Metastability for interacting neurons

Eva Löcherbach and Pierre Monmarché

Journées de Probabilités, Juin 2021

Outline

The Model

Metastability

Where are we ?

The Model

Metastability

Systems of interacting neurons

- *N* spiking neurons described by their membrane potential values $X_t^{N,1}, \ldots, X_t^{N,N} \ge 0, t \ge 0.$
- Each neuron *i* 'spikes' (sends action potentials) at rate $f(X_t^{N,i})$.
- f Lipschitz, increasing, f(0) = 0, bounded:

 $f(x) \leq f_*$.

Systems of interacting neurons

- *N* spiking neurons described by their membrane potential values $X_t^{N,1}, \ldots, X_t^{N,N} \ge 0, t \ge 0.$
- Each neuron *i* 'spikes' (sends action potentials) at rate $f(X_t^{N,i})$.
- f Lipschitz, increasing, f(0) = 0, bounded:

$$f(x) \leq \frac{f_*}{f_*}$$

• When neuron *i* spikes

 \Rightarrow the potential of neuron *i* is reset to a resting value chosen to be 0

 \Rightarrow all other neurons $j \neq i$ get an additional potential value h/N which is added to their current value. We call h > 0 the synaptic weight.

 \bullet Between successive spikes, there is loss of potential at rate $\alpha>$ 0, that is,

$$dX_t^{N,i} = -\alpha X_t^{N,i} dt$$

in between successive jumps of the system.

Remark

Model considered by Cessac 2011, Galves and L. 2013, De Masi, Galves, L. and Presutti 2015, Fournier and L. 2016, Robert and Touboul 2016, Cormier, Tanré and Veltz 2019, and certainly others: Leaky integrate and fire model.

Generator

The above process is a very simple PDMP and has generator

$$L\varphi(x) = \sum_{i=1}^{N} f(x^{i})[\varphi(x + \Delta^{i}(x))) - \varphi(x)] - \alpha \sum_{i=1}^{N} \frac{\partial \varphi}{\partial x^{i}}(x)x^{i},$$
$$(\Delta^{i}(x))^{j} = \left\{\begin{array}{cc} \frac{h}{N} & j \neq i\\ -x^{i} & j = i \end{array}\right\}.$$

Big jumps (unbounded) and collateral jumps of order 1/N !

Remark

Process can also be seen as a system of N interacting non-linear Hawkes processes with kernel function $h_{ij}(t) = (h/N)e^{-\alpha t}, i \neq j$, and variable length memory (reset to 0).

Longtime behavior of the finite system

• Since f(0) = 0, the all-zero state is an invariant state of the system.

Longtime behavior of the finite system

• Since f(0) = 0, the all-zero state is an invariant state of the system.

• It is straightforward to show (Duarte and Ost (MPRF 2016)):

Theorem

If f is differentiable in 0, then the system stops spiking almost surely. As a consequence, the unique invariant measure of the process (X_t^N) is given by δ_0 , where $\mathbf{0} \in \mathbb{R}^N$ denotes the all-zero vector in \mathbb{R}^N .

Longtime behavior of the finite system

• Since f(0) = 0, the all-zero state is an invariant state of the system.

• It is straightforward to show (Duarte and Ost (MPRF 2016)):

Theorem

If f is differentiable in 0, then the system stops spiking almost surely. As a consequence, the unique invariant measure of the process (X_t^N) is given by δ_0 , where $\mathbf{0} \in \mathbb{R}^N$ denotes the all-zero vector in \mathbb{R}^N .

• Situation changes as the number of neurons tends to infinity : As $N \to \infty$, **0** becomes unstable for some values of the parameters.

- If neuron *i* has initially potential value x_i , then the probability that its first spike occurs after time *t* (cond. on the fact that no other spikes have happened in the mean while) is

$$P(T_1^i > t) = \exp\left(-\int_0^t f(e^{-\alpha s}x_i)ds\right).$$

- If neuron *i* has initially potential value x_i , then the probability that its first spike occurs after time *t* (cond. on the fact that no other spikes have happened in the mean while) is

$$P(T_1^i > t) = \exp\left(-\int_0^t f(e^{-\alpha s}x_i)ds\right).$$

- Use change of variables $y = e^{-\alpha s} x_i$, $dy = -\alpha y ds$:

$$P(T_1^i > t) = \exp\left(-\frac{1}{\alpha}\int_{e^{-\alpha t}x_i}^{x_i}\frac{f(y)}{y}dy\right).$$

- If neuron *i* has initially potential value x_i , then the probability that its first spike occurs after time *t* (cond. on the fact that no other spikes have happened in the mean while) is

$$P(T_1^i > t) = \exp\left(-\int_0^t f(e^{-\alpha s}x_i)ds\right).$$

- Use change of variables $y = e^{-\alpha s} x_i$, $dy = -\alpha y ds$:

$$P(T_1^i > t) = \exp\left(-\frac{1}{\alpha}\int_{e^{-\alpha t}x_i}^{x_i}\frac{f(y)}{y}dy\right).$$

– Let $t \to \infty$:

$$P(T_1^i = \infty) = \exp\left(-\frac{1}{\alpha}\int_0^{x_i}\frac{f(y)}{y}dy\right) > 0,$$

since $\int_0 \frac{f(y)}{y} dy < \infty$: $f'(0) < \infty$.

- At each time t such that all potential values of all neurons are simultaneously below some threshold K, there is a strictly positive probability

$$\geq \exp\left(-\frac{N}{\alpha}\int_0^K \frac{f(y)}{y}dy\right)$$

that none of the neurons does ever spike again.

- Use Lyapunov techniques to show that this event (all values below K) happens i.o. almost surely plus **conditional Borel-Cantelli lemma**.

- Finite system possesses a last spiking time $L = L^N < \infty$ almost surely, for any N, after which it is silent.
- Isolated system: we can interpret L^N as the time an initial stimulus survives in the system.
- ► The situation changes however as N → ∞, as we can see on simulations (done by C. Pouzat for a slightly different model including synaptic plasticity)



We will show that this is related to "Metastability".

Where are we ?

The Model

Metastability

- Metastability = transient behavior of the system during which the systems stays close over a very long time to a seemingly stable point = the metastable state.
- Lebowitz and Penrose, Rigorous treatment of metastable states in the van der Waals-Maxwell theory, JSP 1971:
 - A system starting in a metastable state is very likely to stay there for a long time.
 - Once it has left the vicinity of the metastable state, it is very unlikely to return there.
 - This happens after an unpredictable time which is random, and the "unpredictability" is expressed through the fact that it must be *exponentially* distributed.

Initially described in the context of small noise diffusions evolving in an energy landscape



It is the appearance of a statistically rare event that pushes the system out of the metastable state.

- In our case, "Metastability" means that the renormalized last spiking time is asymptotically exponentially distributed, that is L^N/𝔼(L^N) ~ exp(1) as N → ∞.
- Intuitively, this will happen if during a long time, the system does not feel that its actual size is finite (N), but behaves as if it were already in its infinite population limit - and therein, close to its invariant state.

Limit system

Form of the limit? Let us have a look on its generator:

$$A^{N}\varphi(u) = \sum_{i=1}^{N} f(u^{i})[\varphi(u + \Delta_{i}(u))) - \varphi(u)] - \alpha \sum_{i=1}^{N} \frac{\partial \varphi}{\partial u_{i}}(u)u_{i},$$

$$(\Delta_i(u))_j = \left\{ \begin{array}{cc} \frac{h}{N} & j \neq i \\ -u_i & j = i \end{array} \right\}$$

► If there would be no reset -u_i (no big jumps), Taylor's formula would imply

$$\varphi(u + \Delta^{i}(u))) - \varphi(u) = \sum_{j=1}^{N} \frac{\partial \varphi}{\partial u_{j}}(u) \cdot \frac{h}{N} + o(\frac{1}{N}):$$

jumps become a deterministic drift term that pushes each neuron's potential upwards by $h\bar{f}(u)$ where \bar{f} is the total spiking rate of the system.

Under suitable assumptions on the initial potential values, the limit system for neurons without reset (classical Hawkes) is given by a simple ODE for the membrane potential of one typical neuron within the infinite pop limit model

$$x(t) = x(0) - \alpha \int_0^t x(s) ds + \int_0^t \frac{hf(x(s))}{s} ds$$

(every neuron is described by this equation, and neurons are all independent in the limit (*propagation of chaos*)).

Any stationary state x* must satisfy

 $\alpha x^* = hf(x^*).$

• $x^* = 0$ always solution since f(0) = 0.

- for concave f and sufficiently large synaptic weights h, this implies the existence of at least a second non-trivial equilibrium $x^* > 0$ that is attracting.



- these arguments do not apply directly to the true process with reset inducing **big jumps and discontinuities**.

The true limit process for the model with reset

In the true model, the membrane potential process of a typical neuron in the limit remains a stochastic process which is described by

$$d\bar{X}(t) = -\alpha \bar{X}(t)dt + h\mathbb{E}(f(\bar{X}(t)))dt - \bar{X}(t) - d\bar{Z}(t),$$

where \bar{Z} is a jump process having stochastic intensity $f(\bar{X}(t-))$.

- Each neuron's potential undergoes leakage at exponential rate - and has an upward drift given by the current mean firing rate of the system (multiplied by h).
- Moreover, it spikes randomly, at rate f(x), whenever its current value of potential is x, and goes back to the reset value 0 after each spike.

Remark

We have convergence to the limit system (propagation of chaos) plus a rate of convergence, e.g. for the synchronous coupling,

$$\mathbb{E}|ar{X}(t) - X_1^N(t)| \le C(\sqrt{t}+t)e^{(lpha+hk+\|f\|_\infty)t}N^{-1/2},$$

see Fournier and L. (2016), Robert and Touboul (2016), Monmarché and L. (2021).

Invariant states of the true limit

- In any invariant state, the drift term t → hE(f(X(t))) must be constant, say ≡ b.
- This defines for any fixed b a classical renewal Markov process (process coming back to 0 i.o. and thus being recurrent) with generator

$$A^{b}\varphi(x) = -\alpha x \varphi'(x) + \frac{b}{\varphi}\varphi'(x) + f(x)[\varphi(0) - \varphi(x)]$$

and unique invariant probability measure π^{b} .

The true non-linear invariant measure is then solution of the fixed-point equation

$$h\pi^b(f)=b.$$

▶ $b = 0, \pi^b = \delta_0$ is always solution - is there another one ?

Phase transition

- Suppose that $f(x) \ge kx$ for all $x \in [0, u_*]$.
- What is the meaning of k? k = f'(0).
- h the synaptic weight and α the exponential rate of decay : If

$\alpha < \mathbf{kh},$

then in the $N \to \infty-$ limit, there is a second non-trivial (different from the all-zero-state) equilibrium. Robert-Touboul (2016), Cormier, Tanré, Veltz (2020+), Lö-Monmarché (2021).

- Second equilibrium is an absolutely continuous measure having an explicit density depending on the unknown mean spiking rate.

Second equilibrium is an absolutely continuous measure having density

$$g(x) = \frac{p_{\infty}}{V_{\infty}(x)} e^{-\int_0^x \frac{f(y)}{V_{\infty}(y)} dy}, \ x < hp_{\infty}/\alpha,$$

 $V_{\infty}(x) = -\alpha x + hp_{\infty}, \ p_{\infty} = \int_{0}^{\infty} f(x)g(x)dx$ invariant mean spiking rate (UNKNOWN!).

In the sequel, we work with saturating and piecewise linear spiking rates :





Then, for sufficiently high values of *kh*, this second equilibrium is globally attracting (at exponential speed).

We do not have any closed formula to compute e.g. the equilibrium spiking rate explicitly: the limit system is not easy to analyze... **However:**

Series of papers by Quentin Cormier, Etienne Tanré and Romain Veltz, with deep results about local and global stability of invariant state and existence of periodic solutions using functional-analytical tools. Then, for sufficiently high values of kh, this second equilibrium is globally attracting (at exponential speed).

We do not have any closed formula to compute e.g. the equilibrium spiking rate explicitly: the limit system is not easy to analyze... **However:**

- Series of papers by Quentin Cormier, Etienne Tanré and Romain Veltz, with deep results about local and global stability of invariant state and existence of periodic solutions using functional-analytical tools.
- For our purposes we need more: a loss of memory property of the limit process which we will prove by using coupling. This loss of memory will imply

Metastability

For saturating rate functions and sufficiently high values of kh (the product of the derivative of the rate function in 0 and the synaptic weight):

- The finite system stays a long time in a vicinity of the limit equilibrium (e.g. of the limit equilibrium spiking rate).

- It is kicked out of this vicinity after an unpredictably long time τ such that

$$\left|\mathbb{P}\left(au \geq t\mathbb{E}(au)\right) - e^{-t}\right| \leq C \ln N/N^{1/4}
ightarrow 0$$

as $N \rightarrow \infty$ (with P. Monmarché 2021)

Ideas of proof : why exponential distribution?

- Global spiking rate $F^N(t) = \sum_{i=1}^N f(X_i^N(t)) = N\overline{f}(X^N(t))$.
- $\mathcal{D} = \{\bar{f} \ge \delta\}$: mean spiking rate strictly lower-bounded (by some small constant).
- ► If the process has not left D after some time T, then F^N/N is probably in a vicinity of the limit equilibrium
- ► and then it will typically stay in D a time longer than N and forget its initial position.

Ideas of proof : why exponential distribution?

- Global spiking rate $F^N(t) = \sum_{i=1}^N f(X_i^N(t)) = N\overline{f}(X^N(t))$.
- $\mathcal{D} = \{\bar{f} \ge \delta\}$: mean spiking rate strictly lower-bounded (by some small constant).
- ► If the process has not left D after some time T, then F^N/N is probably in a vicinity of the limit equilibrium
- and then it will typically stay in \mathcal{D} a time longer than N and forget its initial position.
- ► So the trials to reach the all-zero configuration become independent ⇒ exponential distribution.

Forgetting the initial position/Coupling in the limit system

- Take two copies of the limit process (one single, typical neuron), with different initial values, and driven by the same Poisson random measure. This leads to a coupling of the jumps and makes the two processes jump together as often as possible.
- Once both processes have reached potential value f_{*}/k, the associated rates are equal and thus the two processes have the same jumps. Once they jump, they are both reset to 0 and hence equal at that time.
- Difference in the drift is the reason why they separate again but drift is only felt before reaching threshold f_{*}/k.
- How long does it take to reach value f_*/k ?

Forgetting the initial position/Coupling in the limit system

- Take two copies of the limit process (one single, typical neuron), with different initial values, and driven by the same Poisson random measure. This leads to a coupling of the jumps and makes the two processes jump together as often as possible.
- Once both processes have reached potential value f_{*}/k, the associated rates are equal and thus the two processes have the same jumps. Once they jump, they are both reset to 0 and hence equal at that time.
- Difference in the drift is the reason why they separate again but drift is only felt before reaching threshold f_{*}/k.
- How long does it take to reach value f_*/k ?
- ► Drift felt by each neuron: $-\alpha x + h\bar{f}$, where \bar{f} is the mean spiking rate of the system.

Forgetting the initial position/Coupling in the limit system

- Take two copies of the limit process (one single, typical neuron), with different initial values, and driven by the same Poisson random measure. This leads to a coupling of the jumps and makes the two processes jump together as often as possible.
- Once both processes have reached potential value f_{*}/k, the associated rates are equal and thus the two processes have the same jumps. Once they jump, they are both reset to 0 and hence equal at that time.
- Difference in the drift is the reason why they separate again but drift is only felt before reaching threshold f_{*}/k.
- How long does it take to reach value f_*/k ?
- ► Drift felt by each neuron: $-\alpha x + h\bar{f}$, where \bar{f} is the mean spiking rate of the system. Need to lower-bound this drift by something strictly positive that we are able to control explicitly.

An auxiliary Markov process for the finite system

- We wish to estimate the time we need for a given neuron to reach the threshold f_*/k without spiking.

- We introduce an auxiliary simple Markov process Z^N such that

$$ar{f}(X^N) = rac{1}{N}\sum_{i=1}^N f(X^N_i(t)) \geq Z^N$$

and large deviation estimates for Z^N are easily obtained (associated limit process z_t is deterministic).

– We know the time the limit process associated to Z^N needs to reach f_*/k .

- Construction of Z^N does only depend on behavior of derivative of f in vicinity of 0 (in particular, it also works for more general choices of rate functions).

More on Z^N Z^N has generator

$$A^{Z^N}\varphi(z) = -lpha z \varphi'(z) + N z [\varphi(m_N(z)) - \varphi(z)],$$

where

$$m_N(z) = z + \frac{kh}{N}(1 - \frac{z}{f_* - \frac{kh}{N}})_+ - f_*/N.$$

More on Z^N Z^N has generator

$$A^{Z^N}\varphi(z) = -lpha z \varphi'(z) + N z [\varphi(m_N(z)) - \varphi(z)],$$

where

$$m_N(z)=z+\frac{kh}{N}(1-\frac{z}{f_*-\frac{kh}{N}})_+-f_*/N.$$

- Indeed, all neurons with potential value $\leq \frac{f^*}{k} - \frac{h}{N}$ will have an increase of their rate – due to the spike of another neuron – given by kh/N.

- Total number of such neurons is lower-bounded by

$$N-\frac{\bar{f}(X^N)}{f(\frac{f^*}{k}-\frac{h}{N})}=N(1-\frac{\frac{1}{N}\bar{f}(X^N)}{f^*-\frac{kh}{N}}).$$

More on Z^N Z^N has generator

$$A^{Z^N}\varphi(z) = -lpha z \varphi'(z) + N z [\varphi(m_N(z)) - \varphi(z)],$$

where

$$m_N(z)=z+\frac{kh}{N}(1-\frac{z}{f_*-\frac{kh}{N}})_+-f_*/N.$$

- Indeed, all neurons with potential value $\leq \frac{f^*}{k} - \frac{h}{N}$ will have an increase of their rate – due to the spike of another neuron – given by kh/N.

- Total number of such neurons is lower-bounded by

$$N-\frac{\bar{f}(X^N)}{f(\frac{f^*}{k}-\frac{h}{N})}=N(1-\frac{\frac{1}{N}\bar{f}(X^N)}{f^*-\frac{kh}{N}}).$$

- Associated limit system is

$$\dot{z} = -lpha z + G(z)z, \ G(z) = (kh(1 - z/f_*) - f_*)_+.$$

Still more on Z^N

- Limit system ż = −αz + G(z)z has unique attracting equilibrium z_∞ > 0. Indeed :
- G is decreasing and has a unique solution z_∞ of G(z_∞) = α (supposing kh sufficiently large).
- ▶ z_{∞} is globally attracting for the limit dynamic on $]0, +\infty[$.
- ... and we have a LDP for Z^N with good rate function (Feng, Kurtz (2006)),

Still more on Z^N

- Limit system ż = −αz + G(z)z has unique attracting equilibrium z_∞ > 0. Indeed :
- G is decreasing and has a unique solution z_∞ of G(z_∞) = α (supposing kh sufficiently large).
- ▶ z_{∞} is globally attracting for the limit dynamic on]0, +∞[.
- Image: marked and m
- ► The limit dynamic ż = -αz + G(z)z given above is only a (pessimistic and very rough) lower bound on the true limit dynamic.



Forgetting the initial position in the limit process

- ► Take two copies \bar{X} and \tilde{X} of the limit system, s.t. $\mathbb{E}f(\bar{X}(0)), \mathbb{E}f(\tilde{X}(0)) \ge z_{\infty} - \delta$. Couple the spiking rates!
- How long does it take to reach value f_{*}/k (above which the rates are equal)?
- Use that

$$\mathbb{E}(f(\bar{X}(s))) \geq z_s \geq z_\infty - \delta$$

such that the drift felt by each neuron at time s

 $-\alpha \bar{X}(s) + h\mathbb{E}(f(\bar{X}(s))) \geq -\alpha \bar{X}(s) + h(z_{\infty} - \delta).$

Yields an explicit bound t_δ : time needed to reach f_{*}/k, without jumping, starting from 0.

Control the distance

$$\delta_t = \mathbb{E}|f(\bar{X}(t)) - f(\tilde{X}(t))|.$$

- It suffices to consider that the last jump before time t has happened within $[t - t_{\delta}, t]$ (if no such jump exists, then $f(\bar{X}(t)) = f(\tilde{X}(t))$, since both are above the saturation threshold!)

Control the distance

$$\delta_t = \mathbb{E}|f(\bar{X}(t)) - f(\tilde{X}(t))|.$$

- It suffices to consider that the last jump before time t has happened within $[t - t_{\delta}, t]$ (if no such jump exists, then $f(\bar{X}(t)) = f(\tilde{X}(t))$, since both are above the saturation threshold!) - Asynchronous jumps happen at rate $|f(\bar{X}(t)) - f(\tilde{X}(t))|$. Control the distance

$$\delta_t = \mathbb{E}|f(\bar{X}(t)) - f(\tilde{X}(t))|.$$

- It suffices to consider that the last jump before time t has happened within $[t - t_{\delta}, t]$ (if no such jump exists, then $f(\bar{X}(t)) = f(\tilde{X}(t))$, since both are above the saturation threshold!) - Asynchronous jumps happen at rate $|f(\bar{X}(t)) - f(\tilde{X}(t))|$. - At synchronous jumps, both processes are reset to 0 and then only feel the differences in the drift upper bounded by

 $h\mathbb{E}(|f(\bar{X}(s)) - f(\tilde{X}(s))|).$

Gives rise to Gronwall with bounded memory

$$\delta_t \leq c_\delta \int_{t-t_\delta}^t \delta_s ds, t \geq t_\delta,$$

plus contribution of initial terms on $[0, t_{\delta}]$. Contraction works if δ sufficiently small.

- If the initial spiking rates are not above z_∞ − δ, apply coupling with the auxiliary process : associated limit system z_t reaches (z_∞ − δ, +∞) after some finite time.
- coupling of rates (bounded and equal above threshold) seems easier than coupling of the true potential values (unbounded and never equal, except right after a common jump).
- To obtain Wasserstein contraction of X and X, main observation is: Asynchronous jumps of X and X automatically imply that the non-jumping particle is below threshold, that is, bounded by f_{*}/k, such that we are able to control this contribution.

Coupling of the rates in the finite system

- Here we use the LD-bounds to compare \overline{F}^N/N with the limit system z_t such that, with exp high probability, time needed to reach level above threshold f_*/k is still comparable to t_{δ} .

TV coupling for the finite system

Total variation coupling of all potential values of the two systems with N neurons succeeds within [0, N] with exp. high probability. It works on the following event:

- ► All spikes are synchronous during [N/2, N] (→ Wasserstein coupling of the rates !)
- Auxiliary process initialized at time N/2, representing lower bound on accumulated drift, is above f_{*}/k within [N/2 + t_δ, N].
- Each process has jumped at least once in $[N/2 + t_{\delta}, N]$.
- Total overall spiking rate is above $z_{\infty} \delta$ at time N/2.

Remark

We have proved a general statement about exponentiality of exit times which extends the results of Brassesco, Olivieri, Vares, JSP 1998 for low-noise diffusions to our frame.



Figure: Summary of the results when $f(u) = (ku) \wedge f_*$ with $a = \alpha/(kh)$ and $b = f_*/(kh)$.

 Other papers on metastability for different systems of interacting neurons: Morgan André (arXiv, 2019), Léo Planché and Morgan André (to appear in SPA).

Ludmila Brochini and Miguel Abadi study a simplified and time discrete version (arXiv 2018)

- Our paper on metastability with Pierre Monmarché is on ArXiv : https://arxiv.org/abs/2004.13353 Other papers on metastability for different systems of interacting neurons: Morgan André (arXiv, 2019), Léo Planché and Morgan André (to appear in SPA).

Ludmila Brochini and Miguel Abadi study a simplified and time discrete version (arXiv 2018)

 Our paper on metastability with Pierre Monmarché is on ArXiv : https://arxiv.org/abs/2004.13353

Thanks for your attention !!!