

# Comment les statisticiens peuvent aider à une meilleure compréhension de l'activité neuronale?

F. Grammont, P. Reynaud-Bouret, A. Rouis, C. Tuleau-Malot (Nice), V. Rivoirard (Dauphine), T. Bessaïh, R. Lambert, N. Leresche, Michaël Quiquempoix (Neuronal Networks and Physiopathological Rhythms, NPA, Paris 6)

6ème rencontre Avignon - Marseille, le 8 juin 2018

# Plan

## 1 Problème

# Plan

- 1 Problème
- 2 Tests paramétriques d'indépendance

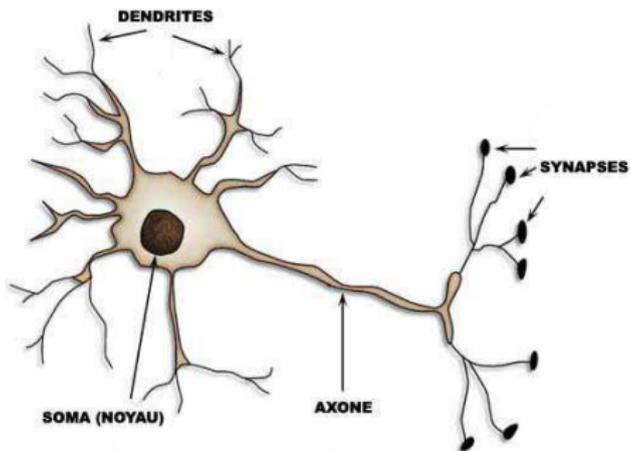
# Plan

- 1 Problème
- 2 Tests paramétriques d'indépendance
- 3 Tests d'adéquation

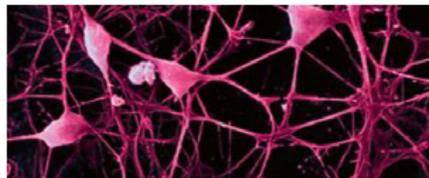
# Plan

- 1 Problème
- 2 Tests paramétriques d'indépendance
- 3 Tests d'adéquation
- 4 Indépendance locale et Processus de Hawkes

## Contexte biologique

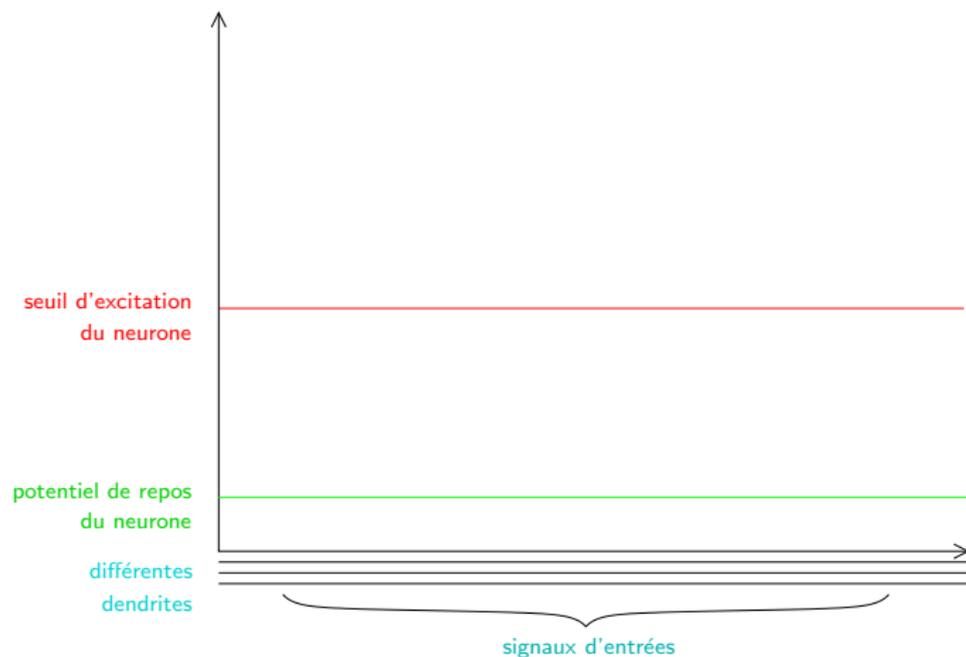


(a) Un neurone

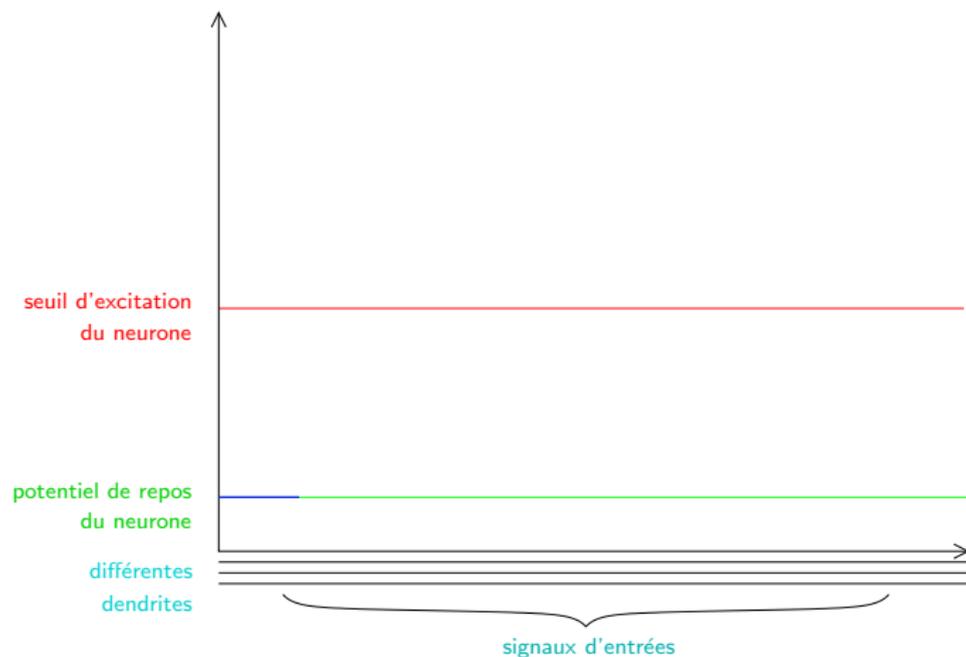


(b) Un réseau de neurones

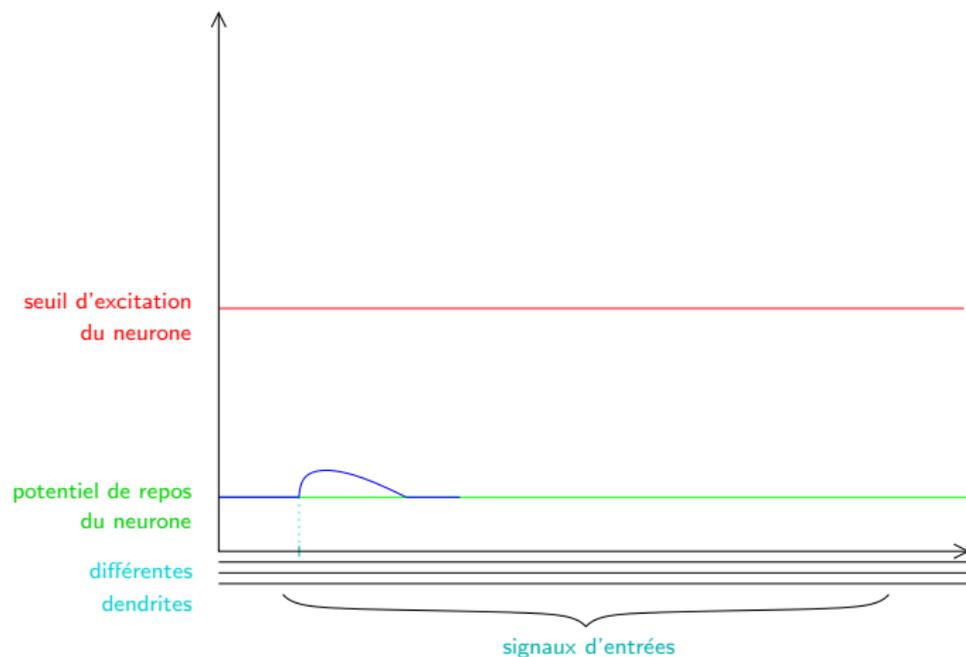
# Emission d'un potentiel d'action sans synchronisation



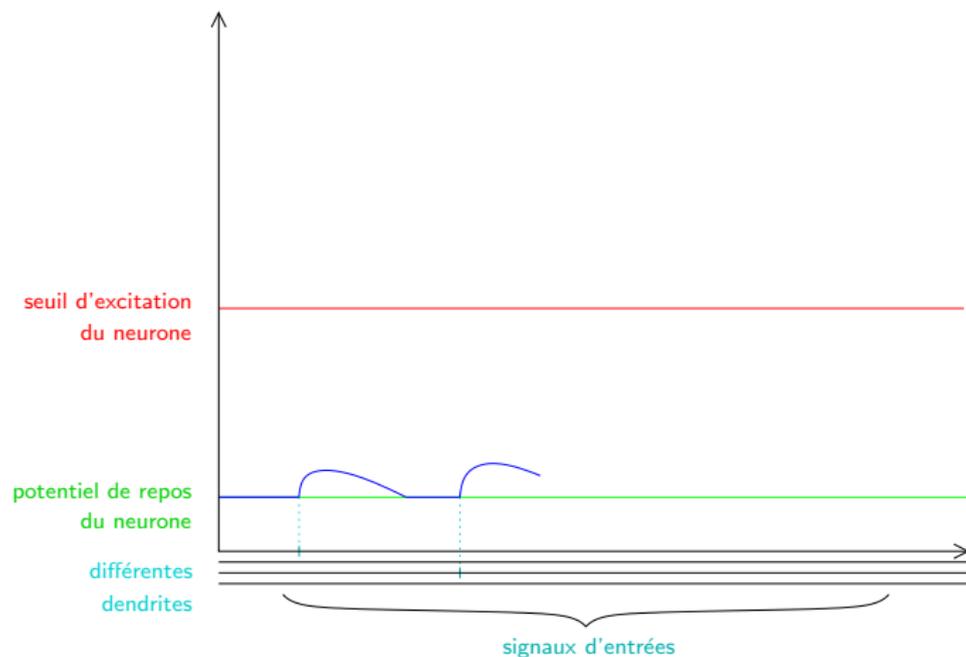
# Emission d'un potentiel d'action sans synchronisation



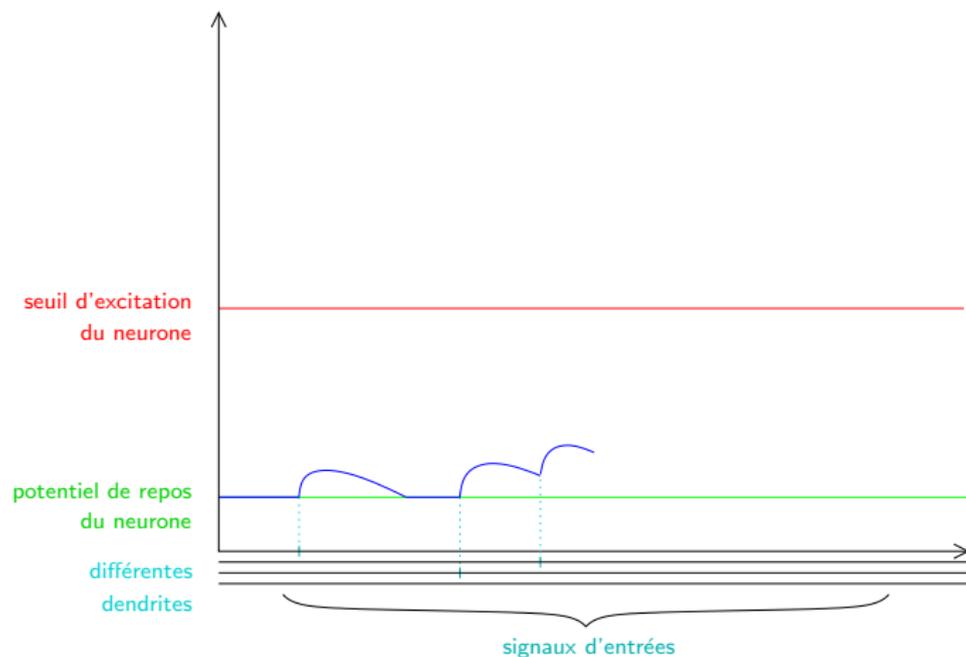
# Emission d'un potentiel d'action sans synchronisation



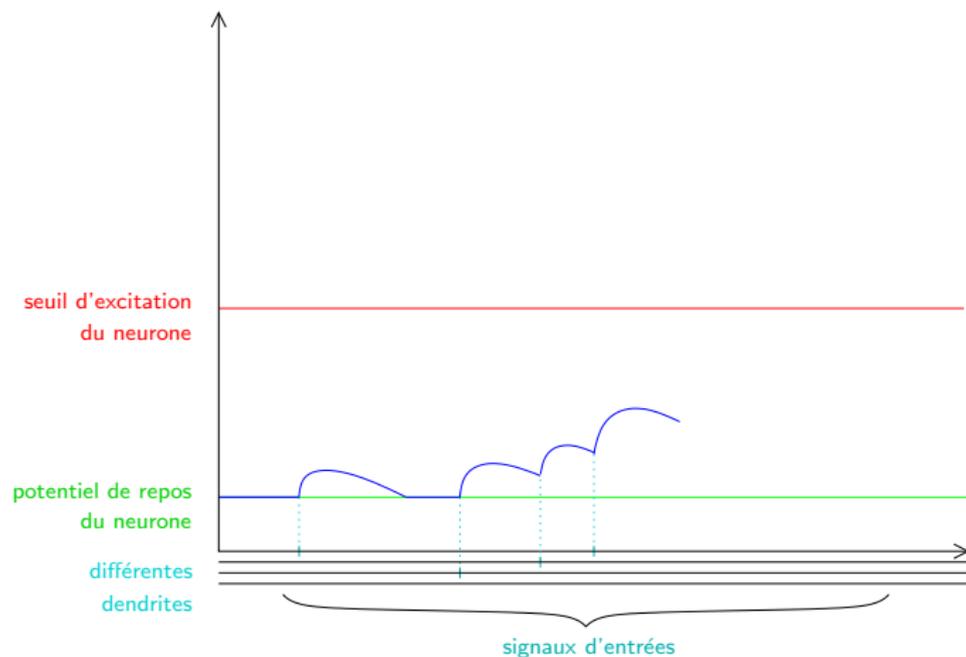
# Emission d'un potentiel d'action sans synchronisation



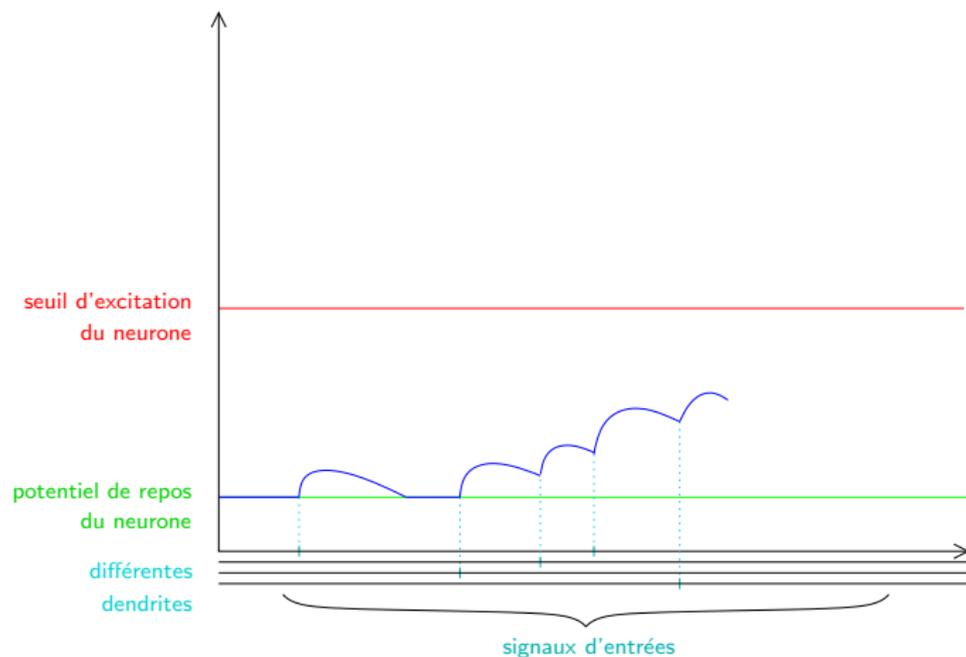
# Emission d'un potentiel d'action sans synchronisation



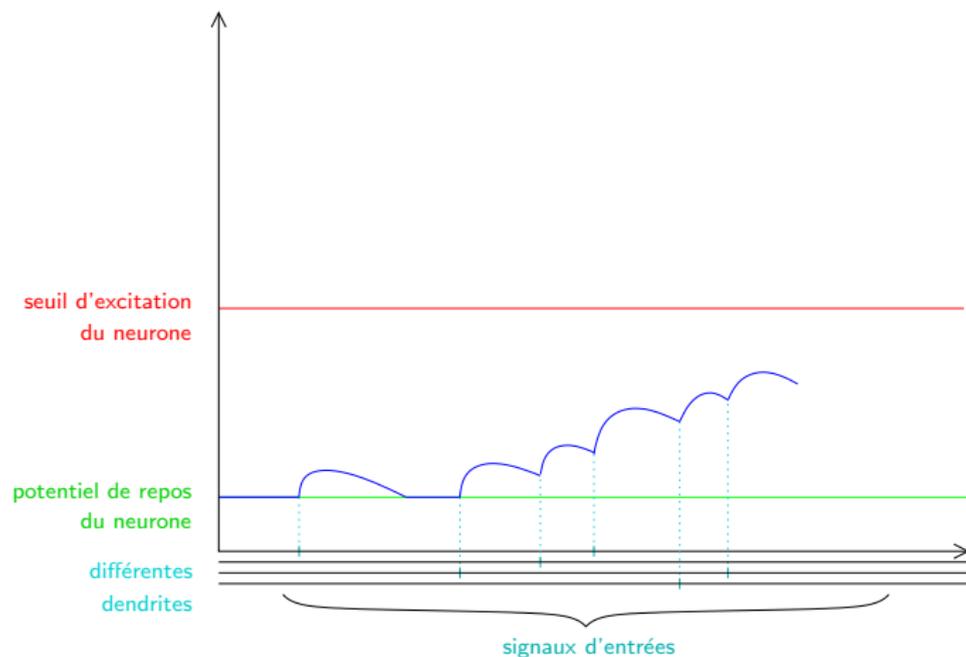
# Emission d'un potentiel d'action sans synchronisation



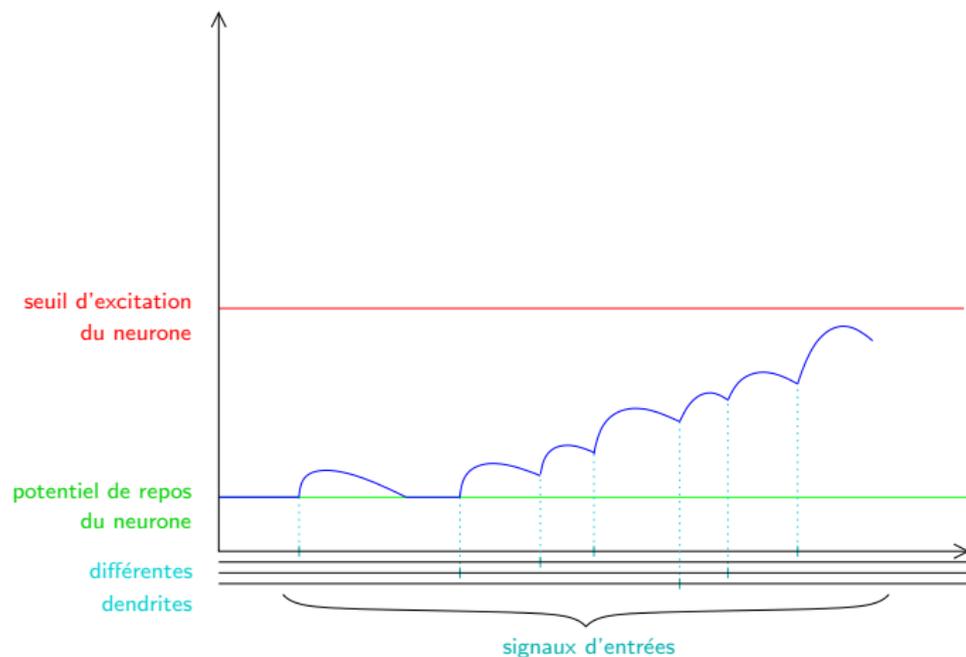
# Emission d'un potentiel d'action sans synchronisation



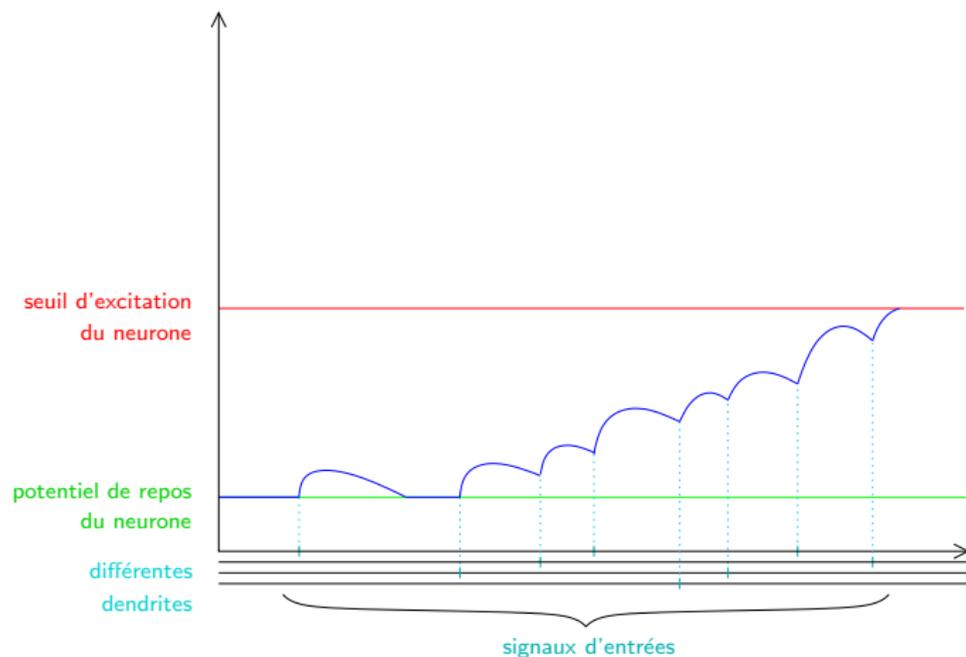
# Emission d'un potentiel d'action sans synchronisation



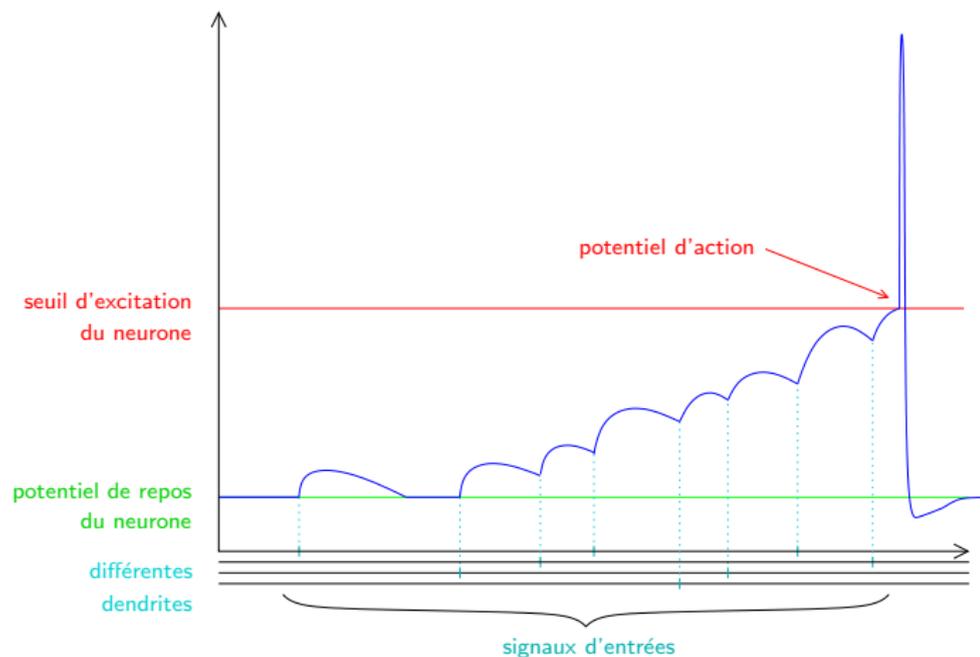
# Emission d'un potentiel d'action sans synchronisation



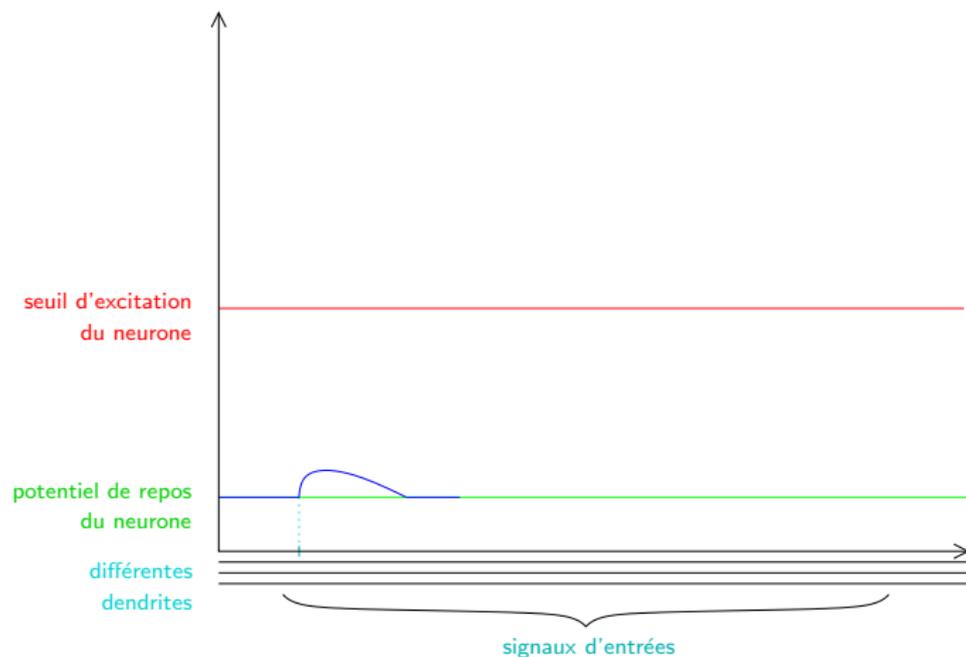
# Emission d'un potentiel d'action sans synchronisation



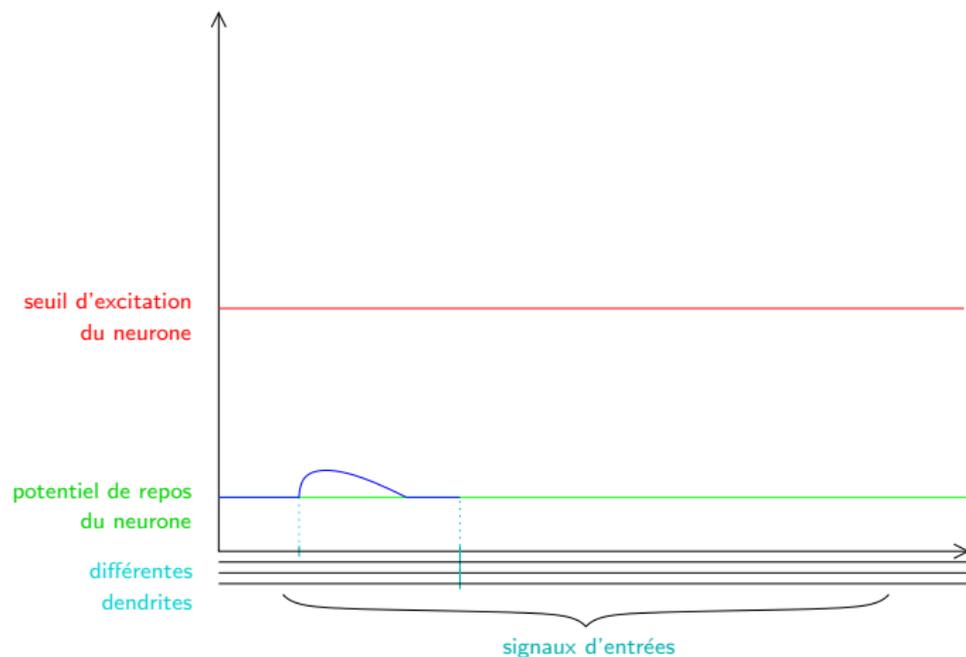
# Emission d'un potentiel d'action sans synchronisation



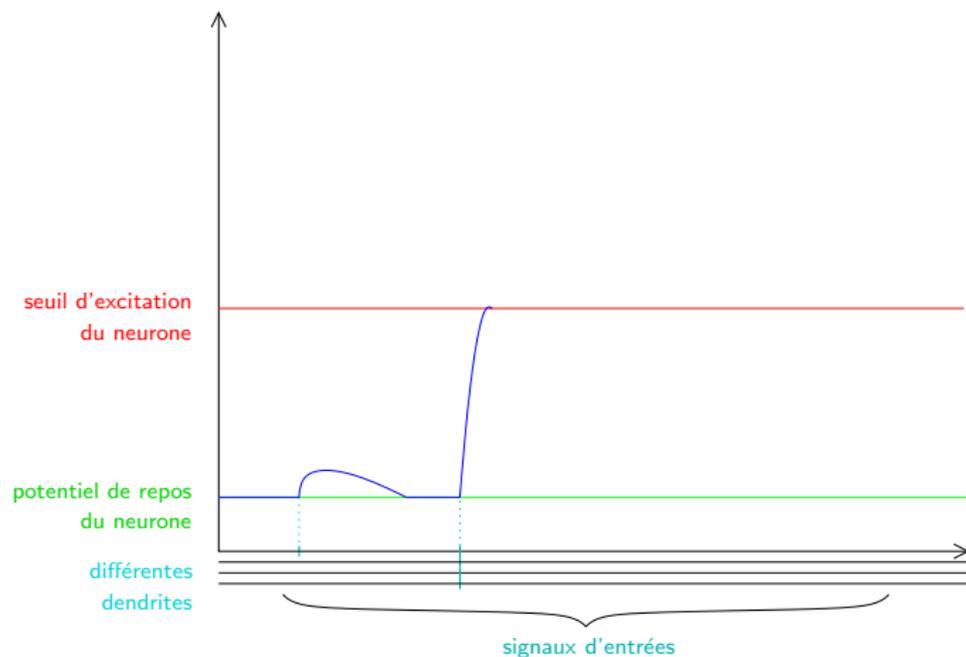
# Emission d'un potentiel d'action avec synchronisation



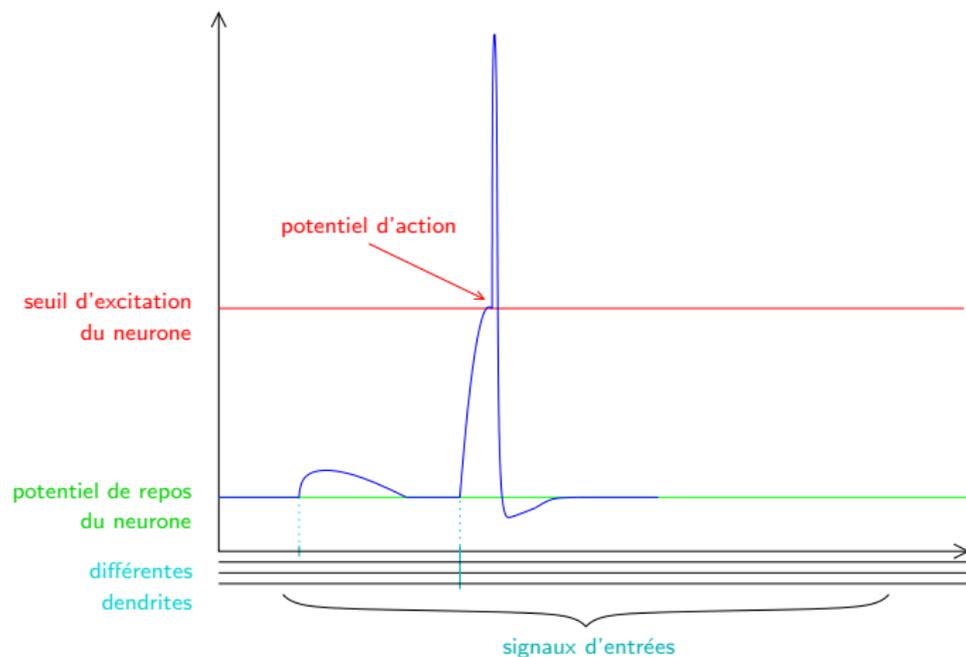
# Emission d'un potentiel d'action avec synchronisation



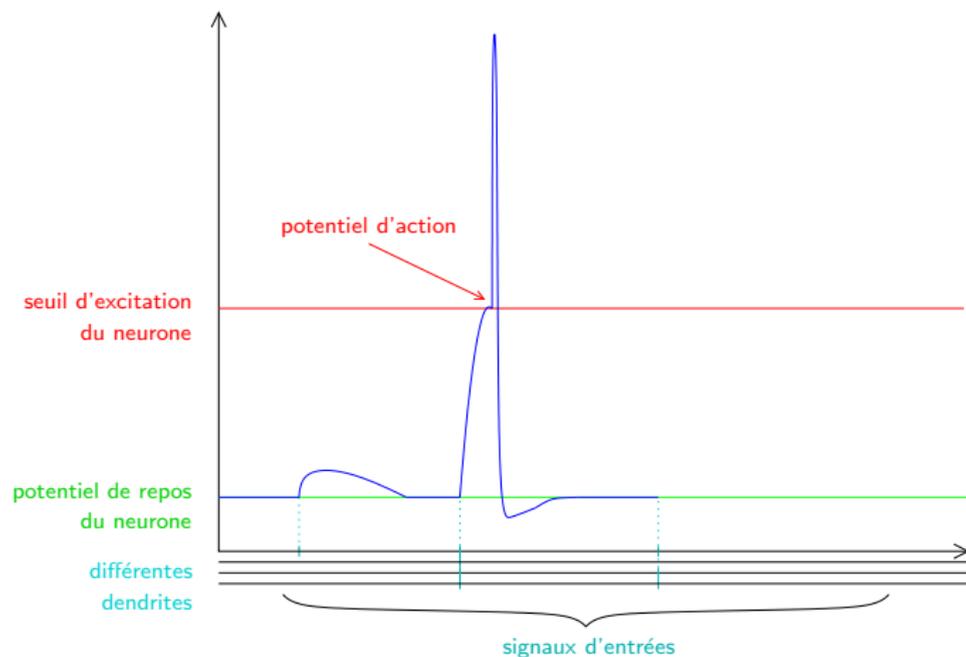
# Emission d'un potentiel d'action avec synchronisation



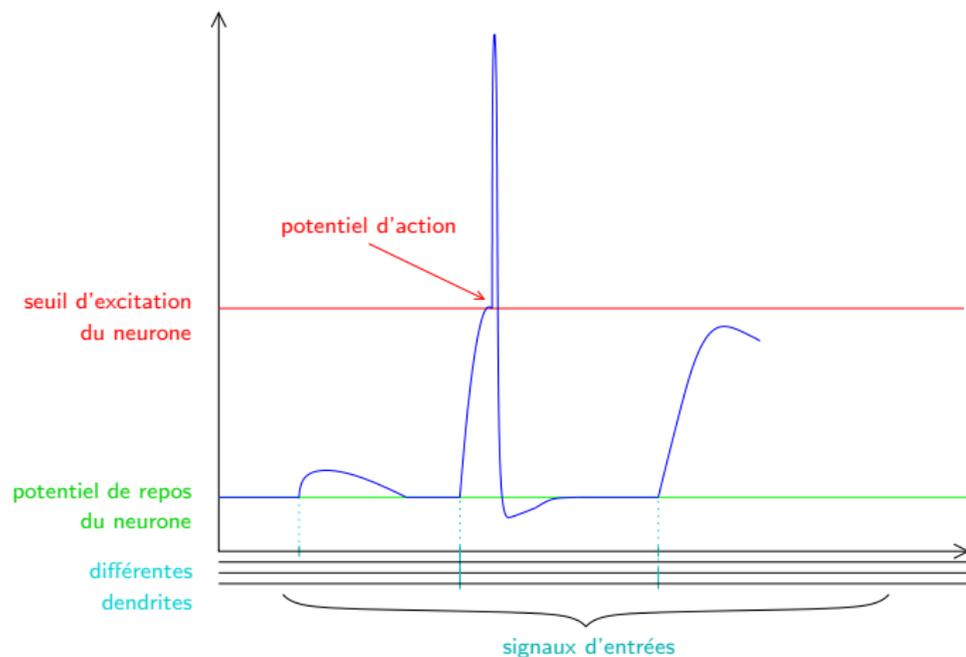
# Emission d'un potentiel d'action avec synchronisation



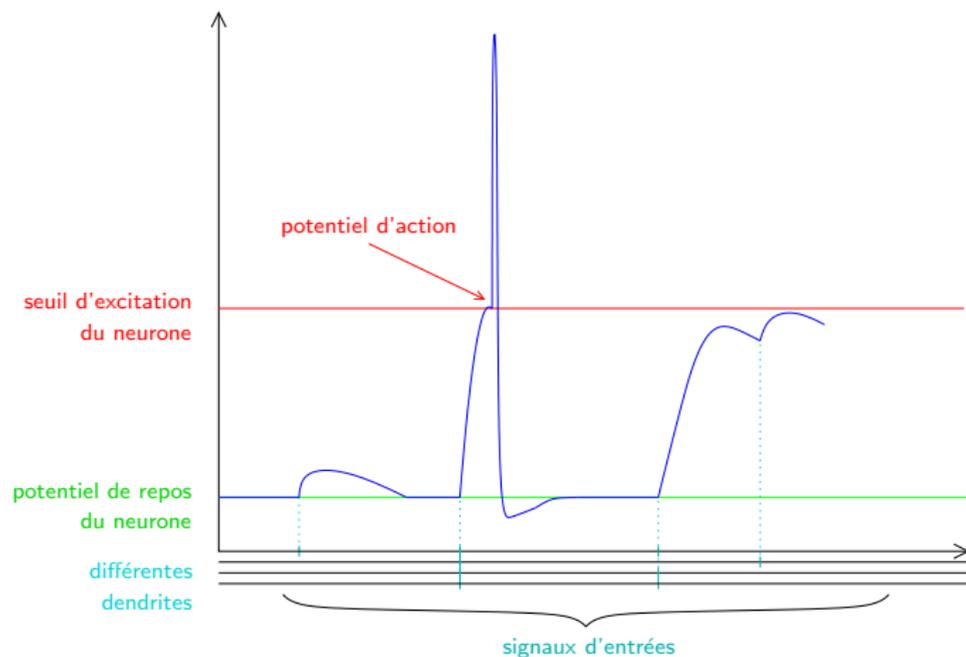
# Emission d'un potentiel d'action avec synchronisation



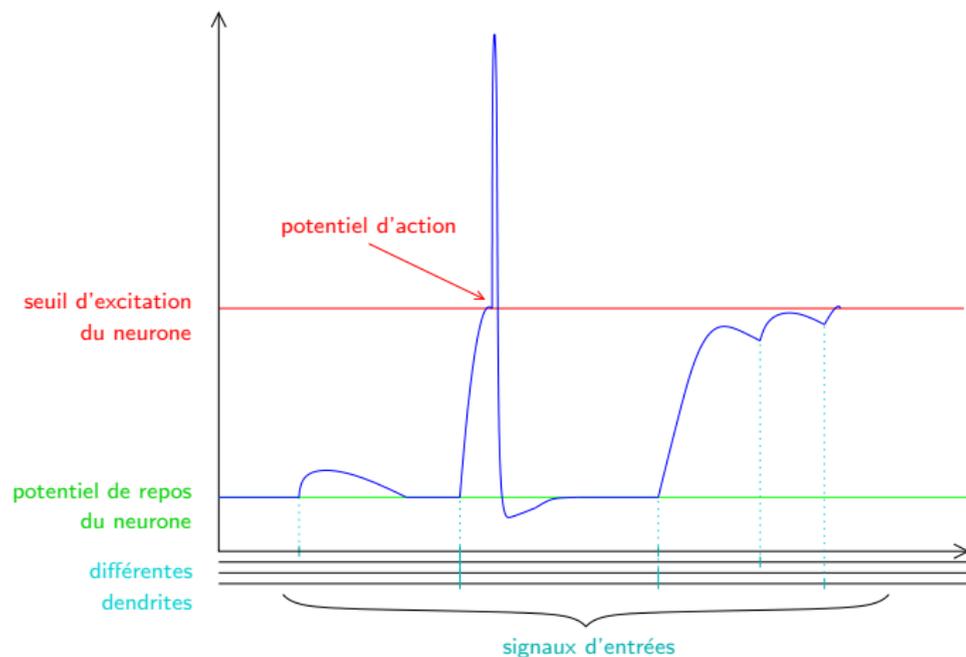
# Emission d'un potentiel d'action avec synchronisation



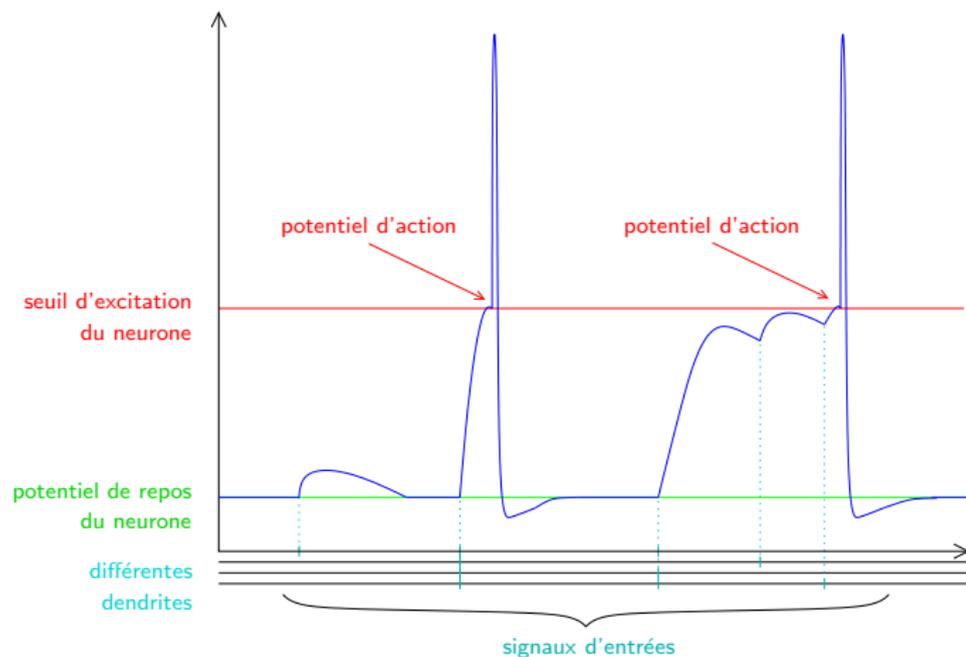
# Emission d'un potentiel d'action avec synchronisation



# Emission d'un potentiel d'action avec synchronisation

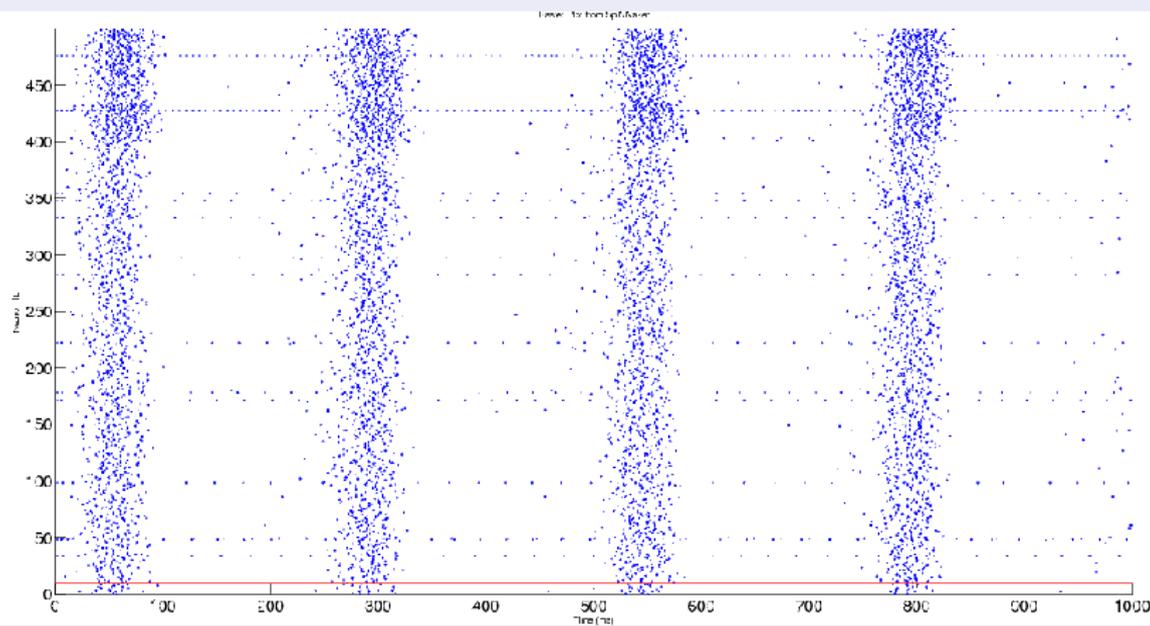


# Emission d'un potentiel d'action avec synchronisation



# Comment détecter un phénomène de synchronisation sur une paire de neurones ?

## Activité neuronale



# Comment détecter un phénomène de synchronisation sur une paire de neurones ?

## Activité neuronale

potentiel d'action (spike) = séquence de 0-1  
⇒ processus ponctuel

# Comment détecter un phénomène de synchronisation sur une paire de neurones ?

## Activité neuronale

potentiel d'action (spike) = séquence de 0-1  
⇒ processus ponctuel

## Caractéristiques

taux de décharge / fréquence / intensité stimulation forte  
⇒ beaucoup de spikes ⇒ plus de réaction.

Synchronisation des spikes si des spikes arrivent en même temps  
⇒ moins d'énergie nécessaire.

# Comment détecter un phénomène de synchronisation sur une paire de neurones ?

## Activité neuronale

potentiel d'action (spike) = séquence de 0-1  
⇒ processus ponctuel

## Caractéristiques

taux de décharge / fréquence / intensité stimulation forte  
⇒ beaucoup de spikes ⇒ plus de réaction.

Synchronisation des spikes si des spikes arrivent en même temps  
⇒ moins d'énergie nécessaire.

**[Grün et al.]** : ce phénomène de synchronisation survient à des moments précis de l'expérience, moments qui doivent être détectés.

# Comment détecter un phénomène de synchronisation sur une paire de neurones ?

## Activité neuronale

potentiel d'action (spike) = séquence de 0-1  
⇒ processus ponctuel

## Caractéristiques

taux de décharge / fréquence / intensité stimulation forte  
⇒ beaucoup de spikes ⇒ plus de réaction.

Synchronisation des spikes si des spikes arrivent en même temps  
⇒ moins d'énergie nécessaire.

**[Grün et al.]** : ce phénomène de synchronisation survient à des moments précis de l'expérience, moments qui doivent être détectés.

## Problème statistique

Identifier la rupture d'indépendance dans l'activité de deux neurones.

## Notion de coïncidences

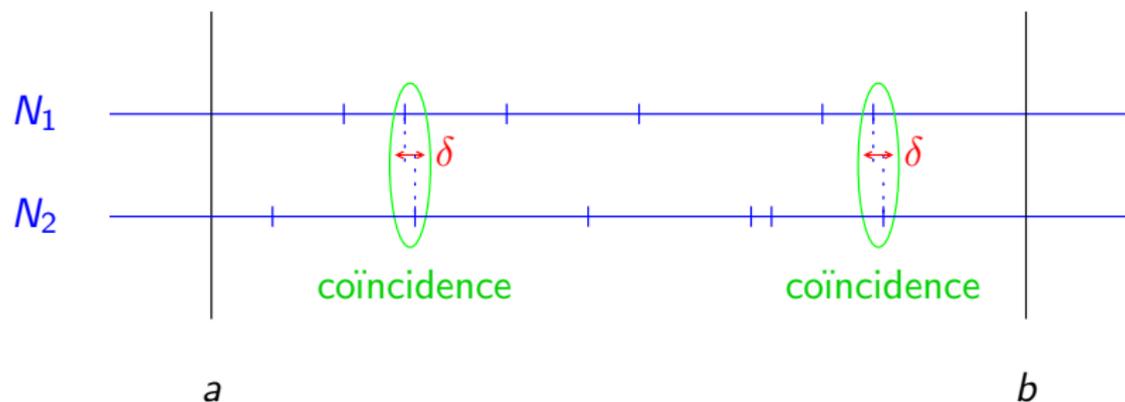
La notion de coïncidence possède plusieurs définitions. Ces définitions dépendent généralement de la modélisation introduite pour l'activité neuronale et des pré-traitements initiés comme le binning.

Généralement, les définitions souffraient d'inconvénients.

## Notion de coïncidences

### Coïncidence

Pour  $\delta \in ]0, (b - a)/2[$ , une coïncidence survient lorsqu'un spike de  $X^1$  est à une distance plus petite que  $\delta$  d'un spike de  $X^2$ .



## Notion de coïncidences

### Le nombre de coïncidences

$$\varphi(X^1, X^2) = \int_a^b \int_a^b \mathbf{1}_{|x-y| \leq \delta} dX^1(x) dX^2(y) = \sum_{x \in X^1, y \in X^2, (x,y) \in [a,b]^2} \mathbf{1}_{|x-y| \leq \delta}.$$

On observe  $n$  enregistrements et à l'aide de la statistique

$$\mathbf{C}^{obs} = \mathbf{C}(\mathbb{X}_n) = \sum_{i=1}^n \varphi(X_i^1, X_i^2),$$

on veut tester  $H_0$  : "  $X^1$  et  $X^2$  sont indépendants" contre  $H_1$  : "ils ne le sont pas".

## L'approximation gaussienne des Unitary Events

On suppose que  $X^1$  et  $X^2$  sont des processus de Poisson stationnaires.

- $\bar{C} = \frac{1}{n} \mathbf{C}^{obs}$  : moyenne empirique du nombre de coïncidences,
- $\hat{\lambda}_i = \frac{1}{n(b-a)} \sum_{k=1}^n X_i^k[a, b]$  : estimée de l'intensité du processus de Poisson associé au neurone  $i$   
↔ Estimation du nombre moyen de coïncidences sous l'hypothèse nulle

# L'approximation gaussienne des Unitary Events

## Statistique de Test

$$T_n = \frac{\sqrt{n} \left( \bar{C} - (2\delta(b-a) - \delta^2) \hat{\lambda}_1 \hat{\lambda}_2 \right)}{\sqrt{\hat{\lambda}_1 \hat{\lambda}_2 \left[ 2\delta(b-a) - \delta^2 + \left( \frac{2}{3}\delta^3 - \frac{\delta^4}{b-a} \right) (\hat{\lambda}_1 + \hat{\lambda}_2) \right]}}.$$

# L'approximation gaussienne des Unitary Events

## Statistique de Test

$$T_n = \frac{\sqrt{n} \left( \bar{C} - (2\delta(b-a) - \delta^2) \hat{\lambda}_1 \hat{\lambda}_2 \right)}{\sqrt{\hat{\lambda}_1 \hat{\lambda}_2 [2\delta(b-a) - \delta^2 + (\frac{2}{3}\delta^3 - \frac{\delta^4}{b-a})(\hat{\lambda}_1 + \hat{\lambda}_2)]}}.$$

## Loi asymptotique sous $(H_0)$ .

Si  $X^1$  et  $X^2$  sont deux processus de Poisson stationnaires et si les enregistrements successifs sont indépendants, alors sous  $(H_0)$ ,

$$T_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1).$$

## Procédure de tests multiples (MTGAUE)

La détection doit être locale

⇒ on divise l'enregistrement total en petites sous-fenêtres sur lesquelles on applique le test

⇒ plusieurs tests sont réalisés pour savoir s'il y a rupture dans le temps de l'indépendance

⇒ application de la procédure de tests multiples de Benjamini et Hochberg.

## Pourquoi cette correction?

Notons  $K$  le nombre de sous-fenêtres considérées.

On a alors :

$$\mathbb{E}(\text{nb de rejets}) = K\mathbb{P}(\text{un test rejette}) \xrightarrow{n \rightarrow +\infty} K.\alpha$$

## Pourquoi cette correction?

Notons  $K$  le nombre de sous-fenêtres considérées.

On a alors :

$$\mathbb{E}(\text{nb de rejets}) = K\mathbb{P}(\text{un test rejette}) \xrightarrow{n \rightarrow +\infty} K.\alpha$$

Donc, si on veut un test global de niveau  $\alpha$ , il ne faut pas avoir des tests de niveau  $\alpha$  sur les sous-fenêtres, mais de niveau  $\frac{\alpha}{K}$ .

## Pourquoi cette correction?

Notons  $K$  le nombre de sous-fenêtres considérées.

On a alors :

$$\mathbb{E}(\text{nb de rejets}) = K\mathbb{P}(\text{un test rejette}) \xrightarrow{n \rightarrow +\infty} K.\alpha$$

Donc, si on veut un test global de niveau  $\alpha$ , il ne faut pas avoir des tests de niveau  $\alpha$  sur les sous-fenêtres, mais de niveau  $\frac{\alpha}{K}$ .

Si sur ces  $K$  sous-fenêtres, il y en a  $L$  disjointes, alors :

$$\mathbb{P}(K \text{ tests acceptent}) \leq \mathbb{P}(L \text{ tests acceptent}) \xrightarrow{n \rightarrow +\infty} (1-\alpha)^L \xrightarrow{L \rightarrow +\infty} 0$$

## Pourquoi cette correction?

Notons  $K$  le nombre de sous-fenêtres considérées.

On a alors :

$$\mathbb{E}(\text{nb de rejets}) = K\mathbb{P}(\text{un test rejette}) \xrightarrow{n \rightarrow +\infty} K.\alpha$$

Donc, si on veut un test global de niveau  $\alpha$ , il ne faut pas avoir des tests de niveau  $\alpha$  sur les sous-fenêtres, mais de niveau  $\frac{\alpha}{K}$ .

Si sur ces  $K$  sous-fenêtres, il y en a  $L$  disjointes, alors :

$$\mathbb{P}(K \text{ tests acceptent}) \leq \mathbb{P}(L \text{ tests acceptent}) \xrightarrow{n \rightarrow +\infty} (1-\alpha)^L \xrightarrow{L \rightarrow +\infty} 0$$

Donc, quand  $L$  grandit, il y a au moins un test de rejeté par la procédure.

## Quelle est cette correction?

Number of $W$ such that	$\Delta_W$ accepts	$\Delta_W$ rejects	Total
Independence on $W$	$T_{nd}$ = "number of correct non discoveries"	$F_d$ = "number of false discoveries"	$K_0$ = "number of windows where independence is satisfied"
Dependence on $W$	$F_{nd}$ = "number of false discoveries"	$T_d$ = "number of correct discoveries"	$K_1$ = "number of windows where dependence exists"
Total	$K - R$	$R$ = "discoveries"	$K$

## Quelle est cette correction?

### FDR

Le *taux de faux positif* est défini par :

$$\text{FDR} = \mathbb{E} \left( \frac{F_d}{R} \mathbf{1}_{R>0} \right).$$

## Quelle est cette correction?

### FDR

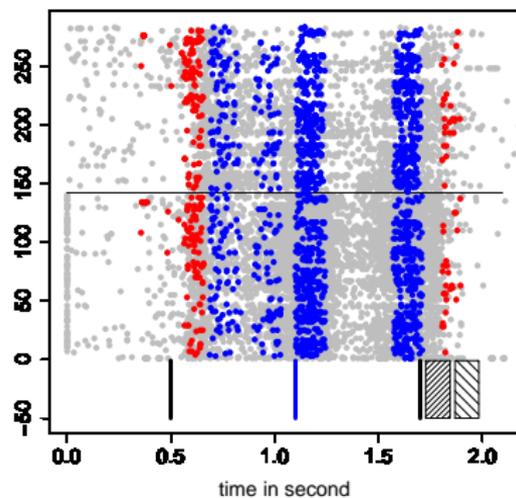
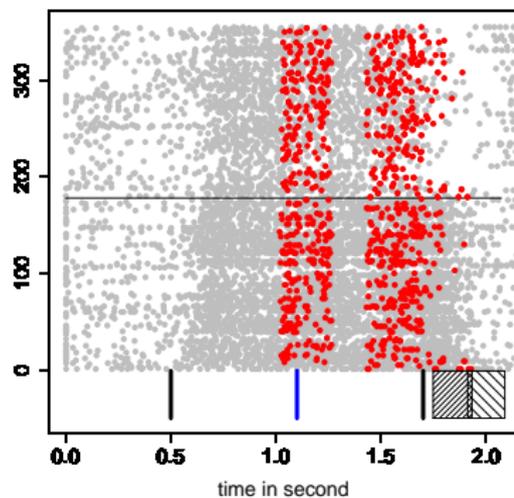
Le *taux de faux positif* est défini par :

$$\text{FDR} = \mathbb{E} \left( \frac{F_d}{R} \mathbf{1}_{R>0} \right).$$

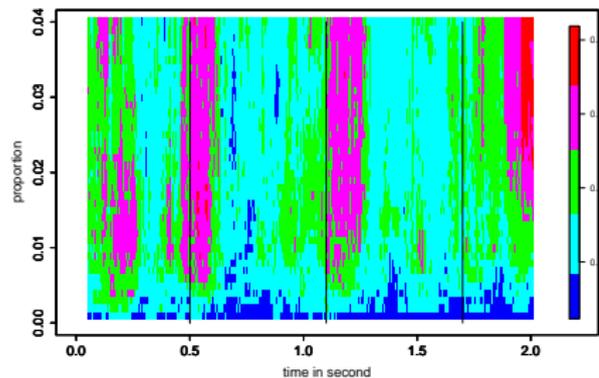
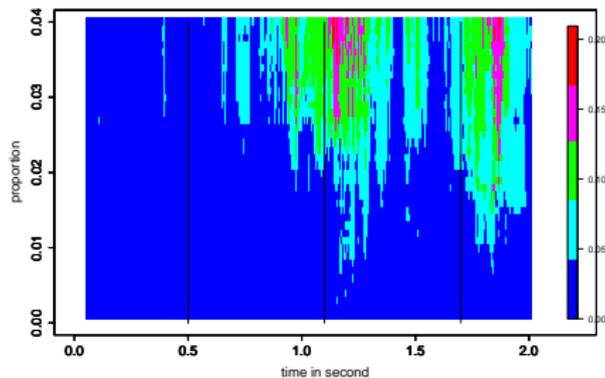
### Procédure :

- on se fixe un paramètre  $q$  pour contrôler le *FDR*
- on applique le test sur toutes les sous-fenêtres  
⇒ une collection de p-valeurs  $p_{W_1}, \dots, p_{W_K}$
- on ordonne ces p-valeurs :  $p_{W^{(1)}}^{(1)} \leq p_{W^{(2)}}^{(2)} \leq \dots \leq p_{W^{(K)}}^{(K)}$
- on choisit  $k$  tel que  $k = \max\{l \text{ such that } p_{W^{(l)}}^{(l)} \leq lq/K\}$
- les sous-fenêtres sur lesquelles on rejette ( $H_0$ ) sont  $W(1), \dots, W(k)$ .

# Présentation des résultats



# Présentation des résultats (2)



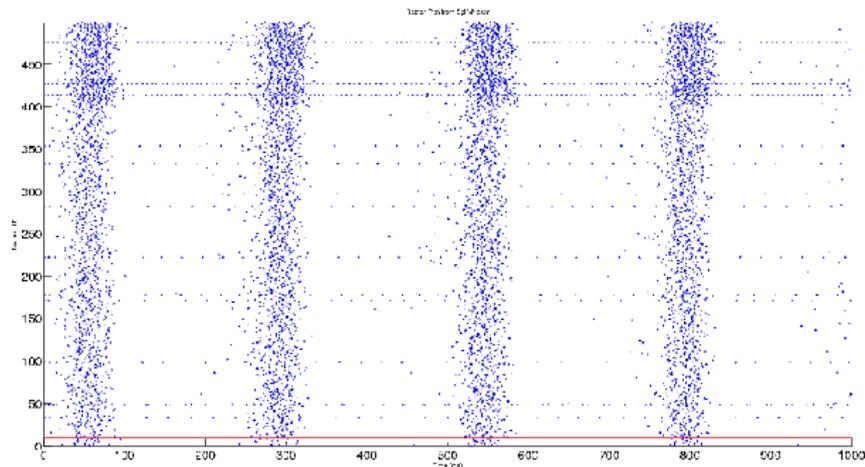
## Premières conclusions

- MTGAUE (TM, Rouis, Grammont, Reynaud-Bouret) améliore la méthode Unitary Events développée par (Grün et al):
  - pas de perte dans la détection des synchronisations
  - le niveau du test est garanti
- Il peut être envisagé d'étendre cette idée à plus de 2 neurones.

⇒ La difficulté est que la sensibilité à l'hypothèse de Poisson devient très forte et le contrôle du FDR devient très complexe.

# Modélisation par un processus de Hawkes

- Processus de Poisson homogène : inadapté



## Modélisation par un processus de Hawkes

- Processus de Poisson homogène : inadapté

Test du caractère poissonnien :

○ correspond à une p-valeur dans  $[10^{-2}, 10^{-1})$ ,

△ à une p-valeur dans  $[10^{-3}, 10^{-2})$ ,

△△ à une p-valeur dans  $[10^{-4}, 10^{-3})$ ,

△△△ à une p-valeur dans  $[0, 10^{-4})$

Les symboles sont noircis si la p-valeur correspond à un rejet par la méthode de Benjamini et Hochberg

Directions	1	2	3	4	5	6	Pooled
N1A	○		○	●	▲		▲▲▲
N2A				▲▲	○	○	▲▲▲
N1B	○	○		○			▲▲▲
N2B	○	●					▲▲▲

## Modélisation par un processus de Hawkes

- Processus de Poisson homogène : inadapté

Test du caractère uniforme :

Directions	1	2	3	4	5	6	Pooled
N1A	▲▲▲	▲▲▲	▲▲▲	•	▲▲▲	▲▲	▲▲▲
N2A	▲▲▲	▲▲▲	▲▲▲	▲▲▲	▲▲▲	▲▲▲	▲▲▲
N1B	▲▲▲	▲▲▲	▲▲▲	▲▲▲	▲▲▲	▲▲▲	▲▲▲
N2B	▲▲▲	▲▲▲	▲▲▲	▲▲▲	▲▲▲	▲	▲▲▲

## Modélisation par un processus de Hawkes

- Processus de Poisson homogène : inadapté

Test du caractère exponentiel :

Directions	1	2	3	4	5	6	Pooled
N1A	▲▲▲	▲▲▲	▲	●		○	▲▲▲
N2A	▲▲▲	▲▲▲	▲▲▲	▲	▲▲▲	▲▲▲	▲▲▲
N1B	▲▲▲	▲▲▲	▲▲▲	●	▲▲▲	▲▲▲	▲▲▲
N2B	●	▲▲▲	▲	○	▲▲	▲	▲▲▲

# Modélisation par un processus de Hawkes

- Processus de Poisson homogène : inadapté
- Processus de Poisson inhomogène : inadapté

## Modélisation par un processus de Hawkes

- Processus de Poisson homogène : inadapté
- Processus de Poisson inhomogène : inadapté
- Processus de Hawkes

Cette modélisation a déjà été introduite dans la littérature :

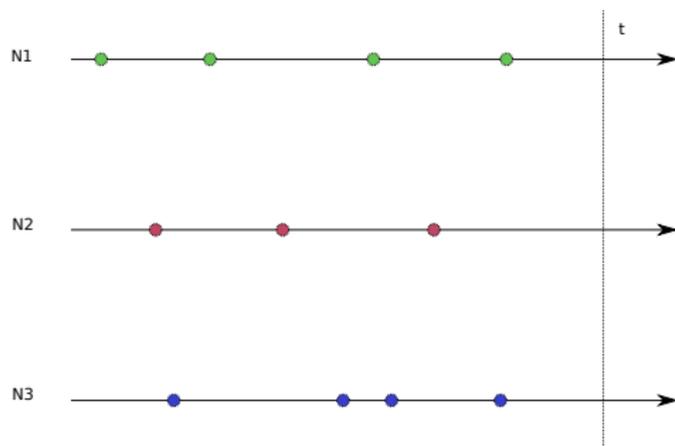
M. Krumin, I. Reutsky, S. Shoham, "Correlation-based analysis and generation of multiple spike trains using Hawkes models with an exogenous input.", *Frontiers in Computational Neuroscience*, 4(article 147), 2010

# Processus ponctuel et intensité conditionnelle

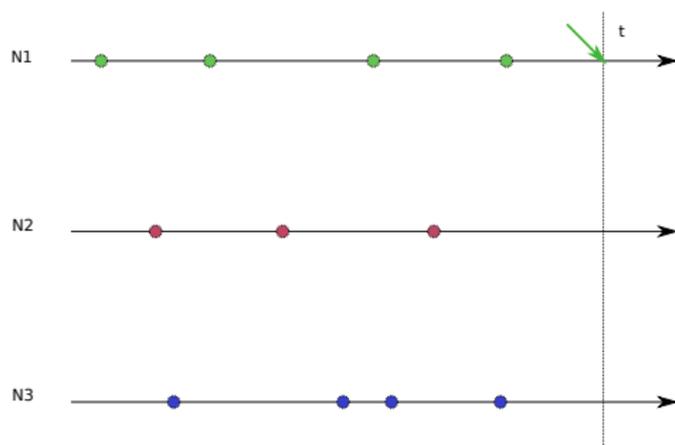
$$\underbrace{dN_t}_{\text{Nbr de points observés dans } [t, t + dt]} = \underbrace{\lambda(t) dt}_{\substack{\text{intensité} \\ \text{Nombre attendu} \\ \text{le passé avant } t \text{ donné}}} + \text{bruit}$$

$$\begin{aligned} \lambda(t) &= \text{fréquence instantanée} \\ &= \text{aléatoire, dépend des points précédents} \end{aligned}$$

# Processus de Hawkes multivarié

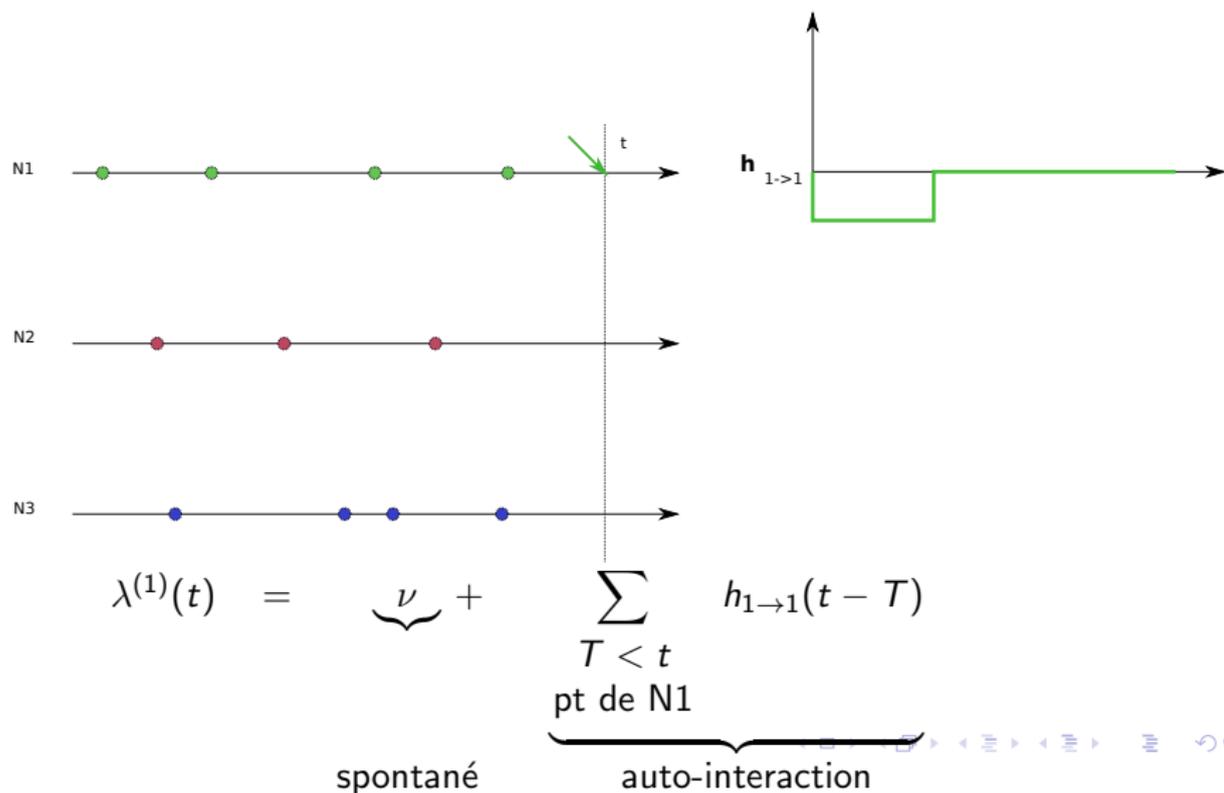


# Processus de Hawkes multivarié

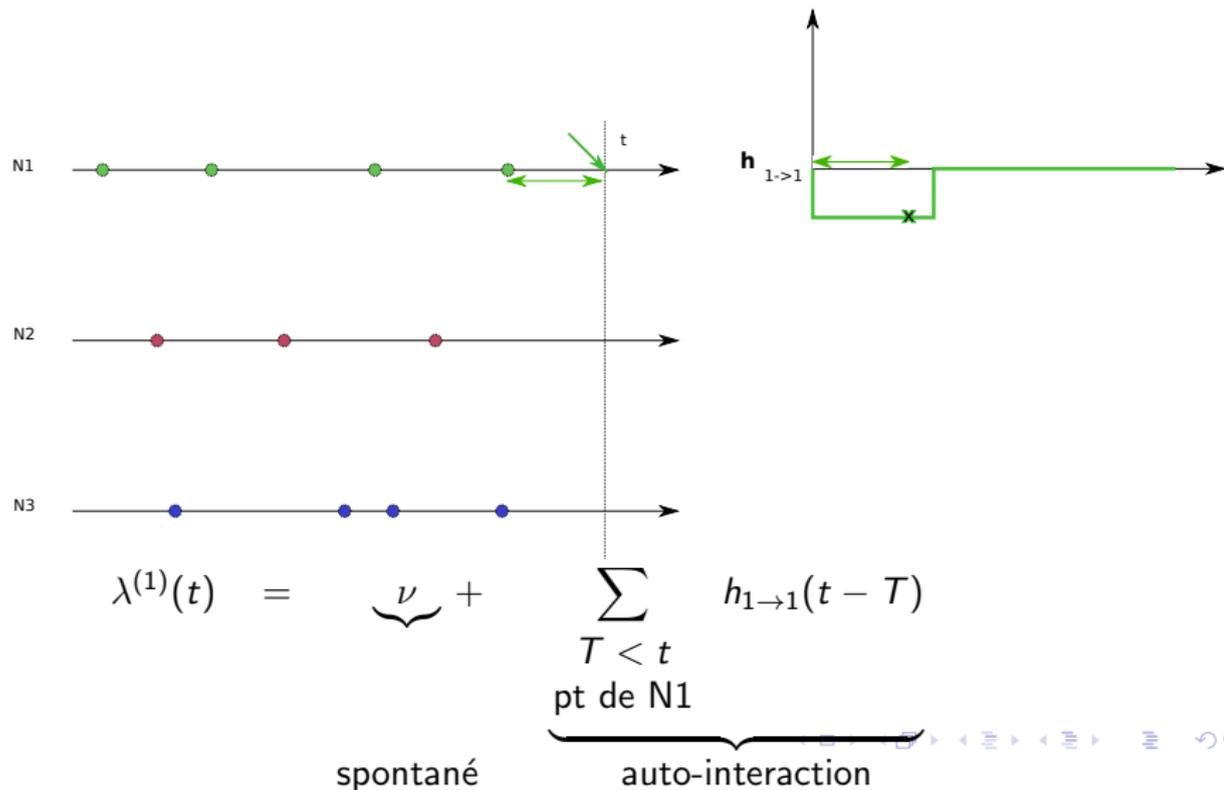


$$\lambda^{(1)}(t) = \underbrace{\nu}_{\text{spontané}}$$

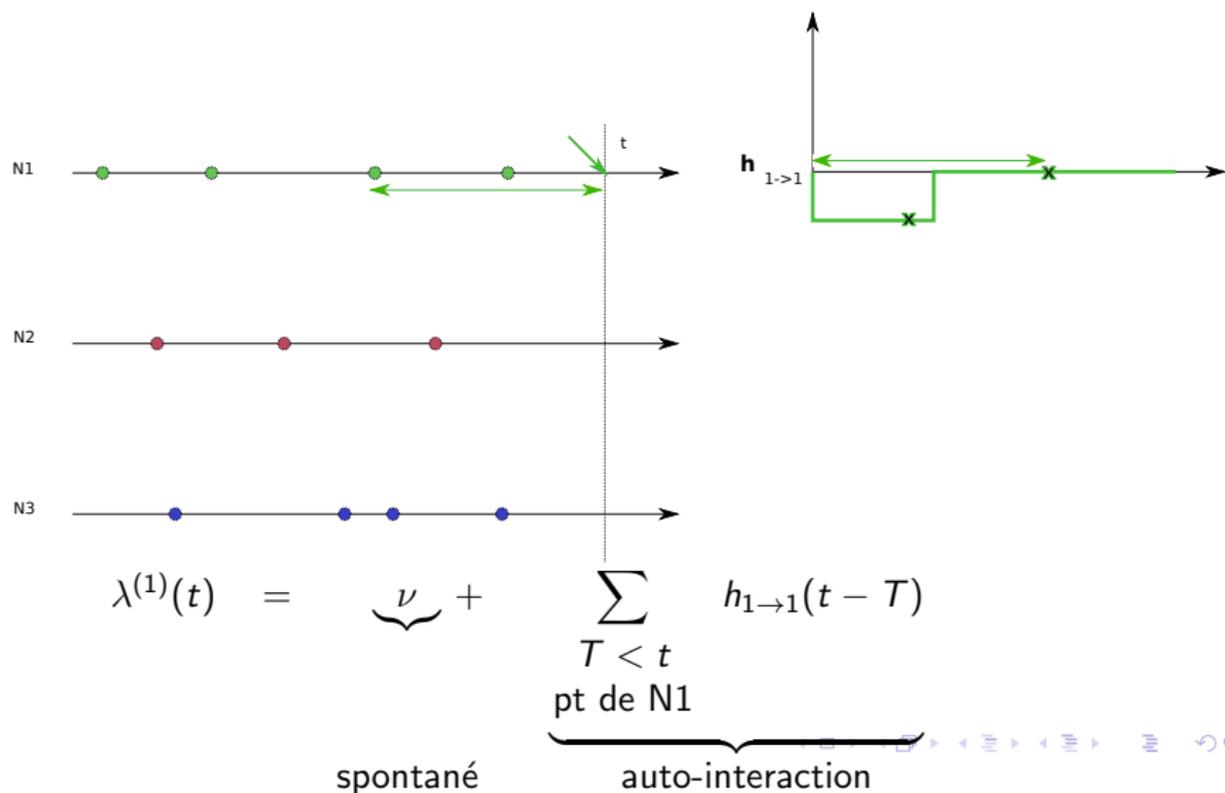
# Processus de Hawkes multivarié



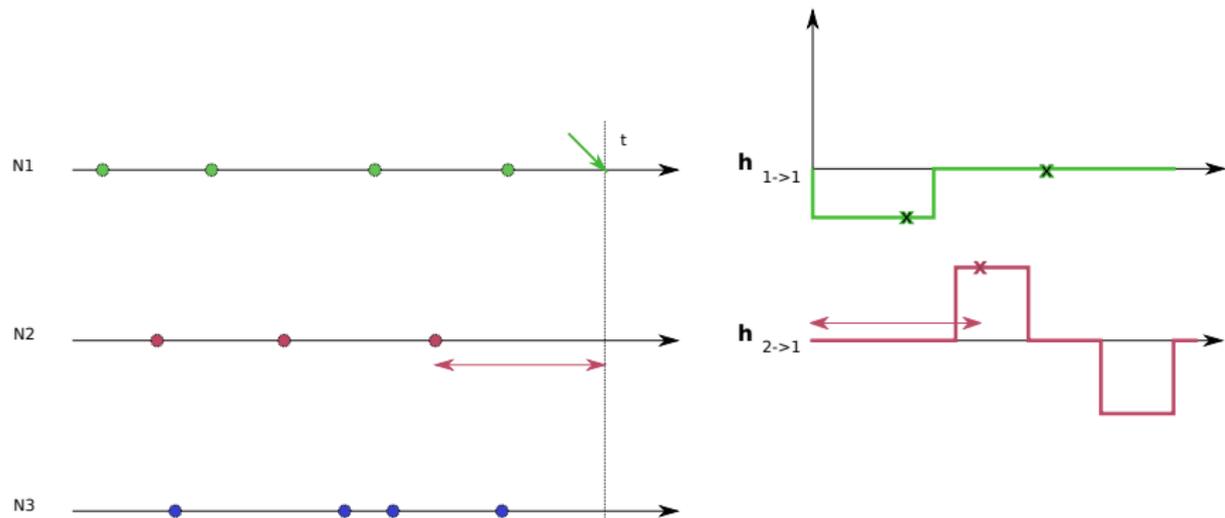
# Processus de Hawkes multivarié



# Processus de Hawkes multivarié



## Processus de Hawkes multivarié



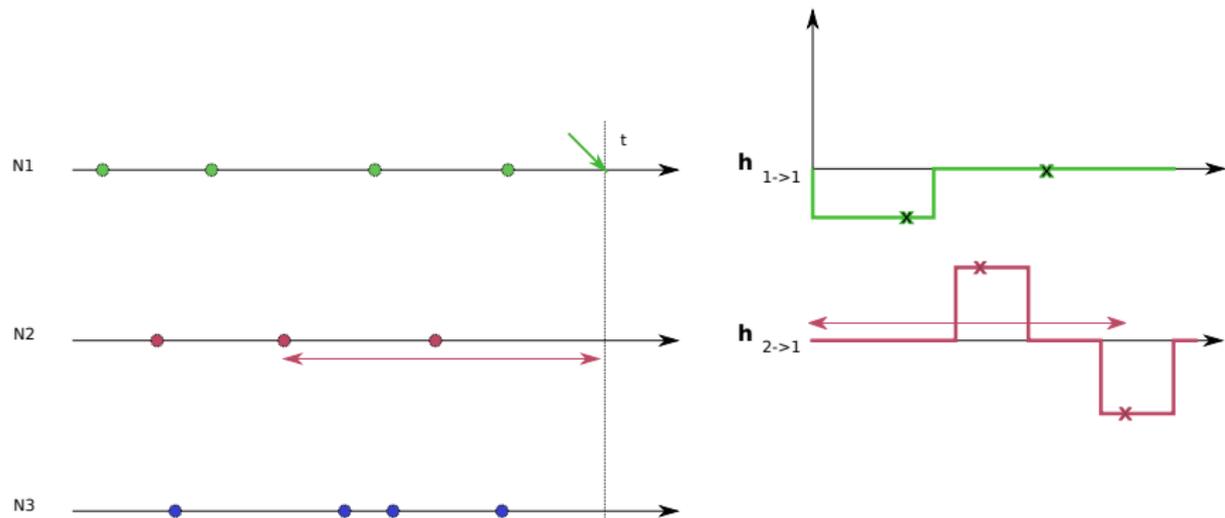
$$\lambda^{(1)}(t) = \underbrace{\nu}_{\text{spontané}} + \underbrace{\sum_{\substack{T < t \\ \text{pt de N1}}} h_{1 \rightarrow 1}(t - T)}_{\text{auto-interaction}} + \underbrace{\sum_{\substack{m \neq 1, \\ T < t, \\ \text{pt de Nm}}} h_{m \rightarrow 1}(t - T)}_{\text{interaction}}$$

spontané

auto-interaction

interaction

## Processus de Hawkes multivarié



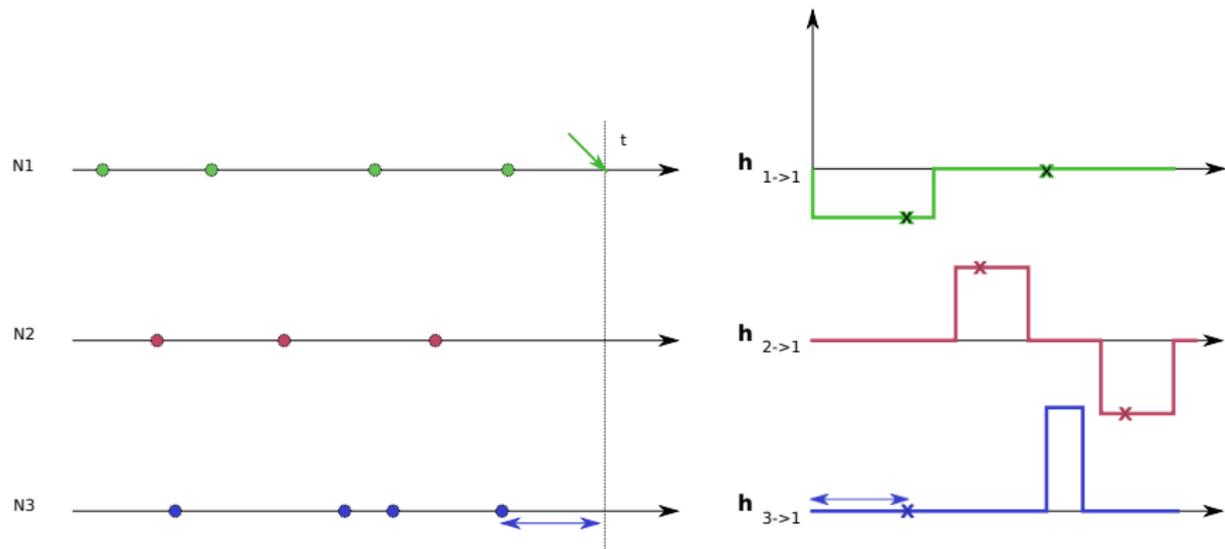
$$\lambda^{(1)}(t) = \underbrace{\nu}_{\text{spontané}} + \underbrace{\sum_{\substack{T < t \\ \text{pt de N1}}} h_{1 \rightarrow 1}(t - T)}_{\text{auto-interaction}} + \underbrace{\sum_{\substack{m \neq 1, \\ T < t, \\ \text{pt de Nm}}} h_{m \rightarrow 1}(t - T)}_{\text{interaction}}$$

spontané

auto-interaction

interaction

## Processus de Hawkes multivarié



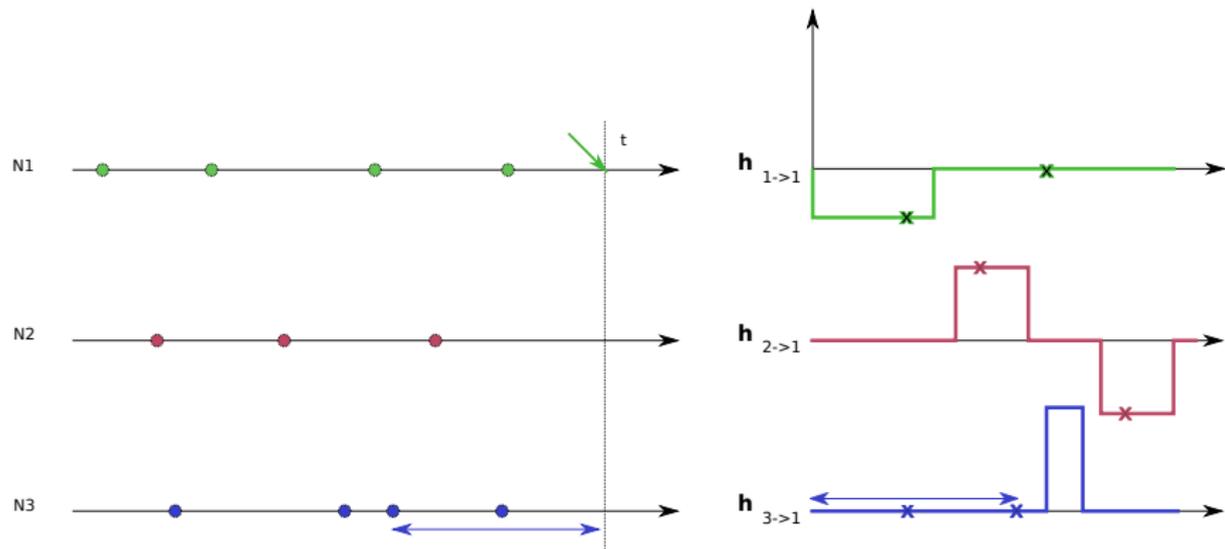
$$\lambda^{(1)}(t) = \underbrace{\nu}_{\text{spontané}} + \underbrace{\sum_{\substack{T < t \\ \text{pt de N1}}} h_{1 \rightarrow 1}(t - T)}_{\text{auto-interaction}} + \underbrace{\sum_{\substack{m \neq 1, \\ T < t, \\ \text{pt de Nm}}} h_{m \rightarrow 1}(t - T)}_{\text{interaction}}$$

spontané

auto-interaction

interaction

## Processus de Hawkes multivarié



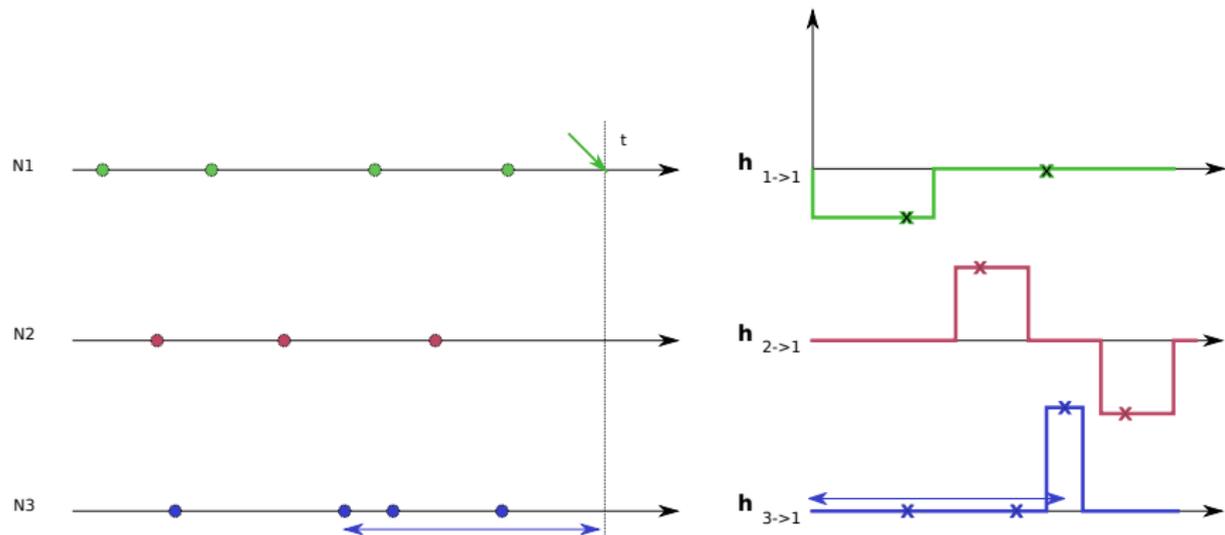
$$\lambda^{(1)}(t) = \underbrace{\nu}_{\text{spontané}} + \underbrace{\sum_{\substack{T < t \\ \text{pt de N1}}} h_{1 \rightarrow 1}(t - T)}_{\text{auto-interaction}} + \underbrace{\sum_{\substack{m \neq 1, \\ T < t, \\ \text{pt de Nm}}} h_{m \rightarrow 1}(t - T)}_{\text{interaction}}$$

spontané

auto-interaction

interaction

## Processus de Hawkes multivarié



$$\lambda^{(1)}(t) = \underbrace{\nu}_{\text{spontané}} + \underbrace{\sum_{\substack{T < t \\ \text{pt de N1}}} h_{1 \rightarrow 1}(t - T)}_{\text{auto-interaction}} + \underbrace{\sum_{\substack{m \neq 1, \\ T < t, \\ \text{pt de Nm}}} h_{m \rightarrow 1}(t - T)}_{\text{interaction}}$$

spontané

auto-interaction

interaction

# Test d'adéquation à un processus de Hawkes

- 1 Avec toutes les observations :  $\hat{\lambda}$  estimateur de  $\lambda$

## Test d'adéquation à un processus de Hawkes

- 1 Avec toutes les observations :  $\hat{\lambda}$  estimateur de  $\lambda$
- 2 On considère un  $m$ -échantillon de  $\{1, \dots, n\}$  avec  $m = n^{(2/3)}$

## Test d'adéquation à un processus de Hawkes

- 1 Avec toutes les observations :  $\hat{\lambda}$  estimateur de  $\lambda$
- 2 On considère un  $m$ -échantillon de  $\{1, \dots, n\}$  avec  $m = n^{(2/3)}$
- 3 On pose  $\hat{\Lambda}_i(t) = \int_0^t (\hat{\lambda}_i(u))_+ du$ , et  
 $\hat{\mathcal{N}}_i = \{X = \hat{\Lambda}_i(T) : T \in N_i\}$  (processus de Poisson sur  $[0, \hat{\Lambda}_i(T_{max})]$ )



## Test d'adéquation à un processus de Hawkes

- 1 Avec toutes les observations :  $\hat{\lambda}$  estimateur de  $\lambda$
- 2 On considère un  $m$ -échantillon de  $\{1, \dots, n\}$  avec  $m = n^{(2/3)}$
- 3 On pose  $\hat{\Lambda}_i(t) = \int_0^t (\hat{\lambda}_i(u))_+ du$ , et  
 $\hat{\mathcal{N}}_i = \{X = \hat{\Lambda}_i(T) : T \in N_i\}$  (processus de Poisson sur  $[0, \hat{\Lambda}_i(T_{max})]$ )
- 4 On cumule les  $m$  processus pour former  $\mathcal{N}^{c,m}$
- 5 Soit  $\tilde{k}_{1-\alpha}$  le  $1 - \alpha$  quantile de la distribution asymptotique  $\mathcal{K}$

## Test d'adéquation à un processus de Hawkes

- ① Avec toutes les observations :  $\hat{\lambda}$  estimateur de  $\lambda$
- ② On considère un  $m$ -échantillon de  $\{1, \dots, n\}$  avec  $m = n^{(2/3)}$
- ③ On pose  $\hat{\Lambda}_i(t) = \int_0^t (\hat{\lambda}_i(u))_+ du$ , et  
 $\hat{\mathcal{N}}_i = \{X = \hat{\Lambda}_i(T) : T \in N_i\}$  (processus de Poisson sur  $[0, \hat{\Lambda}_i(T_{max})]$ )
- ④ On cumule les  $m$  processus pour former  $\mathcal{N}^{c,m}$
- ⑤ Soit  $\tilde{k}_{1-\alpha}$  le  $1 - \alpha$  quantile de la distribution asymptotique  $\mathcal{K}$
- ⑥ Soit  $\theta > 0$ , strictement plus petit que  $m^{-1} \sum_i \hat{\Lambda}_i(T_{max})$

## Test d'adéquation à un processus de Hawkes

- 1 Avec toutes les observations :  $\hat{\lambda}$  estimateur de  $\lambda$
- 2 On considère un  $m$ -échantillon de  $\{1, \dots, n\}$  avec  $m = n^{(2/3)}$
- 3 On pose  $\hat{\Lambda}_i(t) = \int_0^t (\hat{\lambda}_i(u))_+ du$ , et  
 $\hat{\mathcal{N}}_i = \{X = \hat{\Lambda}_i(T) : T \in N_i\}$  (processus de Poisson sur  $[0, \hat{\Lambda}_i(T_{max})]$ )
- 4 On cumule les  $m$  processus pour former  $\mathcal{N}^{c,m}$
- 5 Soit  $\tilde{k}_{1-\alpha}$  le  $1 - \alpha$  quantile de la distribution asymptotique  $\mathcal{K}$
- 6 Soit  $\theta > 0$ , strictement plus petit que  $m^{-1} \sum_i \hat{\Lambda}_i(T_{max})$
- 7 on rejette  $\mathcal{H}_0$ : "Les  $n$  processus sont des processus admettant une intensité conditionnelle  $\lambda$ " si

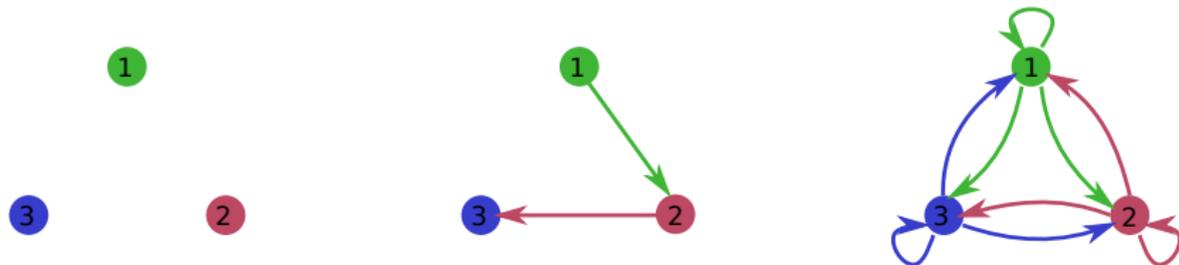
$$\sqrt{\hat{\mathcal{N}}^{c,m}([0, m\theta])} \sup_{u \in [0,1]} \left| \frac{1}{\hat{\mathcal{N}}^{c,m}([0, m\theta])} \sum_{X \in \hat{\mathcal{N}}^{c,m}, X \leq m\theta} \mathbf{1}_{\{X/(m\theta) \leq u\}} - u \right| > \tilde{k}_{1-\alpha}$$

# Application

Directions		1	2	3	4	5	6
N1A	B			○	○	○	
N1A	BO					▲▲▲	○
N2A	B	○	▲▲		▲		○
N2A	BO	○					○
N1B	B		△	▲		○	
N1B	BO						
N2B	B	▲▲▲	▲▲▲	○		▲	
N2B	BO		○	▲	▲		

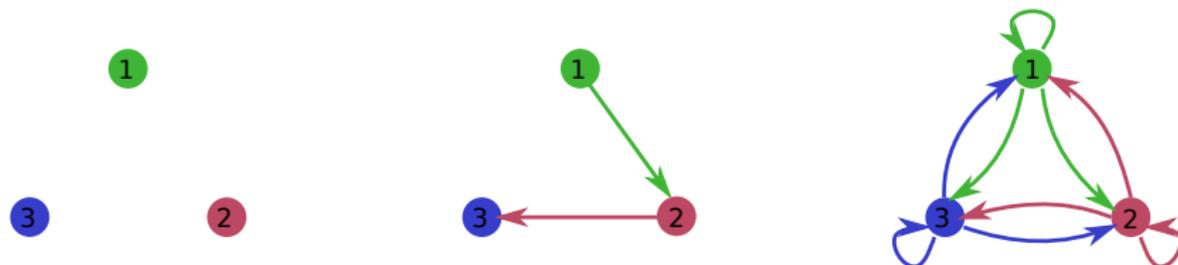
# Graphes d'indépendance locale

(Didelez (2008))



# Graphes d'indépendance locale

(Didelez (2008))



Si nous sommes capables d'accéder aux graphes d'indépendance locale, alors nous aurons accès à la "connectivité fonctionnelle".  
↔ un "modèle" .

## moindre carrés

- une bonne estimation des paramètres doit correspondre à une intensité candidate  $\eta$  proche de la vraie.

## moindre carrés

- une bonne estimation des paramètres doit correspondre à une intensité candidate  $\eta$  proche de la vraie.
- on souhaite minimiser  $\|\eta - \lambda\|^2 = \int_0^T [\eta(t) - \lambda(t)]^2 dt$  en  $\eta$

## moindre carrés

- une bonne estimation des paramètres doit correspondre à une intensité candidate  $\eta$  proche de la vraie.
- on souhaite minimiser  $\|\eta - \lambda\|^2 = \int_0^T [\eta(t) - \lambda(t)]^2 dt$  en  $\eta$
- ou de façon équivalente  $-2 \int_0^T \eta(t)\lambda(t)dt + \int_0^T \eta(t)^2 dt$ .

## moindre carrés

- une bonne estimation des paramètres doit correspondre à une intensité candidate  $\eta$  proche de la vraie.
- on souhaite minimiser  $\|\eta - \lambda\|^2 = \int_0^T [\eta(t) - \lambda(t)]^2 dt$  en  $\eta$
- ou de façon équivalente  $-2 \int_0^T \eta(t)\lambda(t)dt + \int_0^T \eta(t)^2 dt$ .
- pb :  $\lambda(t)dt$  n'est pas observable mais  $dN_t$  fluctue aléatoirement autour de  $\lambda(t)dt$  et est lui observable.

## moindre carrés

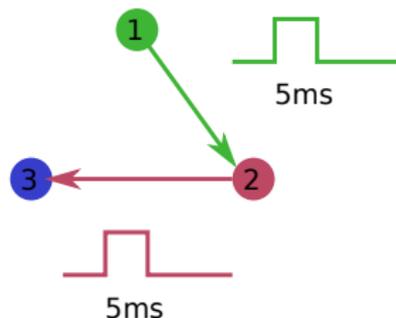
- une bonne estimation des paramètres doit correspondre à une intensité candidate  $\eta$  proche de la vraie.
- on souhaite minimiser  $\|\eta - \lambda\|^2 = \int_0^T [\eta(t) - \lambda(t)]^2 dt$  en  $\eta$
- ou de façon équivalente  $-2 \int_0^T \eta(t)\lambda(t)dt + \int_0^T \eta(t)^2 dt$ .
- pb :  $\lambda(t)dt$  n'est pas observable mais  $dN_t$  fluctue aléatoirement autour de  $\lambda(t)dt$  et est lui observable.

•

$$\Rightarrow -2 \int_0^T \eta(t) dN_t + \int_0^T \eta(t)^2 dt,$$

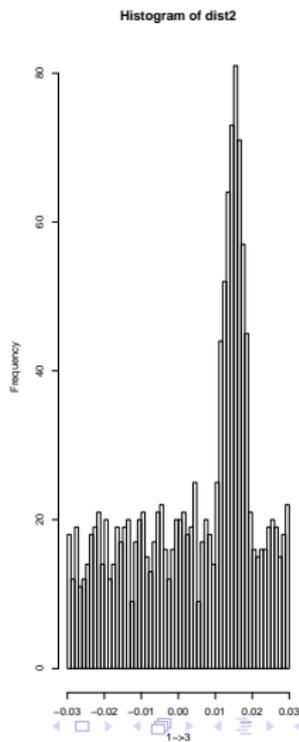
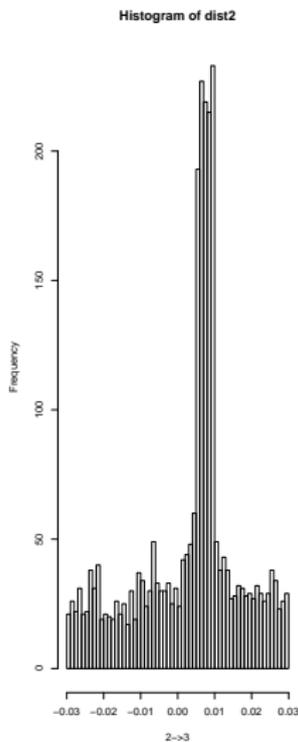
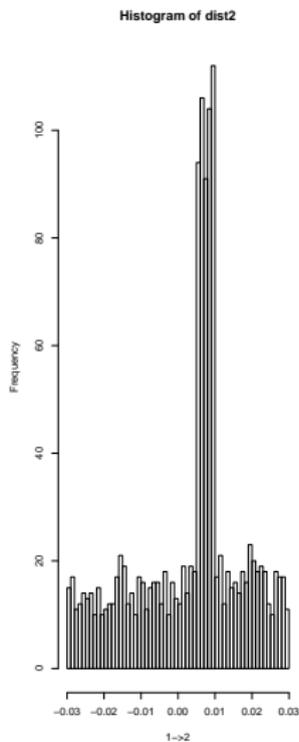
pour  $\eta = \lambda_a(t)$

# Comment déterminer le graphe de dépendance locale?

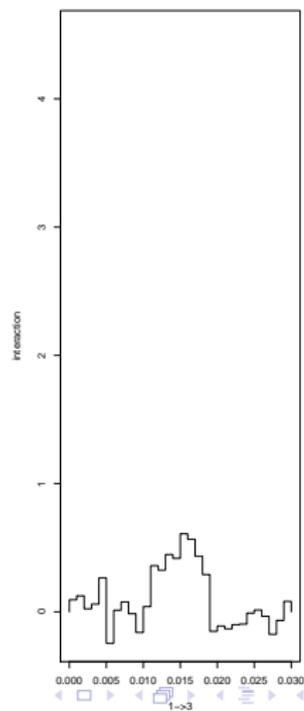
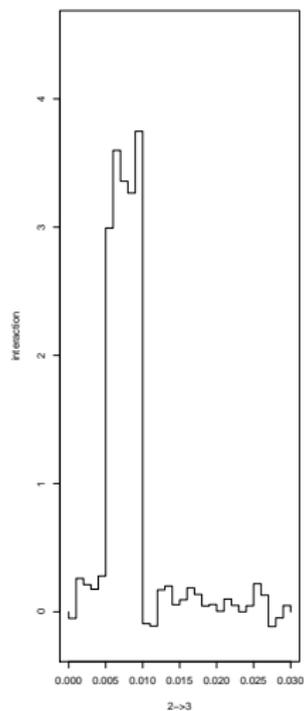
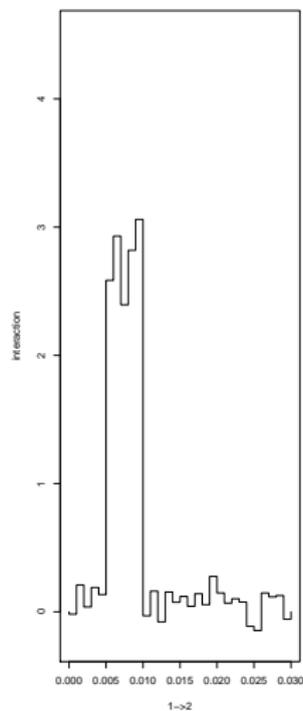


deux fonctions d'interaction non nulles sur les 9 possibles

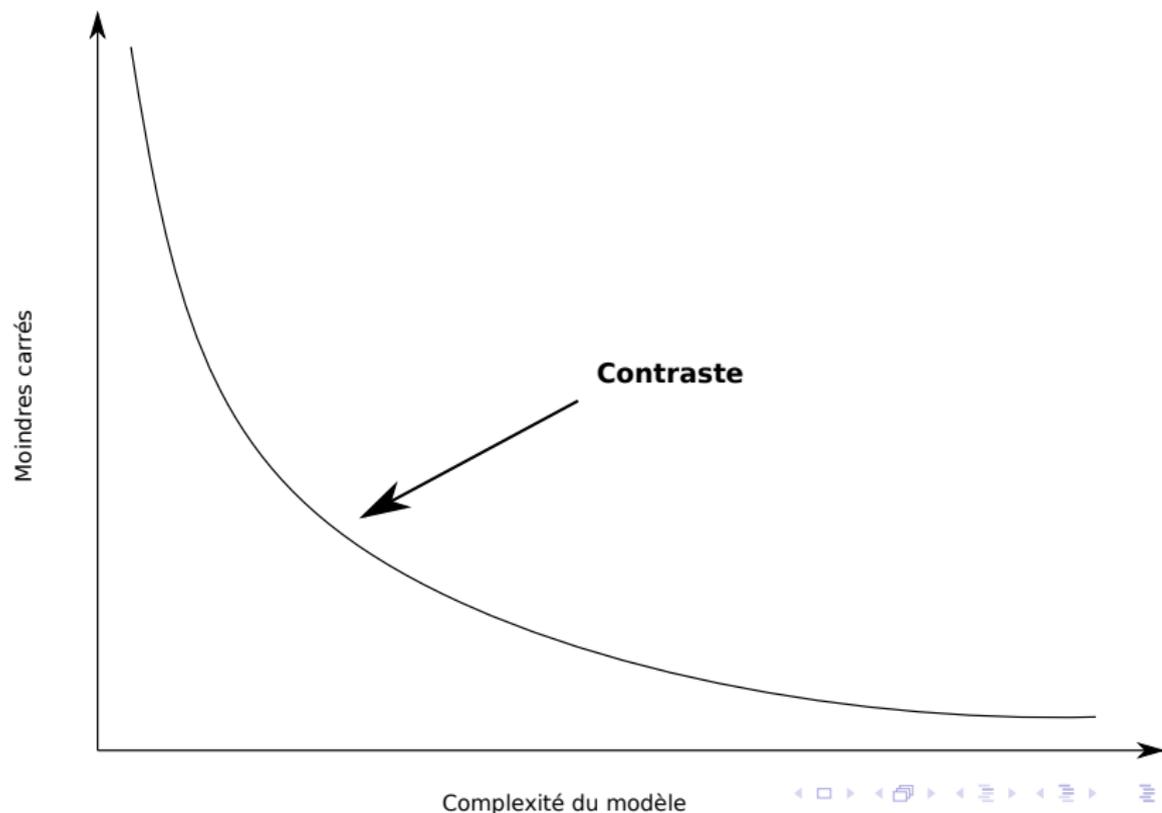
# Comment déterminer le graphe de dépendance locale?



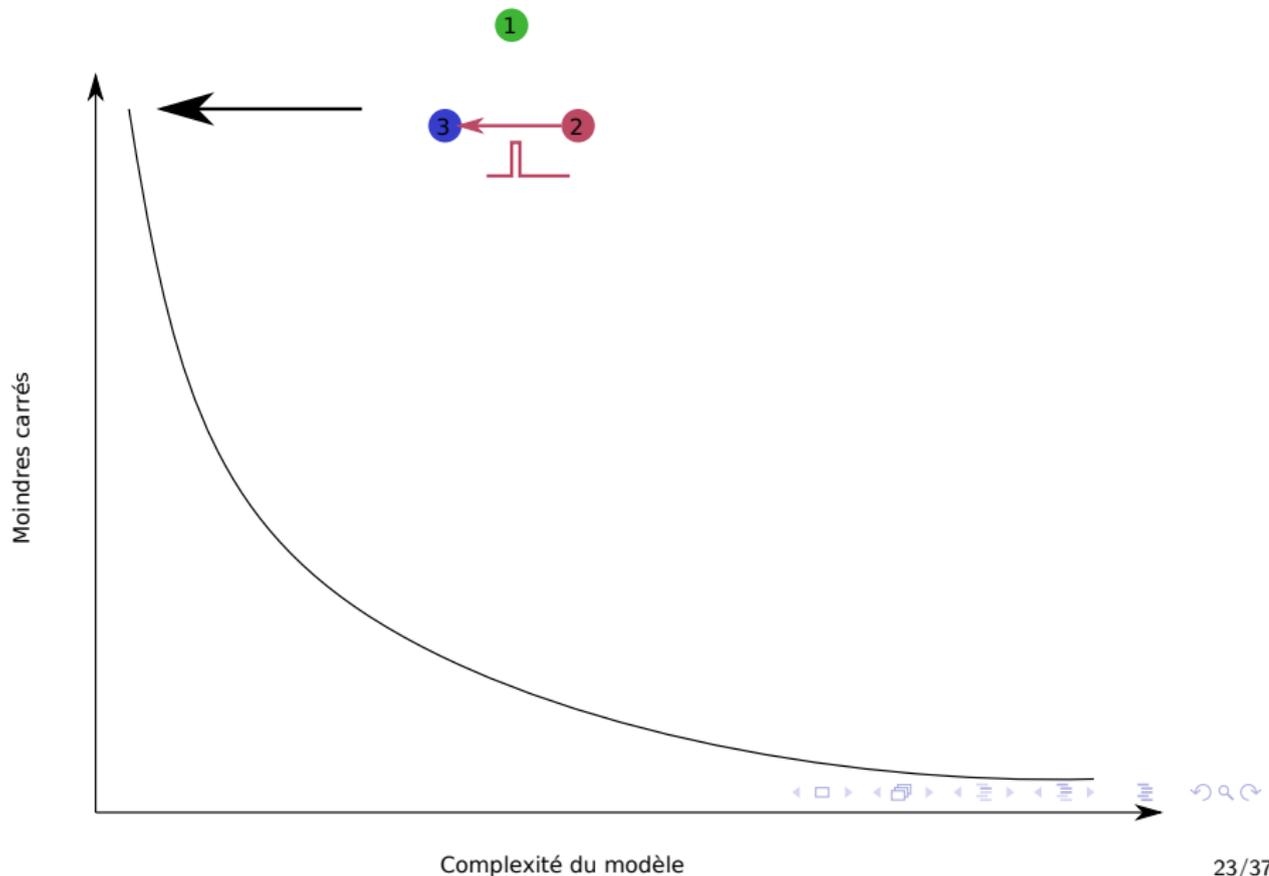
# Comment déterminer le graphe de dépendance locale?



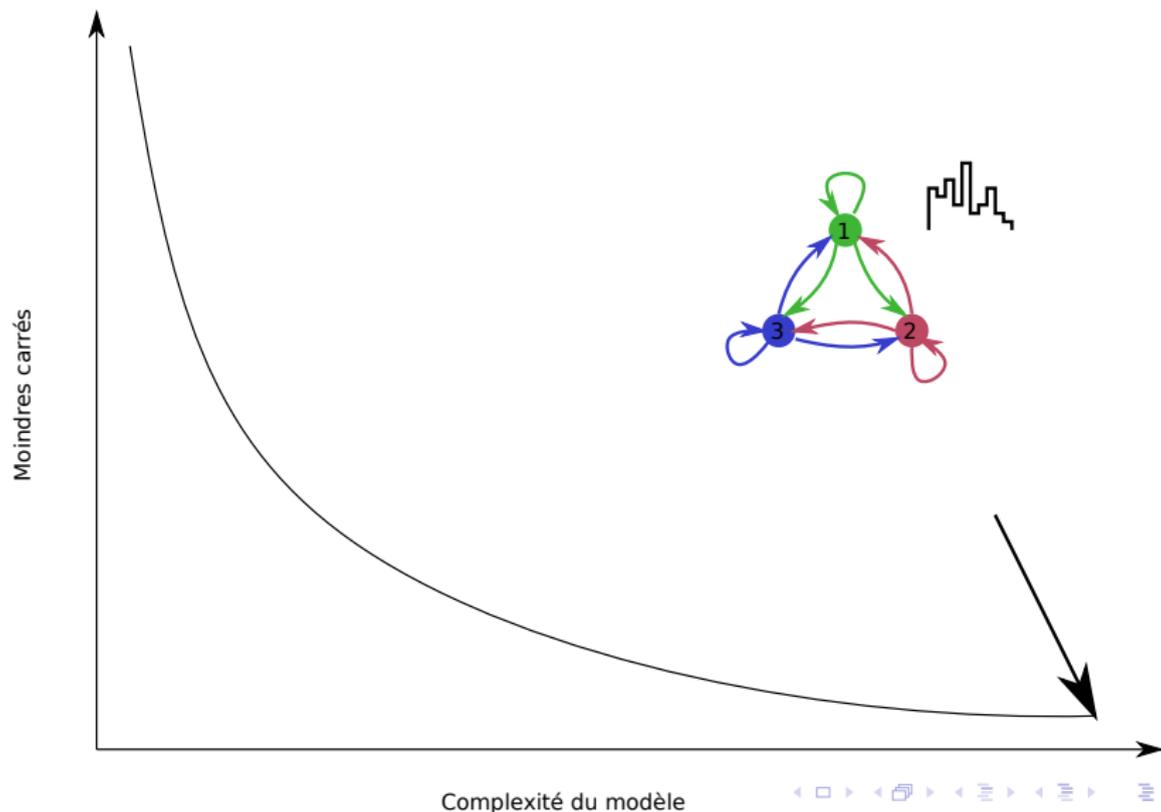
# Pénalisation et choix des paramètres



# Pénalisation et choix des paramètres

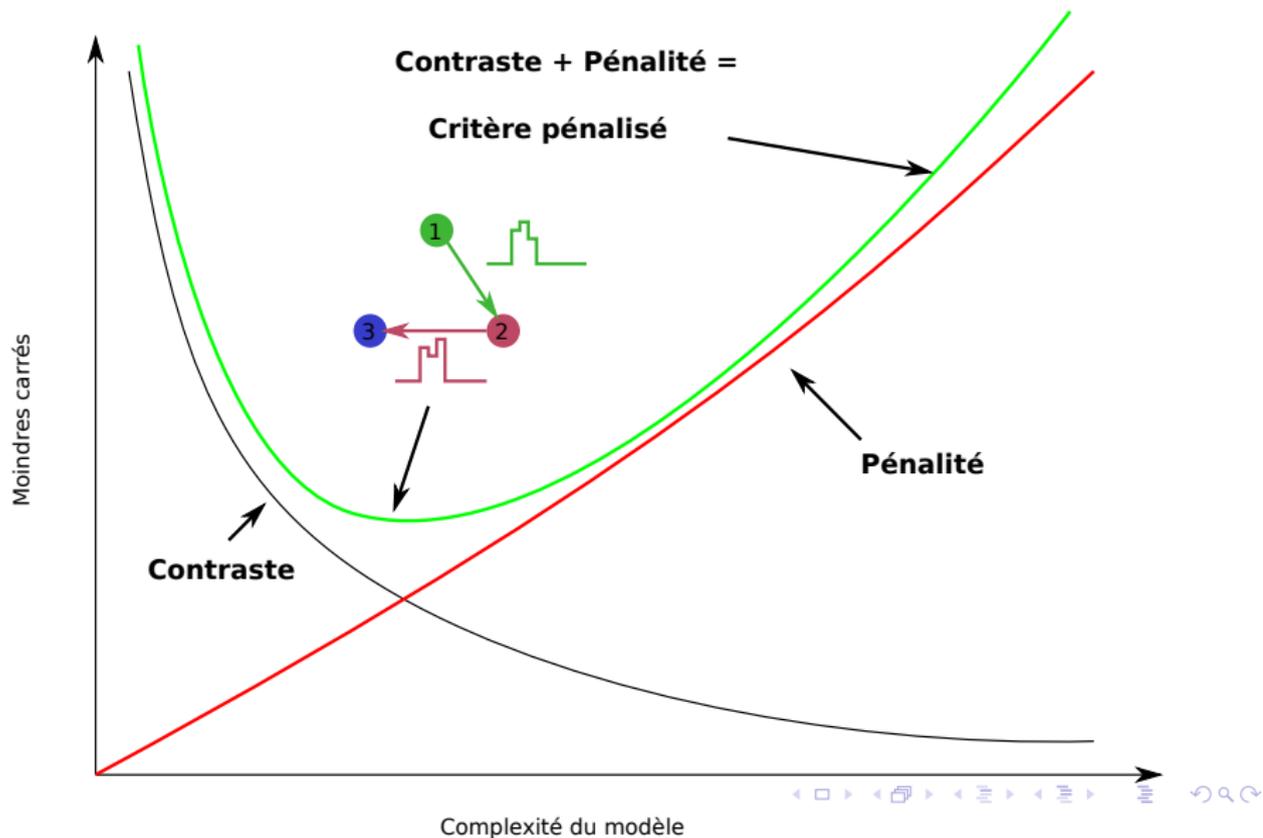


# Pénalisation et choix des paramètres





# Pénalisation et choix des paramètres



## Quelques détails mathématiques

L'intensité du processus  $N^{(r)}$  est donné par

$$\lambda^{(r)}(t) = \nu^{(r)} + \sum_{\ell=1}^M \sum_{T < t, T \text{ in } N^{(\ell)}} h_{\ell}^{(r)}(t - T).$$

## Quelques détails mathématiques

L'intensité du processus  $N^{(r)}$  est donné par

$$\lambda^{(r)}(t) = \nu^{(r)} + \sum_{\ell=1}^M \sum_{T < t, T \text{ in } N^{(\ell)}} h_{\ell}^{(r)}(t - T).$$

On considère que les fonctions d'interaction sont des constantes par morceaux sur une grille fixe que l'on se donne  $\Rightarrow$  vecteur  $a_*^{(r)}$ .

## Quelques détails mathématiques

L'intensité du processus  $N^{(r)}$  est donné par

$$\lambda^{(r)}(t) = \nu^{(r)} + \sum_{\ell=1}^M \sum_{T < t, T \text{ in } N^{(\ell)}} h_{\ell}^{(r)}(t - T).$$

On considère que les fonctions d'interaction sont des constantes par morceaux sur une grille fixe que l'on se donne  $\Rightarrow$  vecteur  $\mathbf{a}_*^{(r)}$ .  
par linéarité, on a alors :

$$\lambda^{(r)}(t) = (\mathbf{Rc}_t)' \mathbf{a}_*^{(r)},$$

avec  $\mathbf{Rc}_t$  défini par

$$(\mathbf{Rc}_t)' = \left( 1, \delta^{-1/2}(\mathbf{c}_t^{(1)})', \dots, \delta^{-1/2}(\mathbf{c}_t^{(M)})' \right),$$

et avec  $\mathbf{c}_t^{(\ell)}$  défini par

$$(\mathbf{c}_t^{(\ell)})' = \left( N_{[t-\delta, t]}^{(\ell)}, \dots, N_{[t-K\delta, t-(K-1)\delta]}^{(\ell)} \right).$$

## Heuristique sur l'estimateur des moindres carrés

On peut écrire que

$$dN^{(r)}(t) \simeq (\mathbf{Rc}_t)' \mathbf{a}_*^{(r)} dt + \text{bruit.}$$

## Heuristique sur l'estimateur des moindres carrés

On peut écrire que

$$dN^{(r)}(t) \simeq (\mathbf{Rc}_t)' \mathbf{a}_*^{(r)} dt + \text{bruit.}$$

Posons

$$\mathbf{G} = \int_0^T \mathbf{Rc}_t (\mathbf{Rc}_t)' dt,$$

la covariance intégrée du nombre de points instantanés renormalisé.

## Heuristique sur l'estimateur des moindres carrés

On peut écrire que

$$dN^{(r)}(t) \simeq (\mathbf{Rc}_t)' \mathbf{a}_*^{(r)} dt + \text{bruit.}$$

Posons

$$\mathbf{G} = \int_0^T \mathbf{Rc}_t (\mathbf{Rc}_t)' dt,$$

la covariance intégrée du nombre de points instantanés renormalisé.

$$\mathbf{b}^{(r)} := \int_0^T \mathbf{Rc}_t dN^{(r)}(t) \simeq \mathbf{G} \mathbf{a}_*^{(r)} + \text{noise.}$$

## Heuristique sur l'estimateur des moindres carrés

On peut écrire que

$$dN^{(r)}(t) \simeq (\mathbf{Rc}_t)' \mathbf{a}_*^{(r)} dt + \text{bruit.}$$

Posons

$$\mathbf{G} = \int_0^T \mathbf{Rc}_t (\mathbf{Rc}_t)' dt,$$

la covariance intégrée du nombre de points instantanés renormalisé.

$$\mathbf{b}^{(r)} := \int_0^T \mathbf{Rc}_t dN^{(r)}(t) \simeq \mathbf{G} \mathbf{a}_*^{(r)} + \text{noise.}$$

L'estimateur des moindres carrés est :

$$\hat{\mathbf{a}}^{(r)} = \mathbf{G}^{-1} \mathbf{b}^{(r)},$$

où dans  $\mathbf{b}$  se trouve une nouvelle fois le nombre de couples avec un certain délai.

## Heuristique sur l'estimateur des moindres carrés

On peut écrire que

$$dN^{(r)}(t) \simeq (\mathbf{Rc}_t)' \mathbf{a}_*^{(r)} dt + \text{bruit.}$$

Posons

$$\mathbf{G} = \int_0^T \mathbf{Rc}_t (\mathbf{Rc}_t)' dt,$$

la covariance intégrée du nombre de points instantanés renormalisé.

$$\mathbf{b}^{(r)} := \int_0^T \mathbf{Rc}_t dN^{(r)}(t) \simeq \mathbf{G} \mathbf{a}_*^{(r)} + \text{noise.}$$

L'estimateur des moindres carrés est :

$$\hat{\mathbf{a}}^{(r)} = \mathbf{G}^{-1} \mathbf{b}^{(r)},$$

où dans  $\mathbf{b}$  se trouve une nouvelle fois le nombre de couples avec un certain délai.

# Estimation Lasso (Hansen, Reynaud-Bouret, Rivoirard)

## Criterion Lasso

$$\hat{\mathbf{a}} = \operatorname{argmin}_{\mathbf{a} \in \mathbb{R}^{\Phi}} \{-2\mathbf{b}^* \mathbf{a} + \mathbf{a}^* \mathbf{G} \mathbf{a} + 2\mathbf{d}^* |\mathbf{a}|\}$$

## Hansen, Rivoirard, RB

Si  $\mathbf{G} \geq cI$  avec  $c > 0$  et si

$$\left| \int_0^T \mathbf{R} \mathbf{c}_t \left( dN^{(r)}(t) - \lambda^{(r)}(t) dt \right) \right| \leq \mathbf{d}^{(r)}, \quad \forall r$$

alors

$$\sum_r \|\lambda^{(r)} - \mathbf{R} \mathbf{c}_t \hat{\mathbf{a}}^{(r)}\|^2 \leq \square \inf_{\mathbf{a}} \left\{ \sum_r \|\lambda^{(r)} - \mathbf{R} \mathbf{c}_t \mathbf{a}\|^2 + \frac{1}{c} \sum_{i \in \operatorname{supp}(\mathbf{a})} (d_i^{(r)})^2 \right\}.$$

# Explications

- **Point principal:**  $d$  contrôle les fluctuations aléatoires / bruit  
→ guidé par les données et fin. Il est dépendant d'un paramètre  $\gamma$  qui lorsque  $\gamma = 0$  on retrouve les moindres carrés ordinaires et lorsque  $\gamma$  croît, le modèle devient de plus en plus parcimonieux !

# Explications

- **Point principal:**  $\mathbf{d}$  contrôle les fluctuations aléatoires / bruit  
→ guidé par les données et fin. Il est dépendant d'un paramètre  $\gamma$  qui lorsque  $\gamma = 0$  on retrouve les moindres carrés ordinaires et lorsque  $\gamma$  croît, le modèle devient de plus en plus parcimonieux !
- Le terme de reste  $(1/T) \sum_{i \in \text{supp}(\mathbf{a})} (d_i^{(r)})^2$  est inévitable. Il doit être vu comme “la capacité d'approximation” + le reste du au bruit.

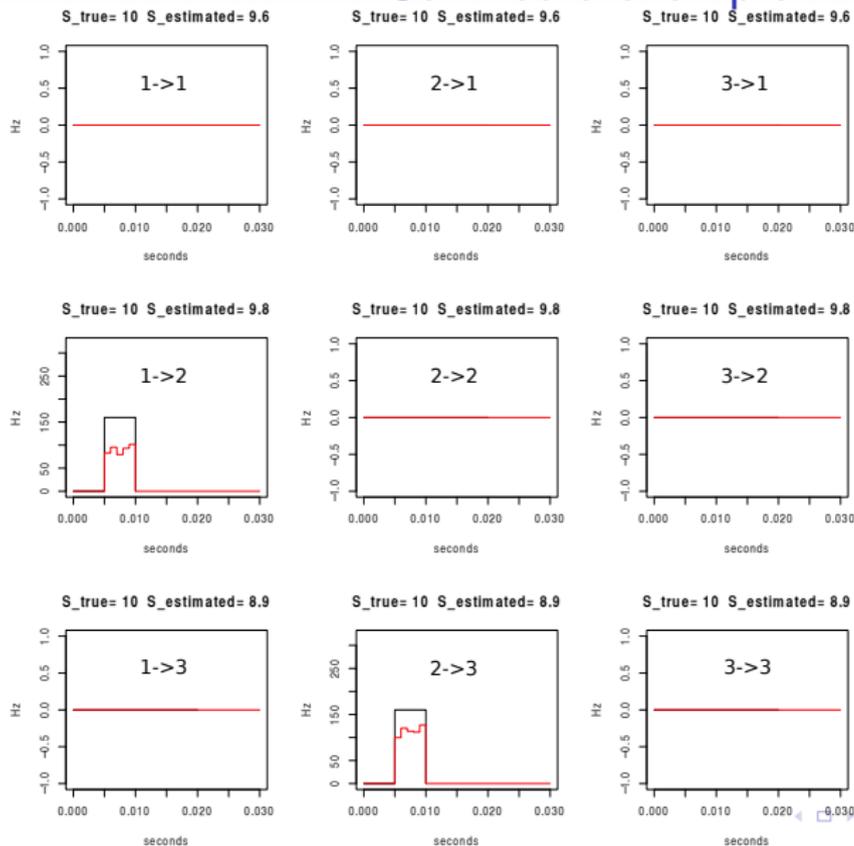
# Explications

- **Point principal:**  $\mathbf{d}$  contrôle les fluctuations aléatoires / bruit  
→ guidé par les données et fin. Il est dépendant d'un paramètre  $\gamma$  qui lorsque  $\gamma = 0$  on retrouve les moindres carrés ordinaires et lorsque  $\gamma$  croît, le modèle devient de plus en plus parcimonieux !
- Le terme de reste  $(1/T) \sum_{i \in \text{supp}(\mathbf{a})} (d_i^{(r)})^2$  est inévitable. Il doit être vu comme “la capacité d'approximation” + le reste du au bruit.
- Le facteur  $1/c$  est un “indicateur de qualité”

## Explications

- **Point principal:**  $\mathbf{d}$  contrôle les fluctuations aléatoires / bruit  
→ guidé par les données et fin. Il est dépendant d'un paramètre  $\gamma$  qui lorsque  $\gamma = 0$  on retrouve les moindres carrés ordinaires et lorsque  $\gamma$  croît, le modèle devient de plus en plus parcimonieux !
- Le terme de reste  $(1/T) \sum_{i \in \text{supp}(\mathbf{a})} (d_i^{(r)})^2$  est inévitable. Il doit être vu comme “la capacité d'approximation” + le reste du au bruit.
- Le facteur  $1/c$  est un “indicateur de qualité”
- Par cette méthode, on obtient la meilleure approximation par Hawkes de l'intensité du processus observé

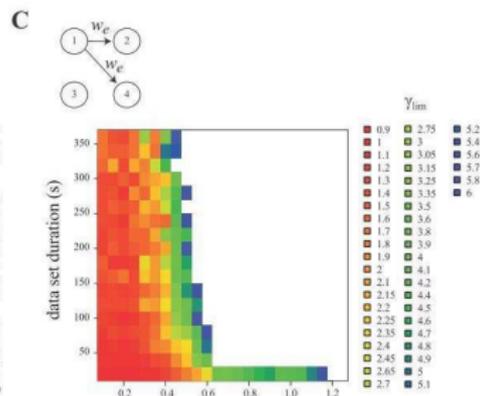
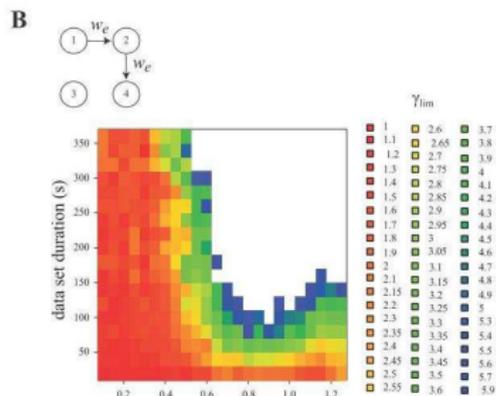
# Sur notre exemple



## Un réseau neuronal plus réaliste

⇒ Tester les limites et la précision de cette méthode quant à la modélisation des tains de spike

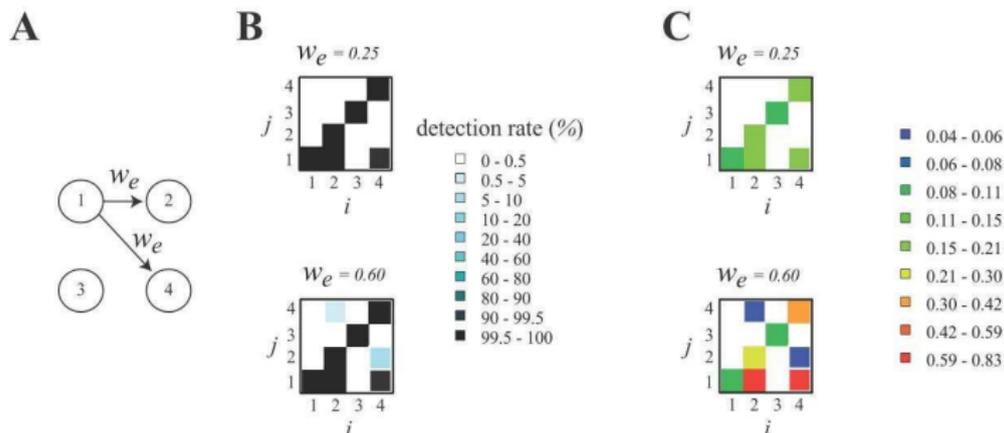
⇒ Simulation d'un modèle Integrate and Fire à  $n$  neurones ( $n < 10$ )



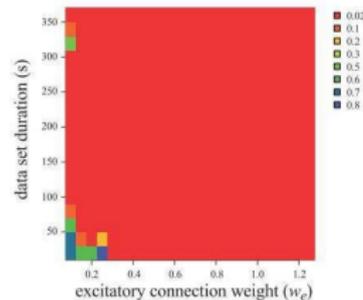
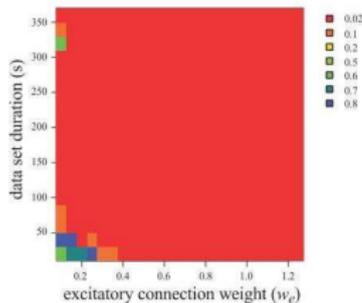
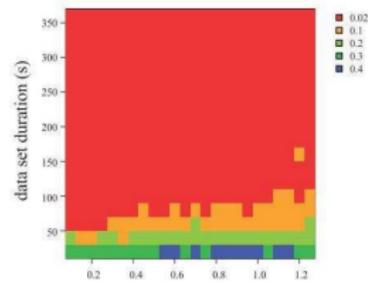
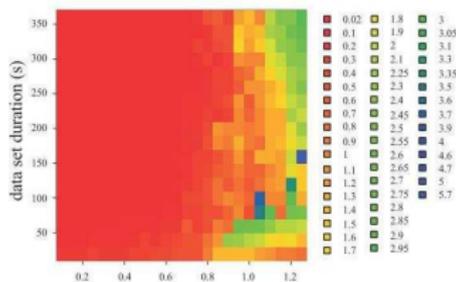
## Un réseau neuronal plus réaliste (2)

⇒ introduction d'un quantificateur de la force d'une interaction :

- force d'une interaction :  $\int h_{j \rightarrow i}(x) dx$
- énergie d'une interaction :  $\int |h_{j \rightarrow i}(x)| dx$



# Un réseau neuronal plus réaliste (3)

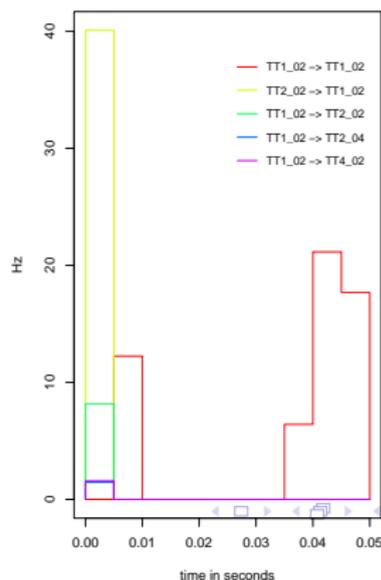
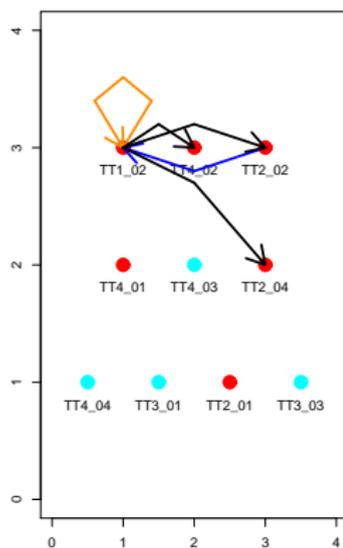


# Sur des données neuronales liées à l'excitation des vibris

Excitation des vibris à faible fréquence.  $T = 90.5$ ,  $M = 10$

Neuron	$TT_{1_02}$	$TT_{2_01}$	$TT_{2_02}$	$TT_{2_04}$	$TT_{3_01}$	$TT_{3_03}$	$TT_{4_01}$	$TT_{4_02}$	$TT_{4_03}$	$TT_{4_04}$
Spikes	9191	99	544	149	15	18	136	282	8	6

Comportement 1 ; k 10 ; delta 0.005 ; gamma 1



## Sur des données neuronales liées à l'excitation des vibris (2)

Data:

Neuron	$TT_{102}$	$TT_{201}$	$TT_{202}$	$TT_{204}$	$TT_{301}$	$TT_{303}$	$TT_{401}$	$TT_{402}$	$TT_{403}$	$TT_{404}$
Spikes	9191	99	544	149	15	18	136	282	8	6

Simulation:

Neuron	$TT_{102}$	$TT_{201}$	$TT_{202}$	$TT_{204}$	$TT_{301}$	$TT_{303}$	$TT_{401}$	$TT_{402}$	$TT_{403}$	$TT_{404}$
Spikes	9327	92	585	148	13	23	133	271	8	3

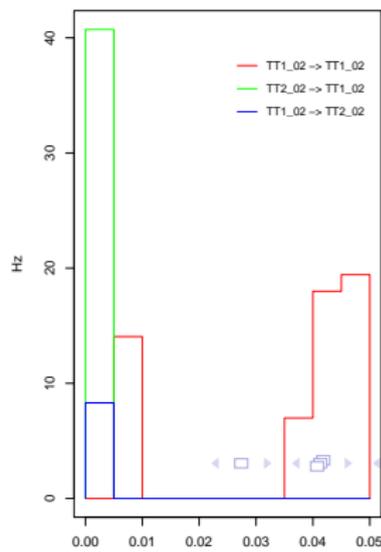
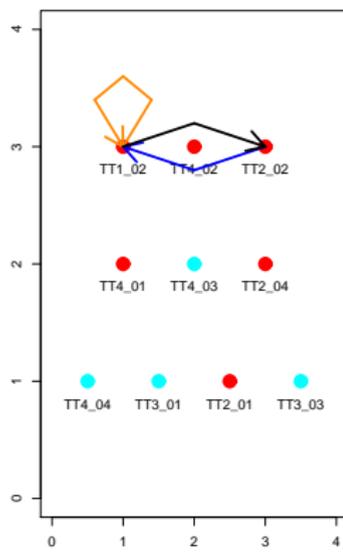
# Sur des données neuronales liées à l'excitation des vibris (2)

Data:

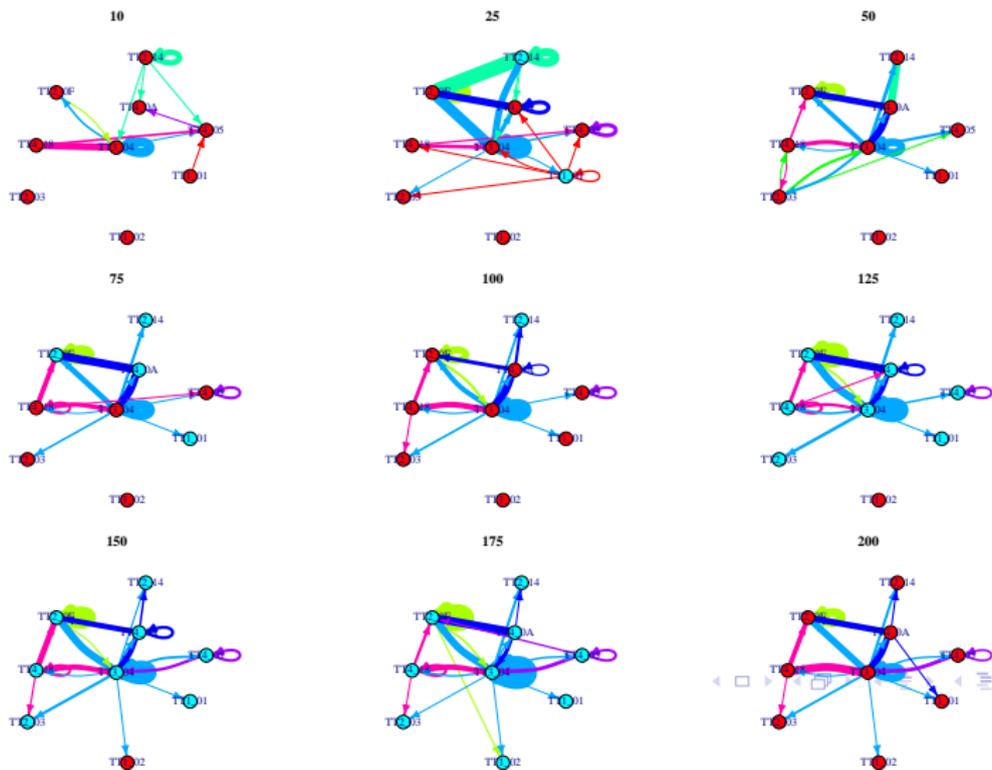
Neuron	$TT1_{02}$	$TT2_{01}$	$TT2_{02}$	$TT2_{04}$	$TT3_{01}$	$TT3_{03}$	$TT4_{01}$	$TT4_{02}$	$TT4_{03}$	$TT4_{04}$
Spikes	9191	99	544	149	15	18	136	282	8	6

Simulation:

Neuron	$TT1_{02}$	$TT2_{01}$	$TT2_{02}$	$TT2_{04}$	$TT3_{01}$	$TT3_{03}$	$TT4_{01}$	$TT4_{02}$	$TT4_{03}$	$TT4_{04}$
Spikes	9327	92	585	148	13	23	133	271	8	3



# Evolution du graphe de dépendance en fonction de l'excitation de la vibris



# Conclusions

## conclusions

- Tests d'indépendance : seulement pour deux neurones  
⇒ ne permet pas d'atteindre les graphes de dépendance
- Pour les graphes de dépendance locale :  $\leftrightarrow$  processus de Hawkes + méthode Lasso
- Cette méthode permet de traiter les cas d'excitation, d'inhibition et permet de retrouver des graphes cohérents même si le modèle n'est pas tout à fait exact

# Conclusions

## questions ouvertes

- On peut tester l'hypothèse de Hawkes, mais comment quantifier la distance au modèle de Hawkes?
- Comment prendre en considération la non-stationnarité? (travail en cours avec Patricia Reynaud-Bouret et Franck Picard (Lyon))
- Comment prendre en compte le fait que nous ne travaillons que sur un faible nombre de neurones?

# Références

-  Grün, S., Diesmann, M., & Aertsen, A.M. *Unitary Events Analysis. In Analysis of Parallel Spike Trains*, Grün, S., & Rotter, S., Springer Series in Computational Neuroscience (2010).
-  Tuleau-Malot, C., Rouis, A., Grammont, F., Reynaud-Bouret, P. *Multiple tests based on a Gaussian Approximation of the Unitary Events method with delayed coincidence count*. Neural Computations, vol.26, 7, 1408–1454, 2014.
-  Hansen, N.R, Reynaud-Bouret, P., Rivoirard, V. *Lasso and probabilistic inequalities for multivariate point processes*. Bernoulli, 21(1), 83–143, 2014.
-  Reynaud-Bouret, P., Rivoirard, V., Grammont, F., Tuleau-Malot, C. *Goodness-of-fit tests and nonparametric adaptive estimation for spike train analysis*. Journal of Mathematical Neuroscience (2014).

## Références (2)

-  Reynaud-Bouret, P., Rivoirard, V., Tuleau-Malot, C. *Inference of functional connectivity in Neurosciences via Hawkes processes*. 1st IEEE, Global Conference on Signal and Information Processing, Austin Texas, 2013.
-  Lambert, R., Tuleau-Malot, C., Bessaih, T., Rivoirard, V., Bourret, Y., Leresche, N., Reynaud-Bouret, P. *Reconstructing the functional connectivity of multiple spike trains using Hawkes models*. Journal of Neuroscience Methods, 297, 9–21, 2018.