Statistical testing of the covariance matrix rank in multidimensional neuronal models: non-asymptotic case

Anna Melnykova

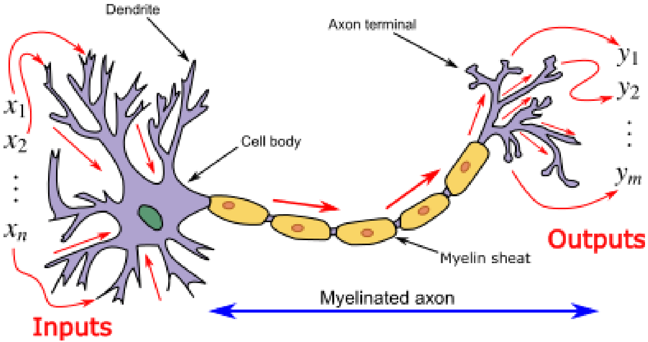
joint work with: Patricia Reynaud-Bouret, Adeline Samson

Université de Cergy-Pontoise, Université de Grenoble Alpes

DynStoch 2019, Delft, Netherlands

13 June 2019

Motivation


Dendrite
Axon terminal
Cell body
Myelin sheat
Myelinated axon
x_1
x_2
$\vdots$
x_n
Inputs
y_1
y_2
$\vdots$
y_m
Outputs

Example 1: Hodgkin-Huxley model

Conductance-based model of **action potential in neurons**:

$$\begin{cases} I & = C_m \frac{\mathrm{d} V_m}{\mathrm{d} t} + \bar{g}_{\mathrm{K}} n^4 (V_m - V_K) + \bar{g}_{\mathrm{Na}} m^3 h (V_m - V_{Na}) + \bar{g}_l (V_m - V_l) \\ \frac{dn}{dt} & = \alpha_n(V_m)(1-n) - \beta_n(V_m) n \\ \frac{dm}{dt} & = \alpha_m(V_m)(1-m) - \beta_m(V_m) m \\ \frac{dh}{dt} & = \alpha_h(V_m)(1-h) - \beta_h(V_m) h \end{cases}$$

- I — membrane potential
- n, m, h — quantities between 0 and 1 that are associated with potassium channel activation, sodium channel activation, and sodium channel inactivation.

References: Hodgkin and Huxley (1952) — *1963 Nobel Prize in Physiology or Medicine*,
Modifications: Fitzhugh (1961), Morris and Lecar (1981)

UNIVERSITÉ Grenoble Alpes

UNIVERSITÉ de Cergy-Pontoise

Example 2: Jansen and Rit Neural Mass model

Convolution-based model of a **neuronal population with excitatory and enhibitory subpopulations**:

$$\begin{cases} dQ(t) = \nabla_P H(Q,P)\,dt, \\ dP(t) = (-\nabla H(Q,P) - 2\Gamma P + G(t,Q))\,dt + \Sigma(t)\,dW_t, \end{cases}$$

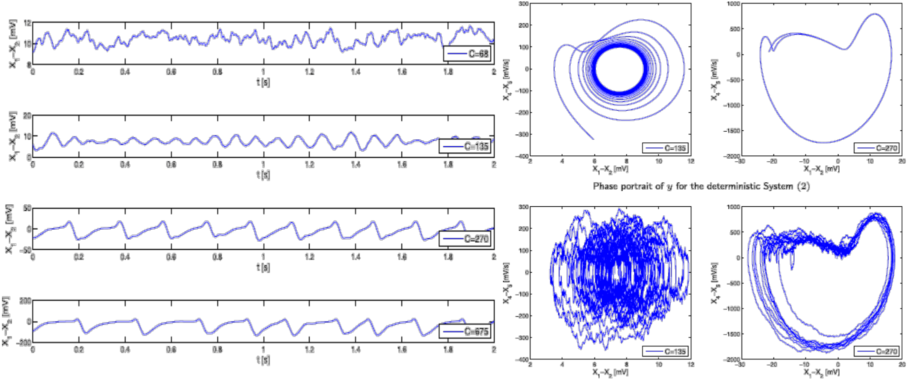
- $Q = (X_0, X_1, X_2) \in \mathbb{R}^3$, $P = (X_3, X_4, X_5) \in \mathbb{R}^3$
- $\Gamma = \mathrm{diag}[a, a, b] \in \mathbb{R}^{3\times 3}$ is a damping part,
- $\Sigma(t) = \mathrm{diag}[\sigma_3(t), \sigma_4(t), \sigma_5(t)] \in \mathbb{R}^{3\times 3}$ is a diffusion part,
- $G(t, Q)$ is a nonlinear displacement term.
- Diffusion components are of **different order** (e.g. $\sigma_3(t), \sigma_5(t) \ll \sigma_4(t)$)!

References: Jansen and Rit (1995), Ableidinger et al. (2017), Buckwar et al. (2019)

UNIVERSITÉ Grenoble Alpes

UNIVERSITÉ de Cergy-Pontoise

Example 2: Jansen and Rit Neural Mass model



Trajectories of the output signal Y for varying coupling parameter C

Phase portrait of one single path of Y for the stochastic System (5)

Figure: Source: Ableidinger et al. (2017)

Example 3: Diffusion approximation of a Hawkes process

Hawkes process (point process with memory), describing the action potentials in a population of neurons, can be approximated by a stochastic diffusion:

$$dZ_t = A(Z_t)dt + B(Z_t)dW_t.$$

- κ populations ($= \kappa$ *rough* variables)
- $\sum_{i=1}^{\kappa} \eta_i$ memory variables (*smooth* variables)
- N neurons

$$A(z) = \begin{pmatrix} -\nu_1 z^1 + z^2 \\ -\nu_1 z^2 + z^3 \\ \vdots \\ -\nu_1 z^{\eta_1+1} + c_1 f_2(z^{\eta_1+2}) \\ -\nu_2 z^{\eta_1+2} + z^{\eta_1+3} \\ \vdots \\ -\nu_n z^{\kappa} + c_n f_1(z^1) \end{pmatrix}, \quad B(z) = \frac{1}{\sqrt{N}} \begin{pmatrix} 0 & 0 \\ \vdots & \vdots \\ 0 & \frac{c_1}{\sqrt{p_2}}\sqrt{f_2(z^{\eta_1+2})} \\ 0 & 0 \\ \vdots & \vdots \\ \frac{c_n}{\sqrt{p_1}}\sqrt{f_1(z^1)} & 0 \end{pmatrix},$$

Reference: Ditlevsen and Löcherbach (2017), Chevallier (2017)

Example 3: Diffusion approximation of a Hawkes process

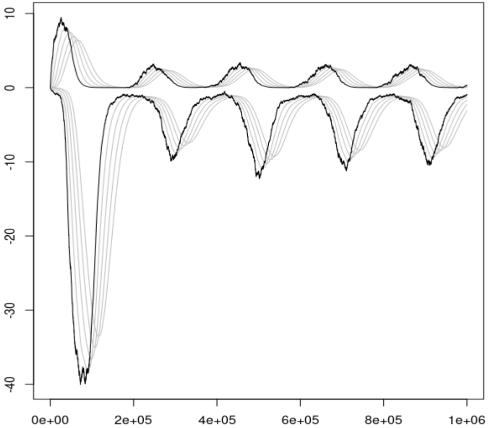


Figure: Diffusion approximation of Hawkes process describing inhibitory and excitatory neuron population (20 neurons in each population)

Where to put noise?

Main challenges:

- Highly non-linear systems
- Computational cost
- Measurements inaccuracy

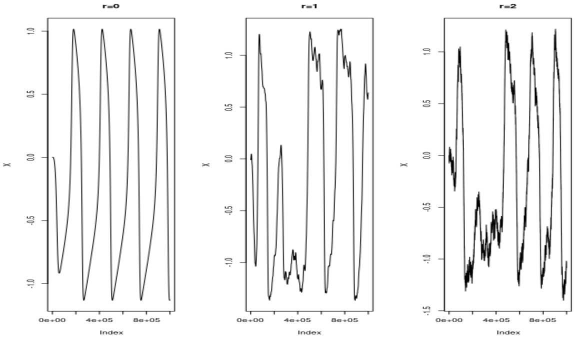


Figure: Membrane potential simulated wiht a FitzHugh-Nagumo model: deterministic, noisy channels, elliptic system

References: see Tuckwell (2005) for a general overview of neuronal models

Formalization

Given:

Discrete observations X_i of the d-dimensional process **with a fixed time step** Δ

$$dX_t = A_t dt + B_t dW_t,\ \ t \in [0, T], \tag{1}$$

$A_t \in \mathbb{R}^d$, $B_t \in \mathbb{R}^{d \times q}$.

Goal:

Propose a test

$$H_0 : rank(\Sigma) = r_0$$
$$H_1 : rank(\Sigma) \neq r_0,$$

where $\Sigma = B_t B_t^T$. If B_t is not constant, we search $\sup\limits_t \operatorname{rank}(\Sigma)$ instead.

Random perturbation approach: asymptotic setting

Main references: Jacod and Podolskij (2013) (see also Jacod et al. (2008))

Given a d-dimensional diffusion process X, consider 2 new processes:

$$\tilde{X}_t^{(k)} = X_t + \sqrt{k\Delta}\tilde{B}\tilde{W}_t,$$

where $k = 1, 2$, and $\tilde{B}$ is such that $\tilde{B}\tilde{B}^T =: \tilde{\Sigma}$ is a non-random matrix of full rank.

$$S_T^k = 2d\Delta \sum_{i=0}^{N-1} \det \begin{pmatrix} \frac{\tilde{X}_{2id+k}^{1,(k)} - \tilde{X}_{2id}^{1,(k)}}{\sqrt{k\Delta}} & \frac{\tilde{X}_{2id+2k}^{1,(k)} - \tilde{X}_{2id+k}^{1,(k)}}{\sqrt{k\Delta}} & \dots & \frac{\tilde{X}_{2id+2d}^{1,(k)} - \tilde{X}_{2id+kd-k}^{1,(k)}}{\sqrt{k\Delta}} \\ \frac{\tilde{X}_{2id+k}^{2,(k)} - \tilde{X}_{2id}^{2,(k)}}{\sqrt{k\Delta}} & \frac{\tilde{X}_{2id+2k}^{2,(k)} - \tilde{X}_{2id+k}^{2,(k)}}{\sqrt{k\Delta}} & \dots & \frac{\tilde{X}_{2id+2d}^{2,(k)} - \tilde{X}_{2id+kd-k}^{2,(k)}}{\sqrt{k\Delta}} \\ \dots & \dots & \dots & \dots \\ \frac{\tilde{X}_{2id+k}^{d,(k)} - \tilde{X}_{2id}^{d,(k)}}{\sqrt{k\Delta}} & \frac{\tilde{X}_{2id+2k}^{d,(k)} - \tilde{X}_{2id+k}^{d,(k)}}{\sqrt{k\Delta}} & \dots & \frac{\tilde{X}_{2id+2d}^{d,(k)} - \tilde{X}_{2id+kd-k}^{d,(k)}}{\sqrt{k\Delta}} \end{pmatrix}^2$$

Random perturbation approach: asymptotic setting

Define:

$$\hat{R}_{T,\Delta} = d - \frac{\log \frac{S_T^2}{S_T^1}}{\log 2}$$

$$V_{T,\Delta} := Var\left[\widehat{R}_{T,\Delta}\right] = \frac{\left(\frac{E[S_T^1]}{E[S_T^2]}\right)^2 Var[S_T^2] - 2\frac{E[S_T^1]}{E[S_T^2]} Cov[S_T^1 S_T^2] + Var[S_T^1]}{(E[S_T^1]\log 2)^2}.$$

Jacod and Podolskij (2013)

$$\frac{\widehat{R}_{T,\Delta} - r_0}{\sqrt{\Delta V_{T,\Delta}}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1), \text{ as } \Delta \to 0$$

UNIVERSITÉ Grenoble Alpes

UNIVERSITÉ de Cergy-Pontoise

How does it work: toy 1d example

Take the process:

$$dX_t = adt + \sigma dW_t$$

Add a random perturbation:

$$\tilde{X}_t^{(1)} = adt + \sigma dW_t + \sqrt{\Delta}\tilde{\sigma}\tilde{W}_t,$$
$$\tilde{X}_t^{(2)} = adt + \sigma dW_t + \sqrt{2\Delta}\tilde{\sigma}\tilde{W}_t$$

Using the first-order approximation, compute:

$$\mathbb{E}\left[\left(\frac{\tilde{X}_{i+1}^{(k)} - \tilde{X}_i^{(k)}}{\sqrt{k\Delta}}\right)^2\right] = \sigma^2 + k\Delta a + k\Delta\tilde{\sigma} =: s_i^k$$

Notice that

$$\frac{s_i^2}{s_i^1} \xrightarrow[\Delta\to 0]{} \begin{cases} 1 & \text{if } \sigma \neq 0 \\ 2 & \text{if } \sigma = 0 \end{cases}$$

How does it work: toy 1d example

Take the process:

$$dX_t = adt + \sigma dW_t$$

Add a random perturbation:

$$\tilde{X}_t^{(1)} = adt + \sigma dW_t + \sqrt{\Delta}\tilde{\sigma}\tilde{W}_t,$$
$$\tilde{X}_t^{(2)} = adt + \sigma dW_t + \sqrt{2\Delta}\tilde{\sigma}\tilde{W}_t$$

Using the first-order approximation, compute:

$$\mathbb{E}\left[\left(\frac{\tilde{X}_{i+1}^{(k)} - \tilde{X}_i^{(k)}}{\sqrt{k\Delta}}\right)^2\right] = \sigma^2 + k\Delta a + k\Delta\tilde{\sigma} =: s_i^k$$

Notice that

$$\frac{s_i^2}{s_i^1} \xrightarrow[\Delta\to 0]{} \begin{cases} 1 & \text{if } \sigma \neq 0 \\ 2 & \text{if } \sigma = 0 \end{cases} \quad \Rightarrow \quad 1 - \frac{\log \frac{S^2}{S^1}}{\log 2} \xrightarrow[\Delta\to 0]{} \begin{cases} 1 & \text{if } \sigma \neq 0 \\ 0 & \text{if } \sigma = 0 \end{cases}$$

What happens if Δ is fixed?

Take the process:

$$dX_t = adt + \sigma dW_t$$

Fix: $\Delta = 0.01$, $\sigma = 0.05$, $a = 1$, $\tilde{\sigma} = 0.01$

Add the random perturbation:

$$\tilde{X}_t^{(1)} = adt + \sigma dW_t + \sqrt{\Delta}\tilde{\sigma}\tilde{W}_t,$$
$$\tilde{X}_t^{(2)} = adt + \sigma dW_t + \sqrt{2\Delta}\tilde{\sigma}\tilde{W}_t$$

Notice that

$$\mathbb{E}\left[\frac{S^2}{S^1}\right] \approx 1.8 \quad \Rightarrow \quad \mathbb{E}\left[\hat{R}_{T,\Delta}\right] \approx 0.15$$

UNIVERSITÉ Grenoble Alpes

UNIVERSITÉ de Cergy-Pontoise

What happens if Δ is fixed?

Ongoing work

A. M., Adeline Samson, Patricia Reynaud-Bouret

Question 1: What can we actually infer in non-asymptotic setting?
Our ultimate goal is to evaluate the following probability:

$$\mathbb{P}\left(\left|\widehat{R}_{T,\Delta} - r_0\right| \geq \varepsilon\right) \tag{2}$$

Question 2: Choice of $\tilde{\Sigma}$.

Concentration inequality for 1d toy model

Consider a one-dimensional process with constant drift and diffusion coefficients:

$$dX_t = adt + \sigma dW_t$$

Statistics:

$$S^k = 2\Delta \sum_{i=0}^{N} \left(\frac{\tilde{X}_{(2i+k)\Delta} - \tilde{X}_{2i\Delta}}{\sqrt{k\Delta}} \right)^2 \tag{3}$$

S^k is difficult to treat, but luckily $\sqrt{S^k}$ is a Lipschitz function of a normal vector!

$$\mathbb{P}\left(\sqrt{S^k} - \mathbb{E}[\sqrt{S^k}] \geq \varepsilon\right) \leq e^{-\frac{\varepsilon^2}{2(\sigma^2+2\Delta\tilde{\sigma}^2)}}$$

How do we pass from bounds on S^k to bounds on $\hat{R}(T,\Delta)$:

$$\mathbb{P}\left(\frac{\mathbb{E}\sqrt{S^2}-\varepsilon}{\mathbb{E}\sqrt{S^1}+\varepsilon} \leq \frac{\sqrt{S^2}}{\sqrt{S^1}} \leq \frac{\mathbb{E}\sqrt{S^2}+\varepsilon}{\mathbb{E}\sqrt{S^1}-\varepsilon}\right) =$$

$$\mathbb{P}\left(\log\frac{\mathbb{E}\sqrt{S^2}-\varepsilon}{\mathbb{E}\sqrt{S^1}+\varepsilon} \leq \log\frac{\sqrt{S^2}}{\sqrt{S^1}} \leq \log\frac{\mathbb{E}\sqrt{S^2}+\varepsilon}{\mathbb{E}\sqrt{S^1}-\varepsilon}\right) =$$

$$\mathbb{P}\left(1-\frac{2}{\log 2}\log\frac{\mathbb{E}\sqrt{S^2}+\varepsilon}{\mathbb{E}\sqrt{S^1}-\varepsilon} \leq \widehat{R}(T,\Delta) \leq 1-\frac{2}{\log 2}\log\frac{\mathbb{E}\sqrt{S^2}-\varepsilon}{\mathbb{E}\sqrt{S^1}+\varepsilon}\right) \leq e^{-\frac{2\varepsilon^2}{(\sigma^2+2\Delta\tilde{\sigma}^2)}}$$

Concentration inequality for 1d toy model

How do we pass to confidence intervals: $\varepsilon := \sqrt{0.5(\sigma^2 + 2\Delta\tilde{\sigma}^2)\log\frac{1}{\beta}}$

Using an approximation of moments, we obtain:

$$\mathbb{P}\left(1 - \frac{2}{\log 2}\log\frac{\sqrt{1+\frac{2\Delta a^2}{\sigma^2+2\Delta\tilde{\sigma}^2}} + \sqrt{\frac{1}{2}\log\frac{1}{\beta}}}{\sqrt{1+\frac{\Delta(a^2-\tilde{\sigma}^2)}{\sigma^2+2\Delta\tilde{\sigma}^2}} - \sqrt{\frac{1}{2}\log\frac{1}{\beta}}} \leq \widehat{R}(T,\Delta) \leq \right.$$
$$\left. 1 - \frac{2}{\log 2}\log\frac{\sqrt{1+\frac{2\Delta a^2}{\sigma^2+2\Delta\tilde{\sigma}^2}} - \sqrt{\frac{1}{2}\log\frac{1}{\beta}}}{\sqrt{1+\frac{\Delta(a^2-\tilde{\sigma}^2)}{\sigma^2+2\Delta\tilde{\sigma}^2}} + \sqrt{\frac{1}{2}\log\frac{1}{\beta}}}\right) \geq 1-\beta$$

UNIVERSITÉ Grenoble Alpes

UNIVERSITÉ de Cergy-Pontoise

Non-asymptotic setting: concentration inequalities

Recall main statistics:

$$S_T^k = 2d\Delta \sum_{i=0}^{N-1} \det \begin{pmatrix} \frac{\tilde{X}^{1,(k)}_{2id+k} - \tilde{X}^{1,(k)}_{2id}}{\sqrt{k\Delta}} & \frac{\tilde{X}^{1,(k)}_{2id+2k} - \tilde{X}^{1,(k)}_{2id+k}}{\sqrt{k\Delta}} & \dots & \frac{\tilde{X}^{1,(k)}_{2id+2d} - \tilde{X}^{1,(k)}_{2id+kd-k}}{\sqrt{k\Delta}} \\ \frac{\tilde{X}^{2,(k)}_{2id+k} - \tilde{X}^{2,(k)}_{2id}}{\sqrt{k\Delta}} & \frac{\tilde{X}^{2,(k)}_{2id+2k} - \tilde{X}^{2,(k)}_{2id+k}}{\sqrt{k\Delta}} & \dots & \frac{\tilde{X}^{2,(k)}_{2id+2d} - \tilde{X}^{2,(k)}_{2id+kd-k}}{\sqrt{k\Delta}} \\ \dots & \dots & \dots & \dots \\ \frac{\tilde{X}^{d,(k)}_{2id+k} - \tilde{X}^{d,(k)}_{2id}}{\sqrt{k\Delta}} & \frac{\tilde{X}^{d,(k)}_{2id+2k} - \tilde{X}^{d,(k)}_{2id+k}}{\sqrt{k\Delta}} & \dots & \frac{\tilde{X}^{d,(k)}_{2id+2d} - \tilde{X}^{d,(k)}_{2id+kd-k}}{\sqrt{k\Delta}} \end{pmatrix}^2$$

General case: lower bound

Lemma (Lower bound for S^k)

$$\mathbb{P}\left(S^k - \mathbb{E}[S^k] \leq -\varepsilon\right) \leq \exp\left(\frac{-\varepsilon^2}{2Nv}\right),$$

where $v = \sup_i \mathbb{E}[(s_i^k)^2]$.

Good news: lower bound is sub-gaussian, because S^k is positive!

Lemma (Lower bound for S^k)

$$\mathbb{P}\left(S^k - \mathbb{E}[S^k] \leq -\varepsilon\right) \leq \exp\left(\frac{-\varepsilon^2}{2Nv}\right),$$

where $v = \sup_i \mathbb{E}[(s_i^k)^2]$.

Good news: lower bound is sub-gaussian, because S^k is positive!

Bad news: ...it depends on the expression of moments, which are difficult to compute!

General case: upper bound

Key: Hadamard inequality

$$\left(\sum_{i=0}^{N} s_i^k\right)^{\frac{1}{2d}} \leq \left(\sum_{i=0}^{N}\left(\sup_t \left\| \frac{X_{t+k\Delta}^{(k)} - X_t^{(k)}}{\sqrt{k\Delta}} \right\|\right)^{2d}\right)^{\frac{1}{2d}} \leq \sum_{i=0}^{N} \sup_t \left\| \frac{X_{t+k\Delta}^{(k)} - X_t^{(k)}}{\sqrt{k\Delta}} \right\|$$

Lemma (Upper bound for S^k)

$$\mathbb{P}\left((S^k)^{\frac{1}{2d}} - \mathbb{E}[(S^k)^{\frac{1}{2d}}] \geq \varepsilon\right) \leq \exp\left(\frac{-\varepsilon^2}{2Nv}\right),$$

where $v = \sup_t \left(BB^T + 2\Delta\tilde{\Sigma}\right)$.

Pasting it all together....

$$\mathbb{P}\left(\frac{1}{2d}\ln\frac{\mathbb{E}[S^2]-\varepsilon}{\mathbb{E}[(S^1)^{\frac{1}{2d}}]+\varepsilon} \geq \ln\frac{S^2}{S^1} \geq 2d\ln\frac{\mathbb{E}[(S^2)^{\frac{1}{2d}}]+\varepsilon}{\mathbb{E}[S^1]-\varepsilon}\right) =$$

$$\mathbb{P}\left(d - \frac{2d}{\ln 2}\ln\frac{\mathbb{E}[(S^2)^{\frac{1}{2d}}]+\varepsilon}{\mathbb{E}[S^1]-\varepsilon} \geq \hat{R}(T,\Delta) \geq d - \frac{1}{2d\ln 2}\ln\frac{\mathbb{E}[S^2]-\varepsilon}{\mathbb{E}[(S^1)^{\frac{1}{2d}}]+\varepsilon}\right)$$

$$\leq \exp\left(-\frac{2\varepsilon^2}{Nv}\right)$$

Open questions: part I

- Is it the sharpest bound? Can we get rid of N?
- Can we have a bound without having to compute $(S^k)^{\frac{1}{2d}}$?

Open questions: part I

- Is it the sharpest bound? Can we get rid of N?
- Can we have a bound without having to compute $(S^k)^{\frac{1}{2d}}$?

Once it is done...

- ...will it help us to choose a ”good” perturbation rate?
- ... does it even exist?

Numerical experiments: general setting

1. Generate 1000 trajectories with $\Delta = 1e - 5$, $T = 10$, using 1.5 strong order scheme (see Kloeden et al. (2003))
2. Subsample the data with a bigger Δ
3. Compute test statistics $\hat{R}_{T,\Delta}$ and $V_{T,\Delta}$
4. Test the "true" and a "wrong" hypothesis for a nominal level $\alpha = 0.05$
5. Report the results

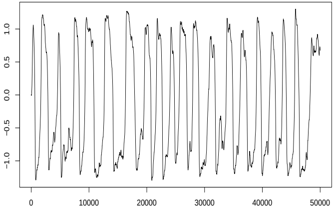
Example 1: FitzHugh-Nagumo model

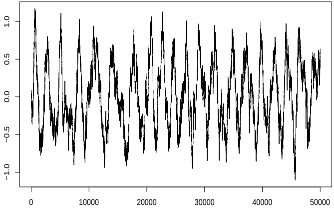
The behaviour of the neuron is defined through:

$$\begin{cases} dX_t = \frac{1}{\varepsilon}(X_t - X_t^3 - Y_t - s)dt + \sigma_1 dW_t^1 \\ dY_t = (\gamma X_t - Y_t + \beta)dt + \sigma_2 dW_t^2 \end{cases}$$

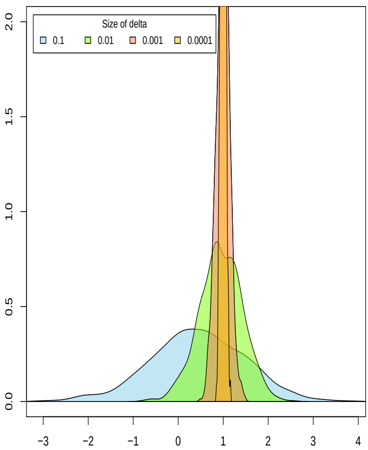
- X_t — membrane potential
- Y_t — recovery variable
- s — magnitude of the stimulus current

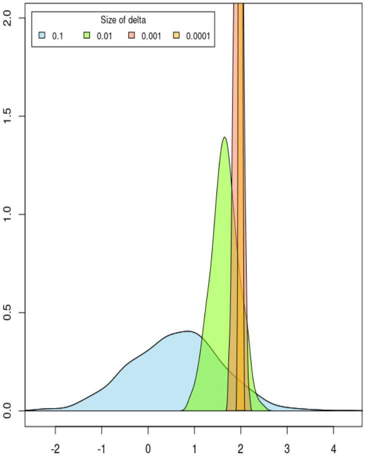
Parameters used in simulations:
$\varepsilon = 0.1, \beta = 0.3, \gamma = 1.5, s = 0.01.$



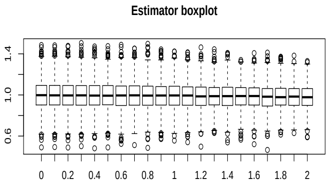


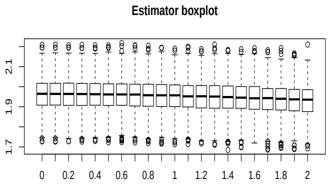
Numerical performance: FitzHugh-Nagumo model

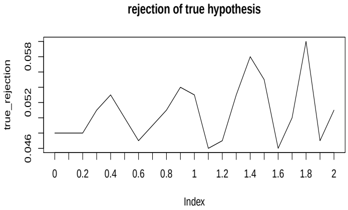


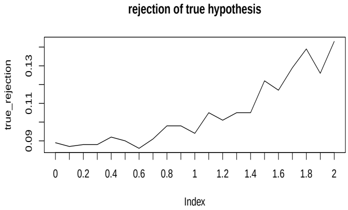


Numerical performance: FitzHugh-Nagumo model









Open questions: part II

- ... can we potentially tell the difference between the "true" (of order $\sqrt{\Delta}$) and the propagated (of order $\Delta^{\frac{3}{2}}$) noise in a given variable?
- ... what is the best empirical choice of $\tilde{\Sigma}$?

UNIVERSITÉ Grenoble Alpes

UNIVERSITÉ de Cergy-Pontoise

Literature I

Ableidinger, M., Buckwar, E., and Hinterleitner, H. (2017). A stochastic version of the jansen and rit neural mass model: Analysis and numerics. *The Journal of Mathematical Neuroscience*, 7(1):1.

Buckwar, E., Tamborrino, M., and Tubikanec, I. (2019). Spectral density-based and measure-preserving abc for partially observed diffusion processes. an illustration on hamiltonian sdes.

Chevallier, J. (2017). Mean-field limit of generalized Hawkes processes. *Stochastic Processes and their Applications*, 127(12).

Ditlevsen, S. and Löcherbach, E. (2017). Multi-class oscillating systems of interacting neurons. *SPA*, 127:1840–1869.

Fitzhugh, R. (1961). Impulses and physiological states in theoretical models of nerve membrane. *Biophysical Journal*, 1(6):445–466.

Hodgkin, A. L. and Huxley, A. F. (1952). A quantitative description of membrane currents and its application to conduction and excitation in nerve. *Journal of Physiology-London*, 117(4):500–544.

Jacod, J., Lejay, A., Talay, D., et al. (2008). Estimation of the brownian dimension of a continuous itô process. *Bernoulli*, 14(2):469–498.

Jacod, J. and Podolskij, M. (2013). A test for the rank of the volatility process: the random perturbation approach. *The Annals of Statistics*, 41(5):2391–2427.

Jansen, B. H. and Rit, V. G. (1995). Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biological Cybernetics*, 73(4):357.

Kloeden, P. E., Platen, E., and Schurz, H. (2003). *Numerical solution of SDE through computer experiments*. Universitext. Springer.

Literature II

Morris, C. and Lecar, H. (1981). Voltage oscillations in the barnacle giant muscle fiber. *Biophysical journal*, 35(1):193–213.

Tuckwell, H. C. (2005). *Introduction to theoretical neurobiology: volume 2, nonlinear and stochastic theories*, volume 8. Cambridge University Press.

Thank you for your attention!