

# MTGAUE : TESTS MULTIPLES POUR TESTER LOCALEMENT L'INDÉPENDANCE ENTRE DEUX NEURONES

Christine Tuleau-Malot <sup>1</sup> & Amel Rouis <sup>2</sup> & Franck Grammont <sup>1</sup> & Patricia  
Reynaud-Bouret <sup>1</sup>

<sup>1</sup> *Univ. Nice Sophia Antipolis, CNRS, LJAD, UMR 7351, 06100 Nice, France. Email :  
malot@unice.fr, grammont@unice.fr, reynaudb@unice.fr.*

<sup>2</sup> *Centre Hospitalier Universitaire de Nice, France. Email : rouis.a@chu-nice.fr.*

**Résumé.** En Neurosciences, un enjeu majeur est d'acquérir une meilleure compréhension de la dynamique de l'activité neuronale et notamment d'identifier les connectivités fonctionnelles pouvant exister entre neurones. Nous proposons ici une procédure, inspirée de la méthode des Unitary Events, permettant de tester localement l'indépendance entre deux neurones.

**Mots-clés.** Processus de Poisson, tests multiples, Unitary Events, Multiple Shift

**Abstract.** In Neurosciences, a major challenge is to obtain a better understanding of the dynamic of the neuronal activity and in particular to identify the functional connectivity that may exist between neurons. We propose a procedure, based on the Unitary Events method, to locally test the independence between two neurons.

**Keywords.** Poisson process, Multiple testing, Unitary Events, Multiple Shift

## 1 Introduction

En Neurosciences, un des intérêts majeurs est d'obtenir une meilleure compréhension de la dynamique du cerveau et notamment d'identifier des connectivités fonctionnelles pouvant exister entre différents neurones en réponse à une stimulation donnée. En effet, s'il semble acquis qu'il existe des assemblées de neurones, toute la question est d'identifier les neurones inter-agissant entre eux, et la manière dont ils inter-agissent. Pour répondre à cette problématique, on considère le temps d'occurrence des potentiels d'actions, ou spikes, qui sont considérés comme le meilleur résumé de l'activité neuronale. Grâce aux avancées technologiques des électrodes, on dispose aujourd'hui de l'enregistrement simultané de plusieurs trains de spikes, chaque train étant associé à un neurone. A partir de ces enregistrements, les chercheurs ont essayé de tester l'indépendance entre deux neurones, à savoir tester "les activités de chacun des deux neurones sont indépendantes" contre "elles sont dépendantes". Une des méthodes les plus populaires et les plus efficaces à ce jour est la méthode des Unitary Events (UE) développée par Grün et collaborateurs (Grün (1996)). Si la méthode originale se base sur la notion de coïncidences binnées (détaillée ci-dessous),

elle a été continuellement améliorée jusqu’à l’obtention de la version actuelle considérant la notion de coïncidences par “Multiple Shift” (cf. Grün et al. (1999)). Cependant, cette méthode souffre d’un certains nombres de défauts que nous avons cherché à corriger en proposant une nouvelle méthode permettant d’étendre la validité de la méthode originelle des Multiple Shift UE.

## 2 La notion de coïncidence

Informellement, en Neurosciences, une coïncidence représente la capacité de deux neurones à produire des spikes proches temporellement. Une question centrale de ce travail est de définir proprement une coïncidence et surtout le nombre de coïncidences. Cependant avant de pouvoir proposer un formalisme pour la notion de coïncidence, il faut modéliser correctement un train de spikes.

En raison de la résolution d’acquisition  $h$  des données (de l’ordre de  $10^{-3}$ ,  $10^{-4}$ ), un train de spikes n’est autre qu’une séquence de 0 et 1, notée  $(H_n)_n$  où  $H_i = 1$  traduit la présence d’un spike sur l’intervalle temporel  $[ih - h/2, ih + h/2)$ . Ainsi, à une telle séquence, il est possible d’associer un processus ponctuel  $N$  défini comme étant l’ensemble des points  $ih$  pour lesquels  $H_i = 1$ .  $N$  est donc un processus discrétisé à la résolution  $h$ .

On note  $N_1$  et  $N_2$  les processus ponctuels, et  $(H_n^1)_n$  et  $(H_n^2)_n$  les séquences, associés respectivement au train de spikes du neurone 1 et du neurone 2 enregistrés simultanément. Nous nous limitons pour l’instant à une fenêtre temporelle  $W$  de longueur  $T$ , avec  $T = nh$ , fenêtre beaucoup plus petite que la fenêtre totale d’enregistrement.

A ce niveau de résolution  $h$ , il est très rare d’avoir des indices  $i$  tels que  $H_i^1 = H_i^2 = 1$ .

La méthode classique des UE, dite du “binning”, commence par un prétraitement des données. Pour un entier  $d > 1$  fixé, on transforme  $(H_n^j)_n$  en une séquence de 0 et 1 de longueur  $r = T/(d * h)$ , notée  $(D_r^j)_r$  de sorte que, pour  $i \in \{1, \dots, r\}$ ,  $D_i^j = 1$  signifie qu’il existe au moins un entier  $k$  dans le “bin”  $\{(i - 1) * d + 1, i * d\}$  tel que  $H_k^j = 1$ . Pratiquement, cela signifie que le niveau de résolution de la nouvelle séquence est plus faible. A partir de cette transformation, on dit qu’il existe une coïncidence au temps  $C = i * d * h$  entre le neurone 1 et le neurone 2, si  $D_i^1 = D_i^2 = 1$ . Le nombre de coïncidences entre le neurone 1 et le neurone 2 sur l’intervalle  $W$  n’est autre que le nombre de tels indices  $i$ .

Si cette construction et ces coïncidences sont encore largement utilisées, elles souffrent d’un inconvénient majeur, le découpage en “bins”, qui peut entraîner une sous-estimation du phénomène de synchronisation.

Dans (Grün et al. (1999)) a été introduite la méthode des “multiple shift” UE qui utilise une nouvelle notion de coïncidences permettant de corriger cette perte de détection dans la synchronisation. Le principe général de cette méthode, qui est détaillé en Figure 1 cas non symétrique, consiste à décaler, jusqu’à un déplacement (“shift”) maximal,  $(H_n^2)_n$  par rapport  $(H_n^1)_n$  qui lui reste fixe et à comptabiliser les coïncidences de chacun des shifts

intermédiaires.

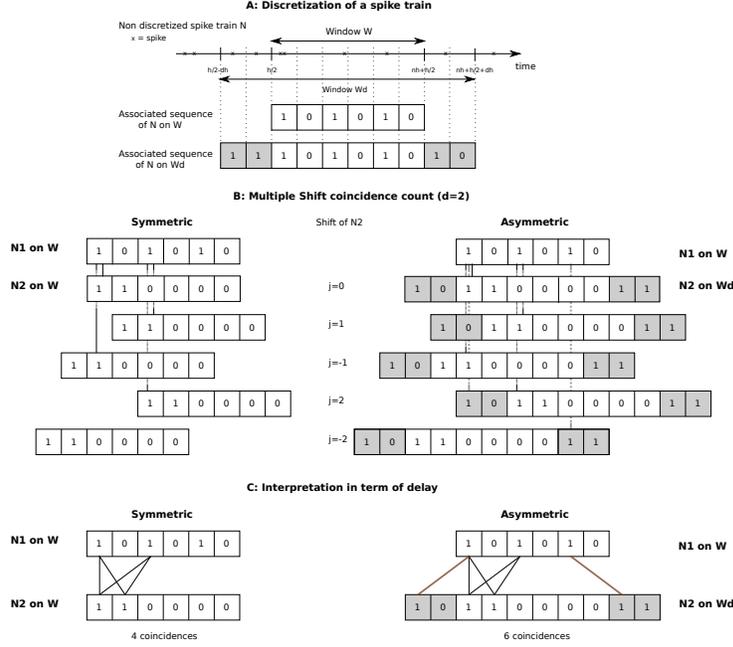


Figure 1: Sur **A**, schéma de la discrétisation. Sur **B** et **C**, visualisation de la notion de coïncidence (symétrique et non symétrique) évaluée par “Multiple Shift”. Les cases grises correspondent aux données en dehors de  $W$  intervenant dans la version non symétrique.

Si cette définition permet de diminuer la perte en détection, elle souffre de d’un inconvénient. En effet, la notion de coïncidence ainsi définie n’est pas symétrique puisque dans cette définition interviennent les points de  $(H_n^1)_n$  contenus dans  $W$  et les points de  $(H_n^2)_n$  contenus dans une fenêtre étendue de  $W$ . Ceci fait que lorsque l’on inverse le rôle de  $(H_n^1)_n$  et  $(H_n^2)_n$  les résultats ne sont pas identiques. C’est la raison pour laquelle nous proposons une version symétrique de cette notion de coïncidence, définition qui se formalise ainsi :

$$X = \sum_{i=1}^n \sum_{k=1}^n \mathbf{1}_{|k-i| \leq d} \mathbf{1}_{H_i^1=1} \mathbf{1}_{H_k^2=1} \quad (1)$$

Si les processus ponctuels actuels sont discrétisés et donc modélisés par des processus de Bernoulli, en raison de l’ordre de grandeur de  $h$ , on peut modéliser les trains de spikes par des processus de Poisson. C’est cette modélisation que nous allons utiliser dans toute la suite et qui va permettre de trouver la loi asymptotique de  $X$ .

La généralisation de l’équation (1) est :

$$X = \int_{W^2} \mathbf{1}_{|x-y| \leq \delta} N_1(dx) N_2(dy).$$

Cette définition est celle d'une coïncidence avec délai  $\delta$ .

### 3 Test local d'indépendance

Afin de construire un test de  $\mathcal{H}_0$  : “ $N_1$  et  $N_2$  sont indépendants” contre  $\mathcal{H}_1$  : “ $N_1$  et  $N_2$  sont dépendants”, nous nous focalisons sur la notion de coïncidence avec délai, notion sur laquelle nous avons prouvé :

**Théorème 1** Fixons  $\delta$  tel que  $0 < 2\delta < T$ .

Si  $N_1$  et  $N_2$  sont des processus de Poisson homogènes indépendants d'intensités respectives  $\lambda_1$  et  $\lambda_2$  sur  $W$ , alors :

$$m_0 := E(X) = \lambda_1 \lambda_2 [2\delta T - \delta^2]$$

et

$$\sigma^2 := Var(X) = \lambda_1 \lambda_2 [2\delta T - \delta^2] + [\lambda_1^2 \lambda_2 + \lambda_1 \lambda_2^2] \left[ 4\delta^2 T - \frac{10}{3} \delta^3 \right].$$

Par ailleurs, si l'on dispose de  $M$  essais i.i.d. de  $N_1$  et  $N_2$  notés  $(N_1^{(m)})_{1 \leq m \leq M}$  et  $(N_2^{(m)})_{1 \leq m \leq M}$ , alors :

$$\sqrt{M} \frac{\bar{m} - m_0}{\sqrt{\sigma^2}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1),$$

où  $\bar{m}$  le nombre moyen de coïncidences avec délai  $\delta$ , i.e.

$$\bar{m} = \frac{1}{M} \sum_{m=1}^M X^{(m)} \quad \text{with} \quad X^{(m)} = \int_{W^2} \mathbf{1}_{|x-y| \leq \delta} N_1^{(m)}(dx) N_2^{(m)}(dy).$$

Ce résultat en lui-même ne permet pas de construire un test car il fait intervenir des paramètres inconnus comme  $\lambda_1$  et  $\lambda_2$ . Nous avons alors prouvé :

**Théorème 2** Avec les mêmes notations que précédemment, soit  $\hat{\lambda}_j$  l'estimateur sans biais de  $\lambda_j$ , de taux de décharge du neurone  $j$ , défini par :

$$\hat{\lambda}_j := \frac{1}{MT} \sum_{m=1}^M N_j^{(m)}(W).$$

Soit  $\hat{m}_0$  un estimateur de  $m_0$  défini par

$$\hat{m}_0 := \hat{\lambda}_1 \hat{\lambda}_2 [2\delta T - \delta^2].$$

Sous les hypothèses du Théorème précédent, on a :

$$\sqrt{M} (\bar{m} - \hat{m}_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, v^2),$$

où

$$v^2 := \lambda_1 \lambda_2 [2\delta T - \delta^2] + \lambda_1 \lambda_2 [\lambda_1 + \lambda_2] \left[ \frac{2}{3} \delta^3 - T^{-1} \delta^4 \right].$$

Par ailleurs  $v^2$  peut être estimé par :

$$\hat{v}^2 := \hat{\lambda}_1 \hat{\lambda}_2 [2\delta T - \delta^2] + \hat{\lambda}_1 \hat{\lambda}_2 [\hat{\lambda}_1 + \hat{\lambda}_2] \left[ \frac{2}{3} \delta^3 - T^{-1} \delta^4 \right]$$

et

$$\sqrt{M} \frac{\bar{m} - \hat{m}_0}{\sqrt{\hat{v}^2}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

Ces résultats nous permettent alors de définir des tests de  $\mathcal{H}_0$  contre  $\mathcal{H}_1$ , tests basés sur l'approximation gaussienne du Théorème 2.

- *le test symétrique*  $\Delta_{GAUE}^{sym}(\alpha)$  of  $\mathcal{H}_0$ : "  $N_1$  et  $N_2$  sont indépendants" contre  $\mathcal{H}_1$ : "  $N_1$  sont  $N_2$  dépendants", qui rejette  $\mathcal{H}_0$  quand  $\bar{m}$  et  $\hat{m}_0$  sont trop différents :

$$|\bar{m} - \hat{m}_0| \geq z_{1-\alpha/2} \sqrt{\frac{\hat{v}^2}{M}}$$

- *le test unilatéral par valeur supérieur*  $\Delta_{GAUE}^+(\alpha)$  qui rejette  $\mathcal{H}_0$  quand  $\bar{m}$  est trop large :

$$\bar{m} \geq \hat{m}_0 + z_{1-\alpha} \sqrt{\frac{\hat{v}^2}{M}}$$

- *le test unilatéral par valeur inférieur*  $\Delta_{GAUE}^-(\alpha)$  qui rejette  $\mathcal{H}_0$  quand  $\bar{m}$  est trop petit :

$$\bar{m} \leq \hat{m}_0 - z_{1-\alpha} \sqrt{\frac{\hat{v}^2}{M}}$$

où  $z_t$  est le  $t$ -quantile de  $\mathcal{N}(0, 1)$ .

Si  $N_1$  et  $N_2$  sont des processus de Poisson homogènes alors ces tests sont asymptotiquement de niveau  $\alpha$ .

La difficulté est que dans la pratique, il est rare d'avoir un régime ne serait-ce que stationnaire sur toute la durée de l'enregistrement. C'est pourquoi la totalité de l'enregistrement a été découpée en des sous-fenêtres  $W$  sur lesquelles nous pouvions considérer les processus homogènes puis nous avons appliqué notre test sur les différents  $W$ .

Pour pouvoir garantir le taux de faux positifs, nous avons appliqué la procédure de tests multiples de Benjamini et Hochberg (Benjamin et Hochberg (1995)).

Voici en définitive la procédure MTGAUE que nous avons proposée :

- Pour chaque sous fenêtre  $W(i)$  de  $\{W(1), \dots, W(K)\}$ , les sous-fenêtres pouvant se chevaucher, on calcule la p-valeur du test GAUE, notée  $P_{W(i)}^{(i)}$ .
- Pour un paramètre fixé  $q$ , qui contrôle le taux de faux positif (FDR), on ordonne ces p-valeurs  $P_{W(1)}^{(1)} \leq \dots \leq P_{W(\ell)}^{(\ell)} \leq \dots \leq P_{W(K)}^{(K)}$  et on détermine  $k$  tel que  $k = \max\{\ell \text{ such that } P_{W(\ell)}^{(\ell)} \leq \ell q/K\}$ .
- Les plages détectées, à savoir celles où  $N_1$  et  $N_2$  sont déclarés dépendants, correspondent aux  $k$  sous-fenêtres associées aux  $k$  plus petites valeurs propres.

Cette procédure, validée théoriquement, a été testée tout d’abord sur données simulées puis sur données réelles. Cela a permis de montrer que :

- la procédure garantit le taux de FDR en pratique, sans utiliser le terme correctif en cas de chevauchement des fenêtres,
- la procédure est réellement utilisable avec des temps de calculs raisonnables, et des résultats tout à fait probants, et meilleurs que ceux obtenus par exemple par la méthode des multiple shift UE,
- la procédure est robuste à la modélisation, ce qui est intéressant pour la pratique, car nous savons que l’activité neuronale se modélise rarement par un processus de Poisson homogène.

## Bibliographie

- [1] Grün, S. (1996), *Unitary joint-events in multiple-neuron spiking activity: Detection, significance and interpretation.*, Thun: Verlag Harri Deutsch.
- [2] Grün, S., Diesmann, M., Grammont, F., Riehle, A., et Aertsen, A.M. (1999), Detecting unitary events without discretization of time, *Journal of Neuroscience Methods*, 93, 67–79.
- [3] Benjamini, Y. et Hochberg, Y., (1995), Controlling the false discovery rate: a practical and powerful approach to multiple testing, *Journal of the Royal Statistical Society, Series B*, 57(1), 289–300.