On Some Mathematical Consequences of Binning Spike Trains

Arnaud Le Ny[(1)]

[(1)] Université Paris Est Créteil- Val de Marne (UPEC)
(Laboratoire d'Analyse et de Mathématiques Appliquées, LAMA - UMR CNRS 8050)

2^d Neuromat worskhop, New Frontiers in Neuromathematics,
São Paulo, November 23, 2016, 15h40–16h30,

Joint work with B. Cessac and E. Löcherbach

Some References

Short Bibliography

[CLNL16] B. Cessac, A. Le Ny, E. Löcherbach. Neural Computation **28**, 2016/17.
On The Mathematical Consequences of Binning Spike Trains.

[CRVV11] B. Cessac, H. Rostro, J. Vasquez, T. Viéville. J. Stat. Phys. **136** :565–602, 2009.
How Gibbs Distributions May Naturally Arise from Synaptic Adaptation Mechanisms.
A Model-Based Argumentation.

[DC06] M. Delecluse, C. Pouzat. J. Neurosci. Methods **150**(1) :16–29, 2006.
Efficient Spike Sorting of Multi-State Neurons Using Inter-Spike Intervals Information.

[LO11] : A. Galves, E. Löcherbach. J. Stat. Phys. **151**(5) :896–921, 2013.
Infinite systems of interacting chains with memory of variable length
- a stochastic model for biological neural nets.

[MM11] : I. Mastromatteo, M. Marsili. J. Stat. Mech. 2011.
On the Criticality of inferred Models.

[VB11] : J. Vasquez. PhD Thesis University of Nice-Sophia Antipolis, 2011.
Analysis of Spike-Train Statistics with Gibbs Distributions : Theory, Implementation and Applications.

Program of the talk

A Stochastic Model of Neural Net

Spike Trains as a Neural Code

Galvani (1791) : Electric nature of nervous signals, responsible of information transmission in animal life.

Ramon y Cajal (1894) : Identification of the nervous network as an assembly of cells (*neurons*) which communicate via *synapses* with a special neural interaction process induced electrically or chemically.

Hodgkin and Huxley (1952) : Neuronal electric signals propagates *via* an electrical impulses called action potentials or *spikes*.

Spike trains : Succession of spikes emitted by neuron(s), possibly (presumably !) interacting, considered as a ”neural code”

Binning and Spike Sorting

- Multi-electrode arrays (MEA) technology to record the spiking activity of populations of neurons. For us, each (k) neuron's activity is characterized by a binary variables **in discrete time**.
- Preliminary specific treatments of MEA data : *spike sorting* to distinguish spikes of "different" natures and **binning of data** : one defines a time window of $\sim 5 - 20$ ms (binning window), larger than the typical duration of a spike ($\sim 1 - 2$ ms), to gather (possibly sparse) spikes. The whole spike train is then divided into contiguous, non overlapping windows : **for each neuron k observed at time n a binary variable $\overline{\omega}_k(m(n))$ is defined** as :

Binned variables : for a (discrete) spike train ω

- $\overline{\omega}_k(m) = 0$ if the neuron k has not spiked in the m^{th} binning window.
- $\overline{\omega}_k(m) = 1$ if the neuron k has spiked *at least* once in this window.

Binning Data of a Spike Train

FIGURE : Binning data of a Spike train

Binning Data of a Spike Train

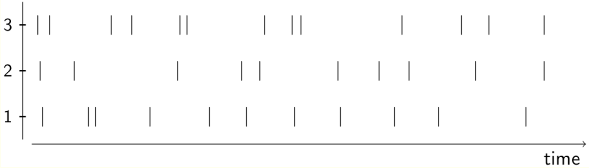

FIGURE : Binning data of a Spike train

Binning Data of a Spike Train

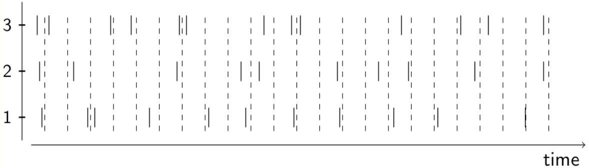

FIGURE : Binning data of a Spike train

Binning Data of a Spike Train

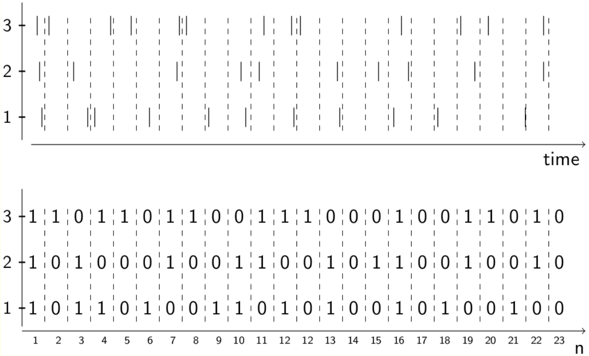

FIGURE : Binning data of a Spike train

(Reversible) 2-states Markov Chains

Markov chain $(X_n)_{n\in\mathbb{N}}$: (Discrete) stochastic process on $\{0,1\}^{\mathbb{N}}$ s.t.

(Conditioning) at time n, the future is independent of the past

Stochastic matrix P : $P(0,1) = p > 0$, $P(1,0) = q > 0$

$$\forall n \geq 0,\ \forall x, y = 0 \text{ or } 1,\ \mathbb{P}[X_{n+1} = y | X_n = x] = P(x,y)$$

In our good cases, $\exists \nu$ on $\{0,1\}$ s.t. $\nu P = \nu$ and *e.g.* for $(0, 1, \cdots, 1, 1)$

$$\mathbb{P}_\nu[X_0 = 0, X_1 = 1, \ldots, X_{n-1} = 1, X_n = 1] = \nu(0)P(0,1)\ldots P(\cdot,1)P(1,1)$$

Detailled balance :

$$\nu(i)P(i,j) = \nu(j)P(j,i)$$

$\Longrightarrow$ Reversible Markov chains whose law $\mathbb{P}_\nu$ can be extended to $\mathbb{Z}$.

Markov of order D : similar but memory of order D :

$$\mathbb{P}[X_{n+1} = x | X_{\leq n} = x_{\leq n}] = \mathbb{P}[X_{n+1} = x_{n+1} | X_{n-D+1}^n = x_{n-D+1}^n]$$

A toy model with $N = 1$ Neuron

$N = 1$ neuron, (p, q)-Markov spike train ($D = 1$), "window size" $\tau = 2$

Original Spike variables : $\omega = (\omega(0), \ldots, \omega(n))$ of weights $\mathbb{P}_\nu(\omega)$

Binned variables : for a (discrete) spike train ω

- $\overline{\omega}(m) = 0$ if the neuron has not spiked in the binning window.
- $\overline{\omega}(m) = 1$ if the neuron has spiked *at least* once.

Binning transformation $T_b : \omega \longmapsto \overline{\omega}, \ \mathbb{P}_\nu \longmapsto \mathbb{P}_\nu^{(b)}$ (factorisation map)

$$\mathbb{P}_\nu^{(b)}[\overline{\omega}] = \mathbb{P}_\nu[T^{-1}(\overline{\omega})] = \mathbb{P}_\nu[\{\omega \text{ s.t. } T_b(\omega) = \overline{\omega}\}]$$

Important fact : different pre-images

- $\overline{\omega}(m) = 0$ corresponds to the event $(0, 0)$ in the initial train.
- $\overline{\omega}(m) = 1$ corresponds either to $(0, 1)$, $(1, 0)$ or $(1, 1)$.

These factorisations can lead to **loss of Markov property.**

Illustrative example : $p = q = \mathbb{P}[0|1] = \mathbb{P}[1|0] = \frac{3}{4}$

Invariant measure $\nu = \left(\frac{q}{p+q}, \frac{p}{p+q}\right)$ i.e. $\nu(0) = \nu(1) = \frac{1}{2}$, but *not i.i.d.*

Claim 1 : The (order 1) Markov property is lost

$$\mathbb{P}^{(b)}[\overline{\omega}(2) = 0|\overline{\omega}(1) = 1, \overline{\omega}(0) = 0] \neq \mathbb{P}^{(b)}[\overline{\omega}(2) = 0|\overline{\omega}(1) = 1].$$

To compute the conditional probabilities, one has to use definition

$$\mathbb{P}^{(b)}[\overline{\omega}(2) = 0|\overline{\omega}(1) = 1] = \frac{\mathbb{P}^{(b)}[\overline{\omega}(2) = 0, \overline{\omega}(1) = 1]}{\mathbb{P}^{(b)}[\overline{\omega}(1) = 1]}.$$

and apply Markov property to the (different) individual pre-images :

$$\begin{aligned}\mathbb{P}^{(b)}[\overline{\omega}(2) = 0, \overline{\omega}(1) = 1] &= \mathbb{P}[\omega(5) = 0, \omega(4) = 0, \omega(3) = 0, \omega(\mathbf{2}) = \mathbf{1}] + \mathbb{P}[\omega(5) = 0, \omega(4) = 0, \omega(3) = 1, \omega(\mathbf{2}) = \mathbf{0}] \\ &+ \mathbb{P}[\omega(5) = 0, \omega(4) = 0, \omega(3) = 1, \omega(\mathbf{2}) = \mathbf{1}]\end{aligned}$$

and $\mathbb{P}^{(b)}[\overline{\omega}(1) = 1] = \mathbb{P}[\omega(3) = 0, \omega(\mathbf{2}) = \mathbf{1}] + \mathbb{P}[\omega(3) = 1, \omega(\mathbf{2}) = \mathbf{0}] + \mathbb{P}[\omega(3) = 1, \omega(\mathbf{2}) = \mathbf{1}]$.

Key : $\omega(2)$ depends whether one has 0 or 1 the step before.

The Binned process as a VLMC

In our numerical example, we (indeed !) get different values

$$\mathbb{P}_{\nu}^{(b)}[\overline{\omega}(2) = 0|\overline{\omega}(1) = 1] = 0,1339 \neq \mathbb{P}_{\nu}^{(b)}[\overline{\omega}(2) = 0|\overline{\omega}(1) = 1,\ \overline{\omega}(0) = 0] = 0,1125$$

In fact, it is the association of 3 symbols of the initial Markov chain to the same the symbol $\overline{\omega} = 1$ that can lead to a loss of the Markov property with the creation of a **memory of variable length** : For any binned block $\overline{\omega}_r^s = (\overline{\omega}(r), \overline{\omega}(r+1), \ldots, \overline{\omega}(s-1), \overline{\omega}(s))$ we have

$$\mathbb{P}^{(b)}[\overline{\omega}(s+1)|\overline{\omega}_r^s] = \mathbb{P}^{(b)}[\overline{\omega}(s+1)|\overline{\omega}_l^s], \qquad (1)$$

where l is the first occurrence of the symbol 0 when going from s to r.

Binning Data of a Spike Train

The binned process is thus a *Variable Length Markov Chain* (VLMC) or *Context Tree Model* in the sense of Rissanen, **where memory goes back up to* the first occurrence of a 0 in the past**(*and not to the last spike !)

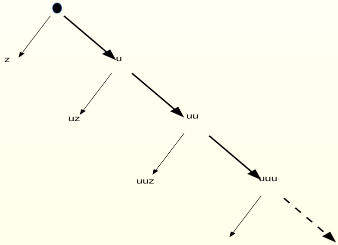

FIGURE : Context Tree for the VLMC (z codes 0, u codes 1)

VLMC for N neurons

Spike train : $\big((\omega(n)) = (\omega_k(n))_{k=1\dots N}\big)_{n\in\mathbb{Z}} \sim \mathbb{P}$ Markov order D,

$$\forall n \in \mathbb{Z},\ \mathbb{P}[\omega(n+1)|\mathcal{F}_{\leq n}](\cdot) = \mathbb{P}[\omega(n+1)|\mathcal{F}_{n-D+1}^{n}](\cdot)\ \ \mathbb{P}-\text{a.s.}$$

Supposed to be nice (*primitive transition matrix*).

Binned raster $\overline{\omega}$: Partition $\mathbb{Z} = \cup_{m\in\mathbb{Z}} F_m,\ F_m = [m\tau, (m+1)\tau - 1] \cap \mathbb{Z}$.

$$\overline{\omega}_k(m) := 1 \text{ if } \exists n \in F_m, \overline{\omega}_k(n) = 1 \text{ vs. } \overline{\omega}_k(m) = 0 \text{ when } \forall n \in F_m, \omega_k(n) = 0.$$

For any past $\overline{\omega}_{-\infty}^{-1}$, the length of the ”variable memory” of the binned law $\mathbb{P}^{(b)}$ will be again the time required to get in the past the ”null” $\mathbf{0} := (0)_{k=1\dots N}$ binned configuration for which no neuron has spiked :

$$l(\overline{\omega}_{-\infty}^{-1}) := \inf\{m : \overline{\omega}(-m) = \mathbf{0}\}.$$

Proposition : Suppose $\tau \geq D$. Then for any infinite past $\overline{\omega}_{-\infty}^{-1}$,

$$\forall a = 0 \text{ or } 1,\ \mathbb{P}^{(b)}\left[\omega(0) = a|\overline{\omega}_{-\infty}^{-1}\right] = \mathbb{P}^{(b)}\left[\omega(0) = a|\overline{\omega}_{-l(\overline{\omega}_{-\infty}^{-1})}^{-1}\right].$$

Continuity of the "one-sided" conditional probabilities

It is important that the initial spike train as no "forbidden" transitions :

- The memory $l(\overline{\omega})$ is variable, unbounded but a.s. finite because :

$$\mathbb{P}^{(b)}[\exists \text{ infinitely windows } F_m \text{ s.t. } \overline{\omega}(m) = \mathbf{0}] = 1$$

- The binned chains is continuous w.r.t the past, with exponential continuity rate : $\exists \alpha = \alpha(N) > 0$, so that as $n \to \infty$:

$$\beta(n) := \sup_{a=0,1} \sup_{x} \sup_{y,z} \left| \mathbb{P}^{(b)}(a|x_{-n}^{-1} y_{-\infty}^{-n-1}) - \mathbb{P}^{(b)}(a|x_{-n}^{-1} z_{-\infty}^{-n-1}) \right| = O\left(e^{-\alpha n}\right).$$

In the context of dynamical systems, such a process is a (particularly regular) g-measure : a measure consistent with a system of conditional probabilities with respect to the past ("One-sided").

Gibbs property : 2-sided conditional probabilities

- **Does binning affect anticipation properties ?** : In general, one-sided and two-sided conditionings are not equivalent....
- Here, starting from a (finite order, primitive) Markov model, the binned process remains a **Gibbs measure** : there always exists a continuous version of two-sided conditional probabilities, i.e. when conditionning on the past **and** on the future.
- For more general models, phase transition might occur and creates some discontinuous memories. Would they be related to "spurious phase transitions" that are around spike trains statistics ? More investigations are needed (long-range models in dimension one, two-sided Gibbs approach, renormalization transformations, non-Gibbsian measures, etc.).

Perspectives

- What about long-range chains with possible phase transitions ?
- What about $N \to \infty$?
  Markov chains with denumerable state space could more easily lead to discontinous g-measures (discontinuous VLMC).
- Case of forbidden transitions – Non-primitives matrices.
- Does one can explain "critical effects" or "spurious phase transitions" by discontinuities (and non-Gibbsianness) by starting with different spike trains distributions ?
- What about random graphs/trees ?

Other References

[CGG14] W. de Carvalho, S. Gallo, N. Garcia. J. Appl. Proba. **53**(1) :216–230, 2016.
Continuity Properties of a Factor of Markov Chains.

[CGL12] M. Cassandro, A. Galves, E. Löcherbach. J. Stat. Phys. **147** :795–807, 2012.
Partially Observed Markov Random Fields are Variable Neighborhood Random Fields.

[COP09] : M. Cassandro, E. Orlandi, P. Picco. J. Math. Phys. **46**, 2009.
Phase Transition in the 1D Random Field Ising Model with Long Range Interaction.

[CU11] J. Chazottes, E. Ugalde. *In* Entropy of Hidden Markov Processes, CUP, 2011.
On the preservation of Gibbsianness Under Amalgamation of Symbols.

[ELN16] : A.C.D. van Enter, A. Le Ny. In revision to Stoch. Proc. Appl., 2016.
Decimation of the Dysin-Ising Ferromagnet.

[FP97] R. Fernández, C. Pfister. Ann. Probab. **25**(3) :1284-1315, 1997.
Global specifications and non-quasilocality of projections of Gibbs measures.

[LO11] : E. Löcherbach, E. Orlandi. Stoch. Proc. Appl. **121**(9), 2011.
Neighborhood radius estimations for variable-neighborhood random fields.

[RW10] : F. Redig, F. Wang. Markov Proc. Relat. Fields **16**, 2010.
Transformations of One-Dimensional Gibbs Measures with Infinite Range Interaction.

[R83] : J. Rissanen. IEEE Trans. Inform. Theory IT **29** :656–664, 1983.
A Universal Data Compression System.

[V11] : E. Verbistkiy. Indagationes Mathematicae **22** :315-329, 2011.
On Factors of g-measures.