INSTITUT DE FRANCE
Académie des sciences

Comptes Rendus
Mathématique

Patricia Reynaud-Bouret, Alexandre Muzy and Ingrid Bethus

Towards a mathematical definition of functional connectivity

Volume 359, issue 4 (2021), p. 481-492

Published online: 17 June 2021

https://doi.org/10.5802/crmath.190

Les Comptes Rendus. Mathématique *sont membres du Centre Mersenne pour l'édition scientifique ouverte*
www.centre-mersenne.org
e-ISSN : 1778-3569

Comptes Rendus
Mathématique
2021, 359, nº 4, p. 481-492
https://doi.org/10.5802/crmath.190

Statistics / *Statistiques*

Towards a mathematical definition of functional connectivity

Patricia Reynaud-Bouret[*, a], **Alexandre Muzy**[b] **and Ingrid Bethus**[c]

[a] Université Côte d'Azur, CNRS, LJAD, France

[b] Université Côte d'Azur, CNRS, I3S, France

[c] Université Côte d'Azur, CNRS, IPMC, France

E-mails: patricia.reynaud-bouret@univ-cotedazur.fr, alexandre.muzy@univ-cotedazur.fr, ingrid.bethus@univ-cotedazur.fr

Abstract. Functional connectivity is a neurobiological notion, informally stating that there would be a strong dependence between neurons and that this dependence might be useful in understanding the way the brain encodes stimuli, programs actions, etc. However, in practice such strong dependencies are often reconstructed via Hawkes processes based on an amazingly small number of neurons, because of the very scarce observation of this very complex and huge network. We derive new simple equations, which explain how the ideal Hawkes reconstruction is linked to the covariance between the observed neurons. These equations help us in particular to understand what the Hawkes reconstruction does in two settings, synchronization and classical point process asymptotics. Moreover they might help us to also understand what is qualitatively happening at the scale of the huge unobserved network, paving the path for a possible mathematical definition of functional connectivity.

Résumé. La connectivité fonctionnelle est une notion neurobiologique, qui affirme informellement qu'il y aurait une forte dépendance entre neurones et que cette dépendance pourrait être utilisée pour comprendre comment le cerveau encode les stimuli, programme les actions, etc. Cependant, en pratique, ces fortes dépendances sont souvent reconstruites, grâce aux processus de Hawkes, sur un nombre incroyablement faible de neurones, parce que l'observation du réseau sous-jacent est excessivement partielle. Nous prouvons de nouvelles équations qui expliquent comment la reconstruction idéale par processus de Hawkes est liée à la covariance entre neurones observés. Ces équations nous aident à comprendre ce que fait exactement la reconstruction par processus de Hawkes dans deux cadres asymptotiques, la synchronization et le cadre classique des processus ponctuels. De plus, elles pourraient nous permettre de comprendre qualitativement ce qui se passe dans l'immense réseau non observé, ouvrant la voie à une possible définition mathématique de la connectivité fonctionnelle.

2020 Mathematics Subject Classification. 62H99, 62M09, 62P10.

Funding. This work was supported by the French government, through the UCAJedi and 3IA Côte d'Azur Investissements d'Avenir managed by the National Research Agency (ANR-15- IDEX-01 and ANR-19-P3IA-0002), by the CNRS through the "Mission pour les Initiatives Transverses et Interdisciplinaires" (Projet DYNAMO, "APP Modélisation du Vivant") and by the interdisciplinary Institute for Modeling in Neuroscience and Cognition (NeuroMod) of the Université Côte d'Azur.

Manuscript received 13th February 2021, accepted 14th February 2021.

[*] Corresponding author.

1. Introduction

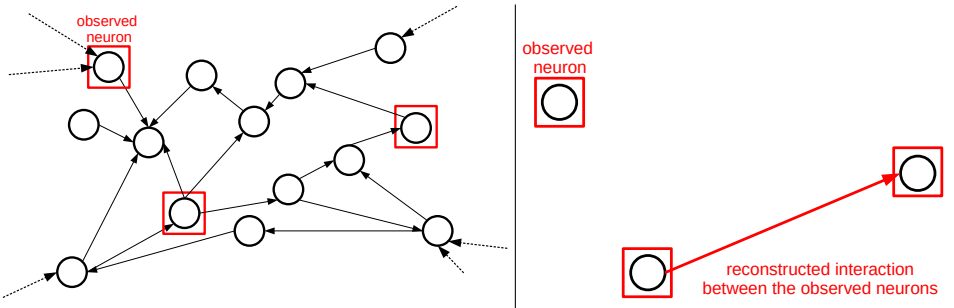
The brain is constituted of several types of cells, including excitable cells, called neurons, which are able to produce electrical signals called action potentials (discretized as spikes). Cognition is the result of millions of neurons which are working together in a fine orchestration. Many experiments have shown that the electrical activity of neurons can depend strongly on the activity of other neurons, in particular through synchronization [12, 20, 31]. But does synchronisation require that the neurons are physically connected to each other? More generally, is it possible for non connected neurons to strongly depend on each other? Can we explain it from a mathematical point of view?

In biology, the synchronization of neurons leads to the concept of *neuronal assemblies*. Historically, Donald Hebb assumed that an assembly consists of neurons 'wiring together and firing together' to achieve a behavior at individual level [16]. At this time, Hebb did not have the experimental apparatus to verify his assumption. Once neuroscientists had the experimental capacity to record the electrical activity of particular neurons *in vivo* (i.e., during the behavior of animals), they validated statistically the firing part assertion - not the wiring one [11]. Assemblies were then populations of neurons whose high dependency may code for stimulus, action or anticipation [28, 29, 32].

More recently a notion has emerged in neuroscience: the *functional connectivity* [10]. This can be thought as highly dependent brain areas thanks to functional neuroimaging approaches such as fMRI, but more microscopically, this pools down to say that neurons may be highly dependent and form a network of strong interactions that can evolve in time, thanks to external stimuli, for instance. This statistical dependence between neurons does not mean necessarily that neurons are wired, that is, physically connected. Indeed, dependent neurons can be indirectly connected. The statistical dependence or graph of dependence obtained may therefore be part of the neural code and used for decoding. However, the significance of this dependence is very hard to establish because there are few data in the recordings. Indeed, at least at the microscopic neuronal scale, the recordings of electrical action potentials are very scarce.

Let us detail a bit more how classical neuronal recordings are made *in vivo* to understand the magnitude of this scarcity. Most *in vivo* experiences in electrophysiology are at the extracellular level. The electrodes implanted in the extracellular medium are recording a difference of electrical potential. In particular, they are sensible to the action potentials emitted by nearby neurons, but the extracellular spike amplitude decreases rapidly as a function of distance from the neuron [17]. It is usual *in vivo* to record the precise activity dynamics of a few tens of neurons with apparatus called tetrodes [15, 18] or few hundreds of neurons with apparatus called silicon probes [7, 19] when an animal is performing a task, whereas the brain area that is recorded contains more than 1000 times this number of actually physically connected neurons. Due to the physical constraints of the recording, the order of magnitude is terrific: even if the apparatus could "sense" the Local Field Potential (LFP), that is, the sum of the electrical activity of thousands of neurons [3, 17], the neurons, whose action potentials are concretely identified and recorded by the same apparatus, are just a few tens or hundreds. The scarcity is so large that, in most experiments, neurobiologists do not dare to assume any sort of physical connectivity between the recorded neurons.

Nevertheless, an amazing phenomenon takes place. These observed neurons are sometimes dependent and one may reconstruct a local dependence graph [8, 14], which we might be tempted to assimilate to the graph of "functional connectivity" and which can be used to decode the neural code [22, 27]. However from a more mathematical point of view, what do we actually reconstruct? How can a few tens of neurons taken more or less at random in the network still be that strongly dependent from a mathematical point of view? What does it tell us on the underlying full network?



(a) Observed and non observed neurons and real underlying network

(b) Observed neurons and reconstructed network

Figure 1. Difference between the real underlying network and the reconstructed one.

To answer these questions, a model of neuronal interactions needs to be settle properly. Hawkes processes and variants have been used in many articles as a model for interacting neurons [23, 25]. Indeed, the shape of their intensity mimics approximately the synaptic integration. In particular, they have been used as a model for statistical purpose to infer functional connectivity [22, 26, 27]. However there is one main problem to this global approach. Most of these articles have assumed that the observed neurons form the totality of the network, whereas this is often not the case. Few have tried to address the problem of statistical inference in this framework. Let us cite in particular [9, 24].

These two approaches tend to prove that if the recordings are made on wired neurons, the reconstruction of the network is locally correct. We want to go further and treat more general cases. More precisely, we want to understand what is the meaning of the reconstructed network on observed neurons with respect to the full network, even if the recorded neurons are not wired (see Figure 1). To do that, we use an ideal estimator : the L^2 projection of the underlying intensity on the Hawkes-like intensities that are based on the observed neurons. It is ideal in the sense that classical (penalized) least-square estimators [14, 22, 24] tend to this ideal, when the observation time tends to infinity.

Mathematically, the aim of the present work is to understand what this projection is and to derive what information on the underlying unobserved network, this projection is providing. We hope that this new approach will help to mathematically define the functional connectivity as a macroscopic phenomenon involving populations of neurons as in the neuronal assemblies concept.

2. Mathematical set-up

We are using a discrete framework as in [6, 24], which is an approximation of the classical continuous-time point processes that represent action potentials (or spikes) [14, 22]. This models the fact that we access information on the presence or absence of spike in a bin of time. We have decided to give an order of magnitude Δ to the size of the bin to see the impact of discretisation and give a qualitative order of magnitude to the different quantities when Δ tends to 0 (see also [20] for similar consideration: in their case, this is used for synchronization). In practice, Δ is at most a few milliseconds.

Let I be the possible infinite set of neurons of the network. Let Δ be the size of the time bin at which we are looking at the network. Then for all $i \in I$ and $t \in \mathbb{Z}$, X_t^i is the 0/1 variable which tells if there is no spike or at least one spike emitted by neuron i in the tth time bin. We denote $X_t = (X_t^i)_{i \in I}$.

Note that when the size of the bin Δ tends to 0, we recover the continuous-time framework of point processes (see in particular [20] for a full mathematical explanation on this limit). Informally, one can view this asymptotic as a fixed limit point process, which is observed at a coarser and discrete time resolution, $\Delta > 0$: when the size of the discretization tends to 0, we recover the limit point process.

We are assuming throughout the present paper that $(X_t)_{t\in\mathbb{Z}}$ is stationary. Moreover, since each neuron cannot fire in practice at more than 100 Hz, we are assuming that for each i, there exists some constant m_i, which represents their firing rate and that are bounded by a fixed constant M that does not depend on Δ, so that,

$$\forall\, i \in I, \quad \mathbb{E}(X_t^i) = \mathbb{P}(X_t^i = 1) = m_i\Delta. \tag{1}$$

Let us call m the vector of firing rates. Note that (1) makes sense as soon as $M\Delta \leq 1$. Indeed in practice, Δ is of the order of at best a few millisecond, whereas $M = 100$ Hz.

Let us denote V the covariance of the system, that is

$$\forall\, i, j \in I, h \in \mathbb{Z}, \quad V(i,j,h) = \mathbb{E}(X_t^i X_{t+h}^j) - m_i m_j \Delta^2.$$

Let us denote the limit by $\to$, whereas $\sim$ denotes the equivalent, that is $f(\Delta) \sim_{\Delta\to 0} h(\Delta)$ if and only if $f(\Delta)/g(\Delta) \to_{\Delta\to 0} 1$. Also in the sequel, $f(\Delta) = O(\Delta^\alpha)$ means that there exists a positive constant C such that for Δ small enough, $f(\Delta) \leq C\Delta^\alpha$.

Note that for all neurons $i, j \in I$ and lag $h \in \mathbb{Z}$, $V(i,j,h) = V(j,i,-h)$ and that when $i = j$ and $h = 0$,

$$V(i,i,0) = m_i\Delta(1 - m_i\Delta) \sim_{\Delta\to 0} m_i\Delta.$$

Remark that since the variables are Bernoulli, this covariance is also measuring the strong mixing coefficients of X_t.

We consider 2 different asymptotic set-ups.

Synchronization. If we follow [20], synchronization between 2 neurons, say j and i, amounts to say that $\mathbb{P}(X_t^i = 1 \text{ and } X_t^j = 1)/\Delta \to_{\Delta\to 0} s_{j,i,0}$, where $s_{j,i,0}$ would be the "firing rate" of the synchronisation. We can generalize it with some lag: j and i synchronize with lag h (where h can eventually be seen as a function of Δ, i.e. $h = h(\Delta)$) if

$$\frac{1}{\Delta}\mathbb{P}(X_t^i = 1 \text{ and } X_{t+h(\Delta)}^j = 1) \to_{\Delta\to 0} s_{i,j,h}.$$

Note that for all neurons $i, j \in I$ and lag h, $s_{j,i,h} = s_{i,j,-h}$ and that these quantities are bounded by M too since

$$s_{j,i,h} \leq \min(m_i, m_j).$$

Finally note that if j and i synchronize with lag h, then

$$V(i,j,h) \sim_{\Delta\to 0} s_{i,j,h}\Delta.$$

As explained in [20], the limiting point processes (N_t^i, N_t^j) do not form a classical multivariate point process with stochastic intensity, because for instance with lag $h = 0$, a spike on N_t^i can occur exactly at the same time as a spike on N_t^j. For more general lag h, one can have a fixed h (e.g. $h = 1$) for instance and therefore, in the limit where $\Delta \to 0$, spikes on each neuron can also happen at the same time, in the limit, with non zero probability. We could also have $\Delta h(\Delta) \to_{\Delta\to 0} \delta > 0$ and then there might be a fixed delay between the spikes even in the limit, with non zero probability.

The first case, when h is fixed, is the most usual. It means in practice that spikes on neuron j and i can occur with non zero probability within a time delay less than $h\Delta$. This matches classical definition and detection of synchronization (see for instance [12]). Indeed for classical recordings analysis as in [12, 13], $\Delta = 10^{-3}$s and h is between 5 and 20.

Classical point process asymptotic. On the other hand, we might think that, if there is an influence between i and j, this is not done at such a small temporal scale and that asymptotically (N_t^i, N_t^j) will be nicely behaved point processes with stochastic intensities with respect to the Lebesgue measure. More generally, this tells us that the probability that two spikes for two different neurons occur in the exact same bin t is much smaller than Δ and of order Δ^2, that is

$$\forall\, i, j \in I, \quad \mathbb{E}\big(X_t^i X_t^j\big) = O(\Delta^2),$$

which somehow prevent too frequent perfect synchronization.

In the same spirit, for any fixed time lag h, we should have

$$\forall\, i, j \in I, \quad \mathbb{E}\big(X_t^i X_{t+h}^j\big) = O(\Delta^2),$$

to ensure that, in the same sense, we cannot have too much synchronization with prescribed lag h.

This implies that except $V(i,i,0)$, which is of order Δ, all other $V(i,j,h)$ are of order Δ^2. As discussed later in Section 4.3, such a small interaction might seem realistic.

3. Ideal reconstruction

We consider the problem where we observe only a finite subset F of I. The statistician then often does as if the set of observed neurons is the totality of the network and he/she tries to fit a Hawkes model on the data at hand. More mathematically, in this discrete time set-up, he/she wants to evaluate μ_i and $g_{j\to i}$ for i and j in the set of observed neurons F (and not in the complete set of neurons I which is not totally observed) such that

$$\phi_{\mu,g}^i(t) = \mu_i + \sum_{i\in F} \sum_{s=t-A}^{t-1} g_{j\to i}(t-s) X_s^j$$

represents the probability for a given neuron i to spike at time t given the past spikes on the observed neurons. Informally, μ_i would correspond to the spontaneous firing rate when none of the other observed neurons have fired whereas $g_{j\to i} : \{1,\dots,A\} \subset \mathbb{N}^* \to \mathbb{R}$ is the interaction function which gives how much neuron j excites or inhibits neuron i after a certain delay in $\{1,\dots,A\}$. In practice, A is small with respect to the classical duration of recordings (typically up to 100 ms whereas a recordings might last several minutes) and A reflects the range of the potential interaction between neurons, but it might be interesting to consider A increasing in Δ or even $A = +\infty$ (see Section 4). Note also that by convention, outside of $\{1,\dots,A\}$, $g_{j\to i}$ is null : since we are looking at prediction of the spiking behavior of X^i given the past, it is logical to look only at non negative indices. Note in particular, that we are also forcing the functions to be null in 0, which means that the reconstruction of the interactions between neurons is always done with some delay and cannot be instantaneous.

In the present article, we do not need to precise conditions for $\phi_{\mu,g}^i(t)$ to be a conditional probability (in $[0,1]$). Indeed, and as proved by many articles using least-squares [14, 22, 24] (see in particular the oracle inequalities), the main point is to say that, for a wide range of ergodic stationary processes $(X_t)_{t\in\mathbb{Z}}$ at least, the resulting estimator aims at minimizing the following quantity, when the duration of the observed spike trains tends to infinity,

$$\mathbb{E}\big([X_t^i - \phi_{\mu,g}^i(t)]^2\big), \tag{2}$$

quantity which does not depend on t, by stationarity. Note that we use this L^2 criterion, because it is particularly well suited to deal with the covariance matrix of the system, which is linked to the strong mixing coefficient in the special case of Bernoulli variables. However in the simple case of Bernoulli variables, the L^2 distance is equivalent to the total variation distance, and is smaller

than the Hellinger or Kullback–Leibler distances, for which computation would have been more difficult.

Because $\mathbb{E}([X_t^i]^2)$ does not depend on $\mu = (\mu_i)_{i\in F}$ and $g = (g_{j\to i}(u))_{i,j\in F, u\in\mathbb{N}^*}$, we can also say that the minimizer in (μ, g) is minimizing for all i,

$$C^i(\mu, g) = -2\mathbb{E}\big(X_0^i \phi^i_{\mu,g}(0)\big) + \mathbb{E}\big([\phi^i_{\mu,g}(0)]^2\big). \tag{3}$$

Theorem 1. *Let A be a positive integer or $A = +\infty$ and let F be a finite subset of I. Let also I_F be the identity matrix of size the cardinal of F, $|F|$.*

The L^2 projection of $(X_0^i)_{i\in F}$ on the vectorial space

$$\left\{ \big(\phi^i_{\mu,g}(0)\big)_{i\in F} \;\middle|\; \mu \in \mathbb{R}^F,\ g \in \mathbb{R}^{F\times F\times\{1,\dots,A\}} \right\}$$

is given by the choice $(\bar\mu, \bar g)$ which satisfies

$$\bar\mu = (I_F - \bar{\mathbb{G}}) m_F \Delta \tag{4}$$

where $\bar{\mathbb{G}}$ is the square matrix of size $|F|$, such that $\bar{\mathbb{G}}_{i,j} = \sum_{u>0} \bar g_{j\to i}(u)$ and $m_F = (m_i)_{i\in F}$ is the vector of firing rates restricted to the set of observed neurons. Moreover, for all $i, j \in F$ and positive integer u,

$$V(j, i, u) = V * \bar g(j, i, u), \tag{5}$$

where

$$V * \bar g(j, i, u) = \sum_{k\in F} \sum_{s>0} V(j, k, u-s)\, \bar g_{k\to i}(s).$$

NB : The last equation corresponds to a convolution operator. Similar equations have been derived in the literature when the whole network is observed and the Hawkes model is true [25]. Close formula have also been derived in the continuous framework and the present equation can be put in perspective with respect to the Wiener Hopf equation [1]. However, up to our knowledge, no such equations were derived on the covariance when the Hawkes model is not assumed to be true and when only a small subset of the neurons are observed.

Proof. Since $\phi^i_{\mu,g}(t)$ is linear in (μ, g), $C(\mu, g)$ is convex and minimal when its derivative is null. Let us call $(\bar\mu, \bar g)$ the value of the minimizer. The minimizer should therefore verify

$$\forall\ i \in F, \quad \partial_{\mu_i}[C^i(\bar\mu, \bar g)] = 0 \tag{6}$$

$$\forall\ i, j \in F, u \in \mathbb{N}^*, \quad \partial_{g_{j\to i}(u)}[C^i(\bar\mu, \bar g)] = 0 \tag{7}$$

This leads to the following interpretation. With (1), we have that for all $i \in F$,

$$\begin{aligned}\partial_{\mu_i} C^i(\bar\mu, \bar g) &= -2\mathbb{E}\big(X_0^i\big) + 2\mathbb{E}\big(\phi^i_{\bar\mu,\bar g}(0)\big) \\ &= 2\left[\bar\mu_i + \sum_{j\in F}\sum_{u>0} \bar g_{j\to i}(u) m_j \Delta - m_i \Delta\right]\end{aligned}$$

Using (6), the vector $\bar\mu$ of size $|F|$ should satisfy

$$\bar\mu = (I_F - \bar{\mathbb{G}}) m_F \Delta. \tag{8}$$

On the other hand,for all $i, j \in F$ and positive integer u,

$$\begin{aligned}\partial_{g_{j\to i}(u)} C^i(\bar{\mu},\bar{g}) &= -2\mathbb{E}\big(X_0^i X_{-u}^j\big) + 2\mathbb{E}\big(X_{-u}^j \phi^i_{\bar{\mu},\bar{g}}(0)\big)\\ &= -2V(j,i,u) - 2m_j m_i \Delta^2 + 2m_j \Delta \bar{\mu}_i + 2\sum_{k\in F}\sum_{s>0} \bar{g}_{k\to i}(s)\mathbb{E}\big(X_{-s}^k X_{-u}^j\big)\\ &= -2V(j,i,u) - 2m_j m_i \Delta^2 + 2m_j \Delta \bar{\mu}_i + 2\sum_{k\in F}\sum_{s>0} \bar{g}_{k\to i}(s)\left[\mathbb{E}\big(X_{-s}^k X_{-u}^j\big) - m_k m_j \Delta^2\right]\\ &\quad + 2\sum_{k\in F}\sum_{s>0} \bar{g}_{k\to i}(s) m_k m_j \Delta^2\end{aligned}$$

But by (8)

$$\sum_{k\in F}\sum_{s>0} \bar{g}_{k\to i}(s) m_k \Delta = m_i \Delta - \bar{\mu}_i.$$

Moreover,

$$\mathbb{E}\big(X_{-s}^k X_{-u}^j\big) - m_k m_j \Delta^2 = V(j,k,u-s),$$

which concludes the proof. □

4. Interpretation

4.1. *About* $\bar{\mu}$

If the process X was a Hawkes process on the set of observed neurons F with parameter $\bar{\mu}$ and $\bar{g}$, that is for all $i \in F, t \in \mathbb{Z}$

$$\mathbb{P}\big(X_t^i = 1 \,\big|\, (X_s^j)_{j\in F, s<t}\big) = \phi^i_{\bar{\mu},\bar{g}}(t),$$

then it is easy to see, by taking expectation on both sides, that

$$m_F \Delta = (I_F - \bar{\mathbb{G}})^{-1} \bar{\mu}.$$

So in this sense, the formula (4):

$$\bar{\mu} = (I_F - \bar{\mathbb{G}}) m_F \Delta,$$

is a generalization, which applies even if X restricted to F is not a Hawkes process. It just tells us that the "projected" $\bar{\mu}$ is what cannot be explained in the mean firing rate of neuron i by the interactions with the other neurons j that have been observed including $j = i$ (that is the auto-interaction). Indeed, the term $\bar{\mathbb{G}} m_F \Delta$ is exactly the fraction of firing rate which can be explained through the interaction, $\bar{g}$. Moreover, in Theorem 1, there is no need to assume any restriction on $\bar{\mathbb{G}}$, such as a spectral radius less than 1 (which guarantees the invertibility of $I_F - \bar{\mathbb{G}}$) and it is totally possible to have negative $\bar{\mu}_i$ as a solution of (4).

The estimation of [14, 22] is performed by a Lasso estimator, which is sparse and therefore might eventually put the estimation of μ_i to 0. The Lasso penalty which enforces sparsity is there to take care of the level of noise and would eventually tend to 0 when the observation time is infinite. Therefore, when this estimator finds 0 for the spontaneous part $\bar{\mu}_i$ of neuron i, it means that up to the inherent level of noise in the observation, the firing rate of i is totally explained by what is seen on the other observed neurons.

4.2. *About* $\bar{g}$

Note that $\bar{g}$ does not need $\bar{\mu}$ to be found. Indeed, by (5), for all $i, j \in F$ and positive integer u,

$$V(j,i,u) = \sum_{k\in F}\sum_{s>0} V(j,k,u-s)\bar{g}_{k\to i}(s).$$

Note that this equation might be read as "There is (at least) a relay neuron k and a positive delay s such that j and k are correlated at lag $u-s$ (either positive or negative) and then, the interaction $\bar{g}_{k\to i}(s)$ projects this dependence onto the covariance of j and i at lag u".

This is a linear system, which might not have a solution, or could have several. However using Z transform or Fourier transform, it is easy to find conditions where this system is invertible. One of the main hidden assumptions, for this kind of inversion to work, would be that X is mixing enough in a certain sense (for instance ρ-mixing), assumption which is quite usual for these kind of ergodic processes. This leads to estimators of $\bar{g}$ that have already been used several times by Bacry and coauthors (see for instance [1]).

We do not want to go into a discussion about the merits of one estimator or another. In any case, all these estimators converge eventually towards the L^2 projection, that is, the ideal reconstruction described in Theorem 1. So let us just consider $\bar{g}$ solution of (5) and let us assume that it is bounded by R when $\Delta \to 0$. Let us now discuss what the existence of such a solution would mean depending on the different asymptotics: synchronization or classical point process asymptotic.

Synchronization. Let us assume that some neurons in the observed set F, are synchronized at very particular small lags, that is very few $V(j,i,u)$ are of order Δ whereas the others are of order Δ^2. Assume also that the range A as well as the number of observed neurons $|F|$ is small and does not grow with Δ.

We can easily see that in the asymptotic $\Delta \to 0$, most of the $V(j,i,u)$ vanish at first order, except the ones with synchronization and the terms $V(j,j,0)$. So what remains is

(i) if j and i synchronizes with lag h,

$$s_{j,i,h} = \bar{g}_{j\to i}(h)m_j + \sum_{k \text{ sync. with } j \text{ at lag } h-\ell} s_{j,k,h-\ell}\bar{g}_{k\to i}(\ell).$$

(ii) if j and i are not synchronized at lag h,

$$0 = \bar{g}_{j\to i}(h)m_j + \sum_{k \text{ sync. with } j \text{ at lag } h-\ell} s_{j,k,h-\ell}\bar{g}_{k\to i}(\ell).$$

The interpretation is close to the one for $\bar{\mu}$ at the level of the system $\bar{g}$: in the limit $\Delta \to 0$, $\bar{g}_{j\to i}$ is here to explain what is not explained by other synchronizations in the system (case (i)), or eventually to cancel out what is explained by the other interactions so that we do not see any synchronization in the limit (case (ii)). So the estimation by the Hawkes process is disentangling the interactions. For instance, if we have synchronization with delay 1 between neuron 1 and 2, and synchronization with delay 1 between neuron 2 and 3, it may happen that synchronization between 1 and 3 occurs at delay 2, but if strong $\bar{g}_{1\to 2}(1)$ and $\bar{g}_{2\to 3}(1)$ are sufficient to explain $s_{1,3,2}$, $\bar{g}_{1\to 3}(2)$ will be 0. On the other hand, if it is not sufficient to explain it then $\bar{g}_{1\to 3}(2)$ will be positive to augment the effect. Of course, if there is no synchronization at all between neuron 1 and 3 at lag 2, then eventually $\bar{g}_{1\to 3}(2)$ will be negative to cancel the effect. This is a very basic example but more generally, Hawkes processes will use the delays between synchronization to get a graph of interactions between the neurons (where j influences i if $\bar{g}_{j\to i} \neq 0$), which would explain the pattern of covariance.

This also underlines why Hawkes processes estimation cannot take well into account synchronization with no lag at all ($h=0$), because the estimation process cannot turn it into a cause-effect link : indeed the graph of interaction is directed because the cause precedes the effect.

Classical point processes asymptotic. Now let us assume that all the $V(j,i,h)$ are of order Δ^2 except $V(j,j,0)$ and let us start again with a fixed range A and a fixed set of observed neurons F. In this case, we see that

$$V(j,i,u) = V(j,j,0)\bar{g}_{j\to i}(u) + O(|F|RA\Delta^2).$$

Hence

$$O(\Delta^2) = \bar{g}_{j\to i}(u) m_j \Delta + O(|F| R A \Delta^2)$$

Hence if $|F|$, R and A are sufficiently small to say that $|F|RA\Delta = o(1)$, we get that $\bar{g}_{j\to i}(u)$ has to be proportional to Δ in the limit. We have gained an order: if we assume $\bar{g} = O(1)$, we have in fact that $\bar{g} = O(\Delta)$.

Let us denote $\bar{g}_{j\to i}(u) = \gamma_{j\to i}(u)\Delta$ so that $\gamma_{j\to i}(u)$ is interpreted as an instantaneous firing rate increase (or decrease) generated by a spike on neuron j with delay u on the firing rate at time 0 of neuron i.

We see in particular that, if we assume that γ is bounded too (say by Γ), then asymptotically

$$\gamma_{j\to i}(u) m_j = V(j,i,u)\Delta^{-2} + O(|F| A \Gamma \Delta). \tag{9}$$

This means that, basically, in the limit, the function $\bar{g}_{j\to i}$, which is reconstructed as the link between j and i, is mainly reflecting the covariance between neuron j and i. The other covariance terms $V(j,k,u-s)$ only add correction terms that are essentially negligible in the limit $\Delta \to 0$.

4.3. *What does it mean for the underlying system ?*

Now let us model the underlying network I.

Let us assume that $(X_t)_{t\in\mathbb{Z}}$ is a Hawkes process with (true) parameters (ν, φ). This means that (ν, φ) are such that

$$\mathbb{P}(X_t^i = 1 | (X_s)_{s<t}) = \nu_i + \sum_{i\in I} \sum_{s<t} \varphi_{j\to i}(t-s) X_s^j.$$

The advantages of the Hawkes model is that the functions $\varphi_{j\to i}$ are mimicking the synaptic integration and that at the same time, the model is tractable. In particular, $\varphi_{j\to i}$ is non zero only for physically connected neurons.

A neuron gets thousands of synapses [30], that is physical connections, so that, if we want the formula to hold, without adding any non linearities, we need some assumptions on the $\varphi_{j\to i}$ to make it small. So let us assume (as we have done in the classical point process asymptotic on the projection) that, from now on, in the full network, for neurons i and j, $\varphi_{j\to i}(.) = \psi_{j\to i}(.)\Delta$, and that the $\psi_{j\to i}$'s are bounded by Ψ.

This model means in particular that the interaction between two physically connected neuron is (i) not instantaneous (there is a delay of at least Δ) (ii) quite small: in particular no individual synapse is so strong that one spike of the presynaptic neuron can force alone the post synaptic neuron to fire with macroscopic probability (this probability is microscopic in $O(\Delta)$).

Now we can look at the consequences of these assumptions on the covariance $V(j,i,u)$. Indeed Theorem 1 also applies to the whole set I. Since

$$\mathbb{E}(X_t^i | (X_s)_{s<t}) = \mathbb{P}(X_t^i = 1 | (X_s)_{s<t}) = \phi_{\nu,\varphi}(t),$$

and since the conditional expectation is an L^2 projection on the past filtration, we have that (ν, φ) should also satisfy (4) and (5). Hence, for all neurons j, i and all positive u

$$\begin{aligned} V(j,i,u) &= \sum_{k\in I} \sum_{s>0} V(j,k,u-s)\psi_{k\to i}(s)\Delta \\ &= \psi_{j\to i}(u) m_j \Delta^2 (1 - m_j\Delta) + \sum_{k\in I, s>0, (k,s)\neq(j,u)} V(j,k,u-s)\psi_{k\to i}(s)\Delta \end{aligned}$$

So let us comment on the order of magnitude in this equation.

- First of all, assume there is synchronization, that is for some j, i, u, $V(j,i,u) > 0$ and $V(j,i,u) = O(\Delta)$. With a Hawkes model, synchronization cannot come from a direct interaction, which would be too small ($O(\Delta^2)$), but from a macroscopic collaboration in

the network in time and space (from now on, by space, we refer to neurons). Indeed we need, for instance, at least M neurons pointing at i ($\psi_{k\to i} \neq 0$) and at least S different lags, with corresponding $V(j,k,u-s) = O(\Delta)$ (meaning that other neurons are also synchronizing), such that $MS = O(\Delta^{-1})$. If no other neurons are synchronizing and $V(j,k,u-s) = O(\Delta^2)$, this could happen only if MS is even larger, that is $MS = O(\Delta^{-2})$.

- Next let us assume that there is no synchronization and that we are in the classical point process asymptotics where $V(j,i,u) = O(\Delta^2)$ for all $(j,u) \neq (i,0)$. Now let us pick two neurons j and i that are observed and imagine that the reconstruction $(\bar{\mu}, \bar{g})$ is such that $\gamma_{j\to i}(u) = \bar{g}_{j\to i}(u)\Delta^{-1} > 0$. From (9), then (i) either $\psi_{j\to i}$ is non zero, and then it means that the neurons j and i are physically connected in I, (ii) or that there is at least M neurons connected to i and correlated to i at S different lags and that $MS = O(\Delta^{-1})$. As already pointed out in the introduction, the case (i) "having *in vivo* two physically connected neurons in the observed data" is not the usual situation, in practice.

5. Conclusion and Perspective

The present article aims at providing a mathematical set-up where the reconstruction of the functional connectivity makes sense, even if the observed neurons are a very small subset of the complete network.

We have shown mathematically in Theorem 1 how we can easily understand what the ideal Hawkes reconstruction on a small number of neurons means in terms of firing rates and covariance between the neurons.

Indeed we consider two cases. The first case is a strong correlation between neurons which lead to synchronization. In this case the Hawkes reconstruction disentangles the reduced network to point out what might be thought of "cause" and "effect". The second case is for a smaller correlation (classical point process asymptotics in terms of the size of the time discretization parameter) and show that the interaction functions in this case, mainly reflects the covariance itself: only j or i are the cause or the effect and there is no need for "disentangling" because the higher order interactions than the direct ones are negligible.

Assume now that a Hawkes model is underlying the full network and that the observed neurons for which we recontructed an interaction are not likely to be physically connected. Imagine also that the reconstructed graph between the observed neurons is not empty and that there are interaction functions $\bar{g}_{j\to i}$ that are reconstructed and that are large, in $O(1)$ (synchronization case) or $O(\Delta)$. Then Equation (5) and its analysis in Section 4.2 imply that there are at least some $j,i,u \neq 0$ such that the covariance between neuron j and i at lag u is in $O(\Delta)$ or $O(\Delta^2)$. Finally the analysis in Section 4.3, lead us to think that this can only happen in practice if the product MS is very large, where M is the number of neurons physically connected to the neurons that are observed and where S is the range of the interaction.

From a biological point of view, this can point out to two phenomenons (and probably both are taking place)

- The physical connection between two neurons modeled by the interaction function $\varphi_{j\to i}$ correspond more or less to a postsynaptic potential. In this case, the range would be the duration of the postsynaptic potential, which is of few milliseconds (hence few Δ). In this sense, to provide such visible connections in the reconstruction on the small set of observed neurons, M has to be huge and therefore the network is massively connected. Moreover this massive connection cannot be made uniformly at random, since in this case, we might have a potential mean-field limit, which might say that two neurons picked at random are almost independent [5]. This could maybe be put in perspective with respect to expander graphs [21].

- The physical connection has a much longer "memory", and might in particular take into account learning phenomenons, such as spike-timing dependent plasticity [4] or long-term potentiation [2] phenomena that are at play at the synapse level and which modifies the strength of the interaction between physically connected neurons.

Of course, such heuristic conclusions need (i) to be precisely derived with adequate and quantified assumptions and (ii) to use a more realistic framework from a biological point of view than just the linear Hawkes process in discrete time. However we think this might give an intuition of which phenomenon might explain such a strong dependence between scarcely observed neurons.

Indeed, in view of all this heuristics results, we want to define functional connectivity from a mathematical point of view as any macroscopic behavior of the network in time and space (across neurons) that would result in high correlation between two neurons that are not necessarily directly connected. This would explain both the local reconstruction that can be made through spike train analysis and the more global approach at the brain area level with EEG. This would also explain how one can decode stimulus and behavior from the connectivity graph of a very few neurons [27]: in fact, the interaction that is estimated between neurons might result from a massive interaction inside the whole network.

Acknowledgment

The authors are very grateful to B. Cessac for his comments and nice discussions on an early version of this manuscript.

References

[1] E. Bacry, J.-F. Muzy, "Second order statistics characterization of Hawkes processes and non-parametric estimation", *IEEE Trans. Inf. Theory* **62** (2016), no. 4, p. 2184-2202.
[2] T. V. Bliss, T. Lømo, "Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path", *J. Physiol.* **232** (1973), p. 331-356.
[3] G. Buzsáki, "Large-scale recording of neuronal ensembles", *Nature Neurosci.* **7** (2004), no. 5, p. 446-451.
[4] N. Caporale, Y. Dan, "Spike Timing–Dependent Plasticity: A Hebbian Learning Rule", *Annu. Rev. Neurosci.* **31** (2008), p. 25-46.
[5] J. Chevallier, "Mean-field limit of generalized Hawkes processes", *Stochastic Processes Appl.* **127** (2017), no. 12, p. 3870-3912.
[6] R. Cofré, B. Cessac, "Exact computation of the maximum-entropy potential of spiking neural-network models", *Phys. Rev. E* **89** (2014), no. 5, p. 1262-1292.
[7] J. Csicsvari, D. A. Henze, B. Jamieson, K. D. Harris, A. Sirota, P. Barthó, K. D. Wise, G. Buzsáki, "Massively parallel recording of unit and local field potentials with silicon-based electrodes", *J. Neurophys.* **90** (2003), p. 1314-1323.
[8] V. Didelez, "Graphical models for marked point processes based on local independence", *J. R. Stat. Soc., Ser. B, Stat. Methodol.* **70** (2008), no. 1, p. 245-264.
[9] A. Duarte, A. Galves, E. Löcherbach, G. Ost, "Estimating the interaction graph of stochastic neural dynamics", *Bernoulli* **25** (2019), no. 1, p. 771-792.
[10] S. B. Eickhoff, V. I. Müller, "Functional connectivity", in *Brain Mapping, an encyclopedic reference*, vol. 2, Academic Press Inc., 2015, p. 187-201.
[11] G. L. Gerstein, P. Bedenbaugh, M. H. Aertsen, "Neuronal Assemblies", *IEEE Trans. Biomed. Eng.* **36** (1989), p. 4-14.
[12] S. Grün, M. Diesmann, A. Aertsen, "Unitary events in multiple single-neuron spiking activity: I. detection and significance", *Neural Comput.* **14** (2002), no. 5, p. 43-80.
[13] S. Grün, M. Diesmann, F. Grammont, A. Riehle, A. Aertsen, "Detecting unitary events without discretization in time", *J. Neurosci. Methods* **94** (1999).
[14] N. R. Hansen, P. Reynaud-Bouret, V. Rivoirard, "Lasso and probabilistic inequalities for multivariate point processes", *Bernoulli* **21** (2015), no. 1, p. 83-143.
[15] K. D. Harris, D. A. Henze, J. Csicsvari, H. Hirase, G. Buzsáki, "Accuracy of Tetrode Spike Separation as Determined by Simultaneous Intracellular and Extracellular Measurements", *J. Neurophys.* **84** (2000), no. 1, p. 402-413.

[16] D. O. Hebb, *The organization of behavior: a neuropsychological theory*, John Wiley & Sons; Chapman & Hall, 1949.
[17] D. A. Henze, Z. Borhegyi, J. Csicsvari, A. Mamiya, K. D. Harris, G. Buzsáki, "Intracellular features predicted by extracellular recordings in the hippocampus in vivo", *J. Neurophys.* **84** (2000), p. 390-400.
[18] M. S. Jog, C. I. Connolly, Y. Kubota, D. R. Iyengar, L. Garrido, R. Harlan, A. M. Graybiel, "Tetrode technology: advances in implantable hardware, neuroimag-ing, and data analysis techniques", *J. Neurosci. Methods* **117** (2002), p. 141-152.
[19] J. J. Jun, N. A. Steinmetz, J. H. Siegle *et al.*, "Fully integrated silicon probes for high-density recording of neural activity", *Nature* **551** (2017), p. 232-236.
[20] R. E. Kass, R. C. Kelly, W.-L. Loh, "Assessment of synchrony in multiple neural spike trains using loglinear point process models", *Ann. Appl. Stat.* **5** (2011), p. 1262-1292.
[21] E. Kowalski, *An introduction to expander graphs*, Contributions in Mathematical and Computational Sciences, Société Mathématique de France, 2019.
[22] R. C. Lambert, C. Tuleau-Malot, T. Bessaih, V. Rivoirard, Y. Bouret, N. Leresche, P. Reynaud-Bouret, "Reconstructing the functional connectivity of multiple spike trains using Hawkes models", *J. Neurosci. Methods* **297** (2018), p. 9-21.
[23] T. Onaga, S. Shinomoto, "Emergence of event cascades in inhomogeneous networks", *Nature Sci. Reports* **6** (2016).
[24] G. Ost, P. Reynaud-Bouret, "Sparse space-time models: concentration inequalities and Lasso", *Ann. Inst. Henri Poincaré, Probab. Stat.* **56** (2020), no. 4, p. 2377-2405.
[25] V. Pernice, B. Staude, S. Cardanobile, S. Rotter, "How Structure Determines Correlations in Neuronal Networks", *PLoS Comput. Biol.* **7** (2011).
[26] J. Pillow, J. Shlens, L. Paninski, A. Sher, A. Litke, E. Chichilnisky, E. Simoncelli, "Spatio-temporal correlations and visual signalling in a complete neuronal population", *Nature* **454** (2008), p. 995-999.
[27] M. Quiquempoix, "Rôle de la connectivité intracorticale dans le traitement des informations sensorielles", PhD Thesis, Université Paris 6, 2017, sous la direction de R. Lambert.
[28] A. Riehle, F. Grammont, M. Diesmann, S. Grün, "Dynamical changes and temporal precision of synchronized spiking activity in monkey motor cortex during movement preparation", *J. Physiol.* **94** (2000), no. 5-6, p. 569-582.
[29] W. Schultz, A. Dickinson, "Neuronal coding of prediction errors", *Annu. Rev. Neurosci.* **23** (2000), p. 473-500.
[30] G. M. Shepherd, *The synaptic organization of the brain*, Oxford University Press, 2004.
[31] W. Singer, "Synchronization of cortical activity and its putative role in information processing and learning", *Annu. Rev. Physiol.* **55** (1993), p. 349-374.
[32] W. Singer, A. K. Engel, A. K. Kreiter, M. H. Munk, S. Neuenschwander, P. R. Roelfsema, "Neuronal assemblies: necessity, signature and detectability", *Trends Cogn. Sci.* **1** (1997), p. 252-261.