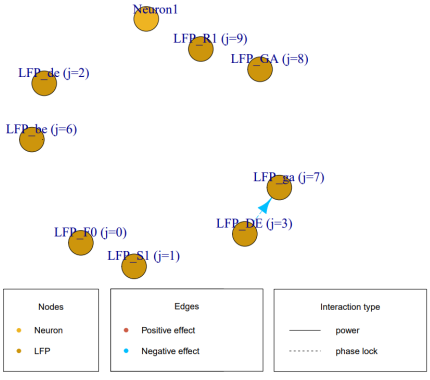
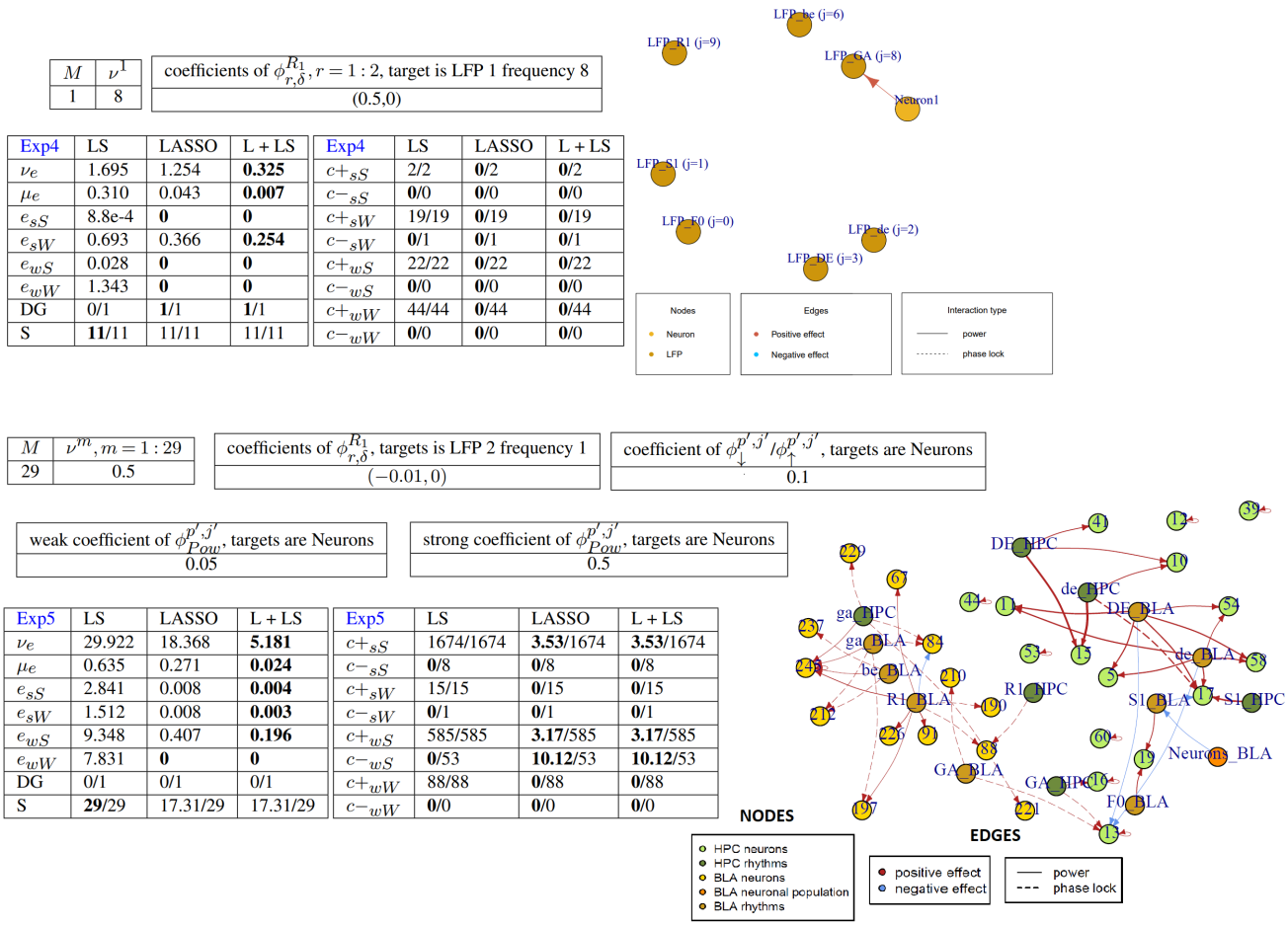


FIG 6. *Performance evaluation for settings 1:3. In each figure, up left: specific parameters for the setting; down left: average errors over 100 simulations; right: true connectivity graph.*



M	ν^1
1	8

coefficients of $\phi^{R_1}_{r,\delta}, r = 1:2$, target is LFP 1 frequency 8: (0.5,0)

Exp4	LS	LASSO	L + LS
ν_e	1.695	1.254	**0.325**
μ_e	0.310	0.043	**0.007**
e_{sS}	8.8e-4	**0**	**0**
e_{sW}	0.693	0.366	**0.254**
e_{wS}	0.028	**0**	**0**
e_{wW}	1.343	**0**	**0**
DG	0/1	**1**/1	**1**/1
S	**11**/11	11/11	11/11

Exp4	LS	LASSO	L + LS
$c+_{sS}$	2/2	**0**/2	**0**/2
$c-_{sS}$	**0**/0	**0**/0	**0**/0
$c+_{sW}$	19/19	**0**/19	**0**/19
$c-_{sW}$	**0**/1	**0**/1	**0**/1
$c+_{wS}$	22/22	**0**/22	**0**/22
$c-_{wS}$	**0**/0	**0**/0	**0**/0
$c+_{wW}$	44/44	**0**/44	**0**/44
$c-_{wW}$	**0**/0	**0**/0	**0**/0

M	$\nu^m, m = 1:29$
29	0.5

coefficients of $\phi^{R_1}_{r,\delta}$, targets is LFP 2 frequency 1: $(-0.01, 0)$

coefficient of $\phi^{p',j'}_{\downarrow}/\phi^{p',j'}_{\uparrow}$, targets are Neurons: 0.1

weak coefficient of $\phi^{p',j'}_{Pow}$, targets are Neurons: 0.05

strong coefficient of $\phi^{p',j'}_{Pow}$, targets are Neurons: 0.5

Exp5	LS	LASSO	L + LS
ν_e	29.922	18.368	**5.181**
μ_e	0.635	0.271	**0.024**
e_{sS}	2.841	0.008	**0.004**
e_{sW}	1.512	0.008	**0.003**
e_{wS}	9.348	0.407	**0.196**
e_{wW}	7.831	**0**	**0**
DG	0/1	0/1	0/1
S	**29**/29	17.31/29	17.31/29

Exp5	LS	LASSO	L + LS
$c+_{sS}$	1674/1674	**3.53**/1674	**3.53**/1674
$c-_{sS}$	**0**/8	**0**/8	**0**/8
$c+_{sW}$	15/15	**0**/15	**0**/15
$c-_{sW}$	**0**/1	**0**/1	**0**/1
$c+_{wS}$	585/585	**3.17**/585	**3.17**/585
$c-_{wS}$	**0**/53	**10.12**/53	**10.12**/53
$c+_{wW}$	88/88	**0**/88	**0**/88
$c-_{wW}$	**0**/0	**0**/0	**0**/0

FIG 7. *Performance evaluation for settings 4:5. In each figure, up left: specific parameters for the setting; down left: average errors over 100 simulations; right: true connectivity graph.*

We can see that as expected, the L+LS method is the one giving the smallest error. More importantly, note that with less than 1 mn of observation, the method is able to recover perfectly the simplest graphs. For the close-to-real data, for which the graph is much more complex, in less than one minute of observation, the estimation method is able to have about 0.5% of false positive and 19% of false negative for the LFP $\rightarrow$ Spikes interactions, whereas the other interactions are most of the time perfectly reconstructed. As explained in more detail in Section 6, we have almost an hour of recordings per condition and we are therefore quite confident that we make almost no false positives, whereas the percentage of false negatives should be smaller (if the model is adequate for these data).

6. Experimental Data . After the numerical validation of the LASSO procedure, let us test its capacity on real data.

The objective of this section is to recover, from a high-quality dataset, relations between neuronal populations and LFP which are well documented in the literature, to see the ability of our procedure to screen complex data and detect automatically relevant interactions.

The data comes from the Buzsaki Lab dataset [33] and consists of the neural activity of a rat, specifically the recording of spikes and LFP belonging to two different brain regions: the hippocampus (HPC) and the basolateral amygdala (BLA). The neural data have been

recorded during the training on a spatial aversive task in which a specific location is associated with an aversive stimulus, and during sleep episodes before and after training. The experiment is explained in detail in [17]. The data presented here come from the spike-LFP recording during a particular sleep phase called non-REM sleep or "slow wave sleep", characterized by widespread slow-frequency (0,5-4 Hz) oscillations in the cortex and ripples in the hippocampus.

In this framework, we treated the data as follows:

1. We isolated the section of the experiment that we wanted to analyze. We considered the recordings performed before (*"prerun"*) and after (*"postrun"*) the aversive spatial task in Rat08 on the 6th day of the experiment. As the rat switched multiple times between slow wave sleep, REM sleep, and wakefulness, we isolated and then joined the segments of slow wave sleep data.
2. We want to infer the functional connectivity via the above L+LS (Lasso + Least Squares Step) procedure, with the same set-up as in the simulations, that is considering $J = 10$ as the finest scale and employing the same dictionary of functions. We decided to use the same penalization parameters used in the Simulations, that is $\gamma_S = 0.6$ and $\gamma_W = 0.3$.
3. Since the LFP sampling frequency of the dataset is 1250 Hz, every block of wavelet coefficients has a duration of $1024/1250 = 0.8192$ seconds. Because of the different duration of each sleep period, we decided to consider 74 segments of $N = 70$ wavelet blocks for the prerun slow wave sleep and 72 segments of $N = 70$ wavelet blocks for the postrun slow wave sleep. This leads to an observation time of $\Upsilon = 57.344$ seconds for every segment and to aggregated observation times of $\approx$ 1h10m40s for the prerun sleep and $\approx$ 1h8m45s for the postrun sleep. The longer the observation time, the higher is the computational cost but we gain in estimator performance.
4. From this time interval we isolate the spikes and the LFP signals coming from the two brain regions of interest : the hippocampus (HPC) and baso-lateral amygdala (BLA).

The resulting connectivity graphs are displayed in Figure 8. We presented only the vertices for which an edge exists to simplify the reading, but in total 156 neurons (with two global nodes for each neuronal population) and 10*2 LFP rhythms (10 rhythms per region)are used for the computation of the model, that is the complete graphs would be on 176 nodes.

How to read the graphs. The hippocampus (HPC) neurons and LFPs are shown in light and dark green, respectively. The BLA neurons and LFPs are shown in yellow and brown. BLA population activity appears in orange, as HPC population activity is not estimated as a node of relevance. Positive interactions are shown as red arrows, and negative interactions as blue arrows. The width of the arrow edges represents the strength (in modulus) of the interaction. The style of the arrows (solid/dashed) represents the type of interaction. As a reminder, "Power" means that the power, $|W_t^{p,j}|$, has an influence on the target rhythm or neuron. "Phase locking" means that the phase of the brain oscillation has an influence on the activity of the target neuron(s) or rhythm.

One of the main differences in the connectivity graphs of slow wave sleep preceding vs. following the aversive task is the emergence of a new positive connection between hippocampal ripples (R1-HPC) and a BLA neuron (neuron 88). This finding echoes the results of [17], in which they show that during sleep following the task, there is an increase in the modulation of a small subpopulation of BLA neurons by hippocampal ripples. The graph also shows a higher effect of beta frequency in the BLA on BLA neurons, a new finding that will be explored in more detail along with other connectivity changes within and across the two structures in follow-up works.

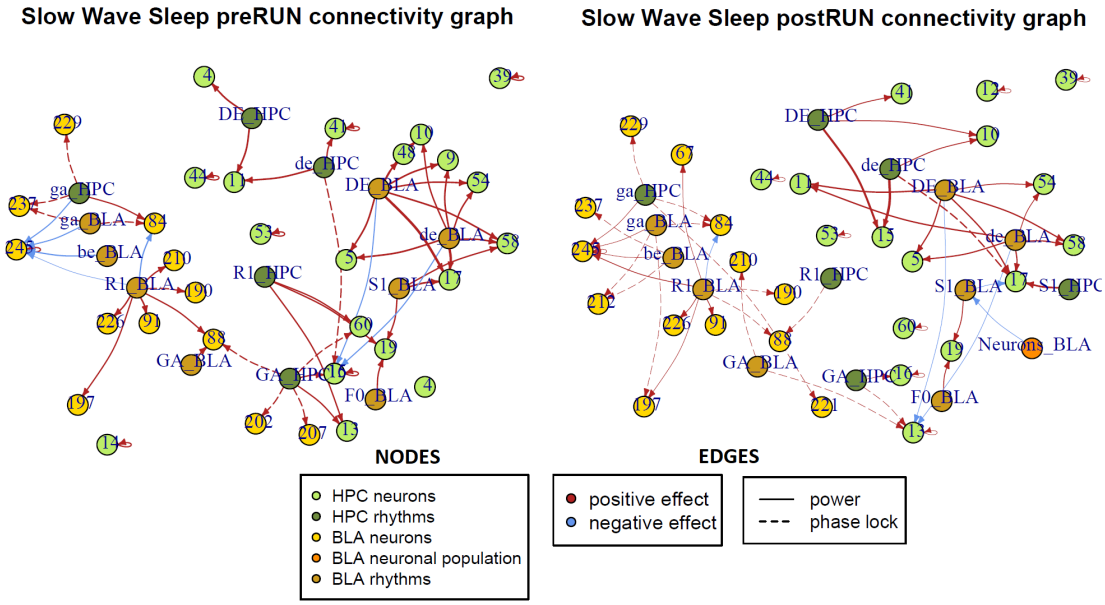


FIG 8. *Fitted connectivity graphs in aggregated segments of slow wave sleep cerebral activity before (left) and after (right) the aversion task. Notice that in the postRUN graph, the Ripples frequency targets amygadala neuron 88, a phenomena described in [17]. Also, after the adversion task we can appreciate a more extensive effect of the BLA beta frequency on BLA neurons.*

7. Conclusion. In conclusion, we proposed a model of spike and LFP rhythm interactions. Conditions under which the model exists in a block stationary version have been made explicit. A weighted Lasso estimator of the interactions has been proposed, for which we proved an oracle inequality. The choice of dictionary is crucial and we decided to fine-tune it to unravel many classical interactions of the literature in neuroscience.

This particular choice of dictionary has the advantage with respect to a more exhaustive dictionary to limit the number of coefficients to estimate. This makes the estimation more robust for small observation times. However, in the future, expanding the dictionaries will potentially allow the unraveling of previously undocumented connections and changes in connectivity. It will also be interesting to further investigate the data of [17] to see if we can establish statistical evidence of the connectivity changes across experimental days.

The choice of the parameters γ_S and γ_W is also difficult. We could use a cross-validation procedure by block to optimize it. However, the code is in R and not optimized yet, which prevented us to implement this computationally costly procedure. This is one of the improvements we would like to make before releasing the code publicly.

Finally, it would be interesting from a mathematical point of view, to see if we can adapt our model to continuous time, mixing point processes in continuous time and continuous Gaussian processes. Moreover the possibility to have a continuum of frequencies would be of interest to fine-tune the detection to meaningful rhythms of the neuroscience literature.

Acknowledgment. This research was supported by the French government, through CNRS (eXplAIn team), the UCA^{Jedi} and 3IA Côte d'Azur Investissements d'Avenir managed by the National Research Agency (ANR-15 IDEX-01 and ANR-19-P3IA-0002), INSERM (Atip-Avenir), directly by the ANR project ChaMaNe (ANR-19-CE40- 0024-02), by the interdisciplinary Institute for Modeling in Neuroscience and Cognition (NeuroMod) and Ville de Paris (Emergence(s)).

SUPPLEMENTARY MATERIAL

Supplementary File

The Supplementary Files contains the missing proofs and the precise formulas for the performance measures.

REFERENCES

[1] ABBASPOURAZAD, H., HSIEH, H.-L. and SHANECHI, M. M. (2019). A multiscale dynamical modeling and identification framework for spike-field activity. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **27** 1128–1138.

[2] ALBO, Z., DI PRISCO, G. V., CHEN, Y., RANGARAJAN, G., TRUCCOLO, W., FENG, J., VERTES, R. P. and DING, M. (2004). Is partial coherence a viable technique for identifying generators of neural oscillations? **90** 318–326.

[3] BACCALA, L. A. and SAMESHIMA, K. (2001). Partial directed coherence: a new concept in neural structure determination. *Biological Cybernetics* **84** 463–474.

[4] BARANOWSKI, R. and FRYZLEWICZ, P. (2020). Multiscale autoregression on adaptively detected timescales. In *Proceedings of the 12th International Conference of the ERCIM WG on Computational and Methodological Statistics (CMStatistics 2019)*.

[5] BONNET, A., DION, C., GINDRAUD, F. and LEMLER, S. (2022). Neuronal Network Inference and Membrane Potential Model using Multivariate Hawkes Processes. *Journal of Neuroscience Methods*.

[6] BRÉMAUD, P. and MASSOULIÉ, L. (1996). Stability of nonlinear Hawkes processes. *The Annals of Probability* **24**. https://doi.org/10.1214/aop/1065725193

[7] BUZSÁKI, G. and WANG, X.-J. (2012). Mechanisms of Gamma Oscillations. *Annual Review of Neuroscience* **35** 203-225. PMID: 22443509. https://doi.org/10.1146/annurev-neuro-062111-150444

[8] BUZSÁKI, G. (2009). *Rhythms of The Brain* xiv, 448 p. https://doi.org/10.1093/acprof:oso/9780195301069.001.0001

[9] CARSTENSEN, L., SANDELIN, A., WINTHER, O. and HANSEN, N. R. (2010). Multivariate Hawkes process models of the occurrence of regulatory elements. *BMC bioinformatics* **11** 1–19.

[10] COHEN, A., DAUBECHIES, I. and FEAUVEAU, J.-C. (1992). Biorthogonal bases of compactly supported wavelets. *Communications on pure and applied mathematics* **45** 485–560.

[11] DE LEON, A. R. and CHOUGH, K. C. (2013). *Analysis of mixed data: methods & applications*. CRC Press.

[12] DION, C. and LEMLER, S. (2020). Nonparametric drift estimation for diffusions with jumps driven by a Hawkes process. *Statistical Inference for Stochastic Processes*. https://doi.org/10.1007/s11203-020-09213-5

[13] DOUKHAN, P. and WINTENBERGER, O. (2007). Weakly dependent chains with infinite memory. https://doi.org/10.48550/ARXIV.0712.3231

[14] FAES, L., ERLA, S. and NOLLO, G. (2012). Measuring Connectivity in Linear Multivariate Processes: Definitions, Interpretation, and Practical Analysis. *Computational and Mathematical Methods in Medicine* **ID 140513**.

[15] FRESE, T., BOUMAN, C. A. and SAUER, K. (2002). Adaptive wavelet graph model for Bayesian tomographic reconstruction. *IEEE Transactions on Image Processing* **11** 756–770.

[16] GALVES, A. and LÖCHERBACH, E. (2015). Modeling networks of spiking neurons as interacting processes with memory of variable length. https://doi.org/10.48550/ARXIV.1502.06446

[17] GIRARDEAU, G., INEMA, I. and BUZSÁKI, G. (2017). Reactivations of emotional memory in the hippocampus–amygdala system during sleep. *Nature neuroscience* **20** 1634–1642.

[18] GONG, X., LI, W. and LIANG, H. (2019). Spike-field Granger causality for hybrid neural data analysis. *Journal of neurophysiology* **122** 809–822.

[19] GRANGER, C. W. J. (1980). Testing for causality: A personal viewpoint. *Journal of Economic Dynamics and Control* **2** 329-352. https://doi.org/10.1016/0165-1889(80)90069-X

[20] GUBNER, J. A. and CHANG, W.-B. (1995). Wavelet transforms for discrete-time periodic signals. *Signal Processing* **42** 167–180.

[21] HANSEN, N. R., REYNAUD-BOURET, P. and RIVOIRARD, V. (2015). Lasso and probabilistic inequalities for multivariate point processes. *Bernoulli* **21**. https://doi.org/10.3150/13-bej562

[22] JENSEN, O. (2005). Reading the hippocampal code by theta phase-locking. *Trends in Cognitive Sciences* **9** 551-553. https://doi.org/10.1016/j.tics.2005.10.003

[23] KOBAYASHI, R., KURITA, S., KURTH, A., KITANO, K., MIZUSEKI, K., DIESMANN, M., RICHMOND, B. J. and SHINOMOTO, S. (2019). Reconstructing neuronal circuitry from parallel spike trains. *Nature communications* **10** 4468.

[24] KREISS, A., MAMMEN, E. and POLONIK, W. (2019). Nonparametric inference for continuous-time event counting and link-based dynamic network models. *Electronic Journal of Statistics* **13** 2764–2829.
[25] KWON, H.-H., LALL, U. and KHALIL, A. F. (2007). Stochastic simulation model for nonstationary time series using an autoregressive wavelet decomposition: Applications to rainfall and temperature. *Water Resources Research* **43**. https://doi.org/10.1029/2006WR005258
[26] LAMBERT, R. C., TULEAU-MALOT, C., BESSAIH, T., RIVOIRARD, V., BOURET, Y., LERESCHE, N. and REYNAUD-BOURET, P. (2018). Reconstructing the functional connectivity of multiple spike trains using Hawkes models. *Journal of Neuroscience Methods* **297** 9-21.
[27] LEPAGE, K. Q., GREGORIOU, G. G., KRAMER, M. A., AOI, M., GOTTS, S. J., EDEN, U. T. and DESIMONE, R. (2013). A procedure for testing across-condition rhythmic spike-field association change. *Journal of neuroscience methods* **213** 43–62.
[28] LEVENSTEIN, D., BUZSÁKI, G. and RINZEL, J. (2019). NREM sleep in the rodent neocortex and hippocampus reflects excitable dynamics. *Nature communications* **10** 2478.
[29] LISMAN, J. and BUZSÁKI, G. (2008). A neural coding scheme formed by the combined function of gamma and theta oscillations. *Schizophrenia bulletin* **34** 974–980.
[30] LISMAN, J. and JENSEN, O. (2013). The Theta-Gamma Neural Code. *Neuron* **77** 1002-16. https://doi.org/10.1016/j.neuron.2013.03.007
[31] MURTAGH, F., STARCK, J. L. and RENAUD, O. (2004). On neuro-wavelet modeling. *Decision Support Systems* **37** 475-484. Data mining for financial decision making. https://doi.org/10.1016/S0167-9236(03)00092-7
[32] OST, G. and REYNAUD-BOURET, P. (2018). Sparse space-time models: Concentration Inequalities and Lasso. https://doi.org/10.48550/ARXIV.1807.07615
[33] PETERSEN, P. C. and HERNANDEZ, G. MICHELLE ANDBUZSÁKI (2020). The Buzsaki Lab Databank – Public electrophysiological datasets from awake animals. http://doi.org/10.5281/zenodo.4307883
[34] RASKUTTI, G., WAINWRIGHT, M. J. and YU, B. (2010). Restricted eigenvalue properties for correlated gaussian designs. *Journal of Machine Learning Research* **11** 2241–2259.
[35] SONG, C. Y., HSIEH, H.-L., PESARAN, B. and SHANECHI, M. M. (2022). Modeling and inference methods for switching regime-dependent dynamical systems with multiscale neural observations. *Journal of Neural Engineering* **19** 066019.
[36] TELEŃCZUK, B., DEHGHANI, N., LE VAN QUYEN, M., CASH, S. S., HALGREN, E., HATSOPOULOS, N. G. and DESTEXHE, A. (2017). Local field potentials primarily reflect inhibitory neuron activity in human and monkey cortex. *Scientific reports* **7** 40211.
[37] VAN DE GEER, S., BÜHLMANN, P. and ZHOU, S. (2011). The adaptive and the thresholded Lasso for potentially misspecified models (and a lower bound for the Lasso). *Electron. J. Stat.* **5** 688–749.
[38] VINCK, M., BATTAGLIA, F., WOMELSDORF, T. and PENNARTZ, C. (2011). Improved measures of phase-coupling between spikes and the Local Field Potential. *Journal of computational neuroscience* **33** 53-75. https://doi.org/10.1007/s10827-011-0374-4
[39] WANG, C. and SHANECHI, M. M. (2019). Estimating multiscale direct causality graphs in neural spike-field networks. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **27** 857–866.
[40] WILLSKY, A. S. (2002). Multiresolution Markov models for signal and image processing. *Proceedings of the IEEE* **90** 1396-1458. https://doi.org/10.1109/JPROC.2002.800717
[41] ZHANG, L., WU, X., PAN, Q. and ZHANG, H. (2004). Multiresolution modeling and estimation of multisensor data. *IEEE transactions on Signal Processing* **52** 3170–3182.
[42] ZHANG, Y., LIU, B., JI, X. and HUANG, D. (2017). Classification of EEG signals based on autoregressive model and wavelet packet decomposition. *Neural Processing Letters* **45** 365–378.
[43] ZHU, Y. and WANG, R. (2015). Research on phase synchronization with spike-LFP coherence analysis. *Neurocomputing* **168** 655–660.