

Détection dynamique de graphes appliquée à l'analyse de signaux électroencéphalographiques issus du sommeil de rat

Guillaume BECQ¹, Pierre-Olivier AMBLARD¹, Paul SALIN², Jean-Christophe COMTE²

¹Laboratoire Gipsa-lab, UMR 5216 CNRS

²Centre de Recherche en Neurosciences de Lyon U1028 / UMR 5292

guillaume.becq@gipsa-lab.fr, pierre-olivier.amblard@gipsa-lab.fr,
paul.salin@sommeil.univ-lyon1.fr, jchrist.comte@gmail.com

Résumé – L'objectif de ce papier est de proposer une méthode pour l'estimation de graphes de connectivité orientés entre des signaux provenant de différentes aires du cerveau au cours de différentes phases de veille et sommeil. La non stationnarité des données est prise en compte pour évaluer les liens entre les activités des différentes aires de façon à rendre compte de l'aspect dynamique des graphes sous-jacents. Pour cela, une méthode d'estimation de cohérences partielles dirigées est testée et analysée sur un modèle qui génère des signaux à partir d'une séquence de modèles autorégressifs multivariés supposée mimer l'activité du cerveau durant différents états de vigilance. Cette méthode est ensuite utilisée sur des signaux réels électroencéphalographiques provenant de trois aires du cerveau d'un rat durant plusieurs phases de veille et sommeil. La méthode met en évidence, d'une part le caractère dynamique de l'échange entre régions du cerveau, d'autre part la distribution des principaux graphes de connectivité au cours des différents états de vigilance.

Abstract – The aim of this paper is to propose a method for the estimation of oriented graphs of connectivity between signals from different areas of the brain during different epochs of sleep and wake. The non stationarity is taken into account to evaluate the links between the activities of the different areas to put in evidence the dynamical aspect of the underlying graphs. For that, a method to estimate the partial directed coherence is tested and analysed on a model that generates signals from sequences of multivariate autoregressive models supposed to mimic the dynamical aspect of changes of the links of the areas of the brain during different vigilance states. The method is tested on real data, three electroencephalographic signals coming from three different areas of the brain of a rat recorded during three vigilance states. The method puts in evidence, on one hand, the dynamical behavior of the exchange between the areas of the brain, on the other hand, the distribution of the principal connectivity graphs for the different states of vigilance.

1 Introduction

L'analyse du sommeil ou des différents états de vigilance sur l'animal est un des premiers champs d'investigation en neurosciences depuis l'observation des variations au cours de la nuit des grapho-éléments de l'électroencéphalogramme (EEG) [5]. L'analyse des signaux EEG sur des fenêtres de quelques secondes, 5, 10, 20 ou 30 s, conduit à une classification de ces fenêtres en différentes phases de veille et sommeil que l'on observe chez différentes espèces en expérimentation, principalement l'homme, le rat, la souris et le chat. Dans cette étude seules 3 phases fondamentales du sommeil sont retenues : éveil (wake - W), sommeil lent profond (slow wave sleep - SWS) et sommeil paradoxal (paradoxical sleep - PS). L'évolution de ces phases au cours de la nuit conduit à une représentation nommée hypnogramme. Un exemple d'évolution des phases de veille et sommeil d'un rat sur une période de vingt minutes est proposé Fig. 1. L'étude des électroencéphalogrammes et des hypnogrammes permet de mettre en évidence des pathologies ou des désordres psychiques et fait partie des examens cliniques habituellement réalisés à l'hôpital [6, 7, 4]. De façon à améliorer la compréhension de ces phases, il est intéressant

d'analyser l'évolution des interactions des différentes parties du cerveau au cours de celles-ci [3]. Dans cet article, nous proposons une méthode pour estimer les relations entre mesures EEG enregistrées à différents endroits du cerveau en fonction des phases de veille et sommeil. Pour analyser ces données, une méthode d'estimation de cohérences partielles dirigées proposées dans [1] est testée et analysée sur un modèle qui génère des signaux à partir d'une séquence de modèles autorégressifs multivariés supposée mimer l'aspect dynamique des changements de connectivité au cours des différents états de vigilance. La méthode est ensuite évaluée sur des mesures électroencéphalographiques provenant de trois électrodes implantées dans trois régions du cerveau d'un rat. Une comparaison des résultats est proposée et discutée sur les données simulées ou réelles.

2 Description du modèle simulé et de la méthode de détection des états

Dans cette partie nous proposons un modèle simple qui mime des changements d'activités correspondant à un changement de

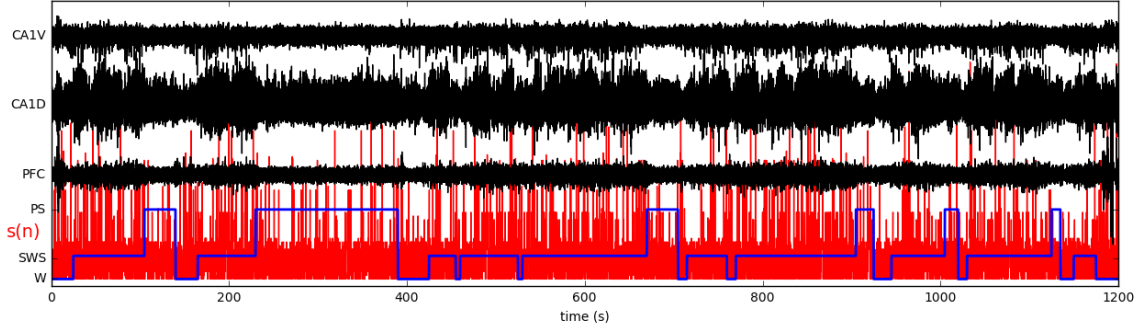


FIGURE 1 – Succession des états obtenus par DPC en confrontation avec l'hypnogramme du rat sur une durée de 20 min. En noir, les données EEG provenant de PFC, CA1D et CA1V. Un hypnogramme en bleu a été réalisé visuellement par un expert par observation des données EEG sur des périodes de 5 s. La courbe rouge correspond à la succession d'états trouvés par la machine sur des fenêtres de 200 ms non chevauchantes. $s(n)$ correspond à la succession des états trouvés par la machine (voir texte pour détail).

connectivités entre différentes variables. Une méthode d'estimation des cohérences partielles dirigées est utilisée sur des petites fenêtres d'analyse. Ceci permet de rendre compte de l'état instantané du système. Une règle de décision simple permet d'associer à cet état un graphe de connectivité qui met en évidence les interactions entre variables.

On considère le modèle autorégressif multivarié reliant N variables :

$$x_i(n) = \sum_{j=1}^N w_{s(n-t_{i,j}),i,j} x_j(n-t_{i,j}) + u_i(n)$$

qui fournit la valeur à l'instant n de la variable x_i en fonction des valeurs des autres variables x_j aux instants $n-t_{i,j}$. $w_{s(n),i,j}$ correspond à une pondération des entrées et $u_i(n)$ à un bruit. Les poids $w_{s(n),i,j}$ varient au cours du temps en fonction de l'état à l'instant n du système $s(n)$. Pour notre étude on fixe : le nombre de variables à $N = 3$; u_i une variable aléatoire qui suit une loi gaussienne de moyenne $m = 0$ et d'écart-type $s = 0.1$; Les délais entre signaux sont donnés par : $\tau = \begin{pmatrix} 1 & 3 & 5 \\ 3 & 1 & 7 \\ 5 & 7 & 1 \end{pmatrix}$. L'état du système $s(n)$ est une valeur associée à un graphe de connectivité correspondant à un ensemble de 3 noeuds en relations. Pour cet ensemble de 3 noeuds, dans cette étude, toutes les connections intra groupe sont exclues, et on s'intéresse uniquement aux $64 = 2^6$ matrices de connectivité possibles sur l'ensemble des $512 = 2^9$ matrices qui prennent en compte toutes les connections. En d'autres termes les matrices avec des 0 ou des 1 sur la diagonale seront considérées comme ayant un 0 sur la diagonale. Dans la suite de l'article les connections $\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ seront notées en version plate (abcdefghi). Ainsi la matrice de connectivité de l'état 32 $s_{32} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ sera notée (000100000), qui peut être vu comme un mot binaire dont la valeur décimale vaut 32. Les états suivants 0 (000000000), 2 (000000010), 32 (000100000), 36 (000100100), 160 (010100000) seront utilisés pour simuler des signaux. Ils correspondent respectivement à : s_0 , un état déconnecté; s_{32} , un état où x_1 influence x_2 ; s_{36} ,

x_1 influence x_2 et x_3 ; s_2 , x_2 influence x_3 ; s_{160} , une boucle avec x_1 qui influence x_2 , et x_2 qui influence x_1 . Les poids des matrices $w_{i,\dots}$ sont obtenus à partir des s_i décrit précédemment avec des valeurs fixées à 0.5 sur la diagonale et à 0.4 aux emplacements des 1 des s_i . Par exemple $w_{160,\dots} = \begin{pmatrix} 0.5 & 0.4 & 0 \\ 0.4 & 0.5 & 0 \\ 0 & 0 & 0.5 \end{pmatrix}$

Pour retrouver la connectivité dynamique sur ces données, les signaux sont extraits sur des fenêtres non chevauchantes de 200 échantillons. Pour les signaux extraits sur chaque fenêtre, la méthode d'estimation des cohérences partielles dirigées (directed partial coherence - DPC) proposée dans [1] est utilisée. Cette méthode consiste en une factorisation spectrale forte de l'inverse de la matrice de densité spectrale pour des signaux multivariés. Ceci permet d'obtenir des paramètres autorégressifs correspondant à des systèmes causaux, tout en estimant proprement les matrices d'intercorrélation des bruits en entrée du système. Elle propose aussi de tester, de façon non paramétrique, les bandes de fréquence qui montre une DPC supérieure à celles moyennes obtenues sur plusieurs réalisations obtenues à partir de données simulées aléatoirement en se basant sur les modèles estimés. Elle permet de ne pas avoir à fixer arbitrairement différents paramètres. De façon à indiquer si une connection existe entre x_i et x_j , un test sur toutes les bandes de fréquences est réalisé. Ce test est positif si au moins une bande de fréquence montre un résultat positif avec un seuil de signification α fixé à 0.05. Pour notre étude les spectres ont été estimés par la méthode de Welch sur 11 fenêtres de 16 échantillons, pour chaque fenêtre de 200 échantillons, sans recouvrement. Les hypothèses de non cohérence ont été testées sur 30 essais. Pour les données réelles, les signaux sont échantillonnés à 1000 Hz. Les données simulées peuvent aussi être considérées comme échantillonnées à cette fréquence. Les paramètres de détection utilisées sur les jeux de données simulées ou réelles sont les mêmes.

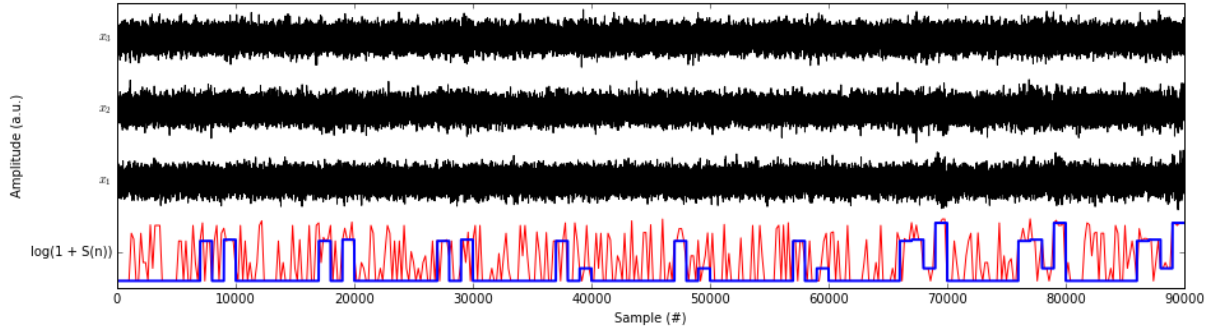


FIGURE 2 – Signaux simulés et séquences d'états. La courbe bleue représente la séquence des états (0, 32, 36, 2, 160) pour générer les trois signaux en noir, de haut en bas de x_1 vers x_3 . En rouge, la séquence des états, trouvée par la méthode de PDC sur des fenêtres non chevauchantes de 200 ms.

3 Resultats

3.1 Test sur les données simulées

La séquence des états simulés et les signaux résultants sont présentés Fig 2. Les états 2, 32, 36 et 160 ont été maintenus pendant 1 s, séparés ou non par des repos à l'état 0 suivant 3 phases avec 3 motifs périodiques. On remarque que les états estimés en rouge sont souvent proches des états simulés en bleu. Cet exemple montre que la méthode proposée permet de retrouver partiellement les états qui ont servi à obtenir les données. On remarque de nombreuses erreurs, mais comme le numéro de l'état n'est pas ordonné, cela ne précise pas la gravité de l'erreur par rapport à l'état espéré. Une analyse plus fine des distributions des confusions entre états trouvés par l'algorithme, en colonne, et ceux simulés, en ligne, est proposé dans la table 1. Les éléments diagonaux de la matrice de confusion sont représentés en gras. Les taux de bonnes classification sont respectivement de 57 %, 40 %, 38 %, 23 % et 67 % pour les états 0, 2, 32, 36 et 160 utilisés pour la simulation. Une distribution aléatoire sur les 64 états possible conduirait à un taux de 1.6 %. Etant donné la complexité du système, on constate donc que la méthode proposée se révèle très efficace.

Pour l'état 0, les confusions se font favorablement sur les petits réseaux avec une seule arête dirigée (4, 8, 32, 64, 128), puis ensuite sur les réseaux avec 2 arêtes (6, 10, 12, 34, 36 etc.). Les états 2, 32, et 36 sont principalement confondus avec 0, ce qui doit correspondre à un niveau de bruit relativement haut qui empêche de détecter le modèle sous-jacent sur de si petites fenêtres d'analyse. L'état 160 n'est pas confondu avec 0 mais principalement avec 128. On remarque sur la Fig. 2, que pour cet état, les signaux générés avec les paramètres du modèle autorégressif retenu montrent une dynamique liée au modèle qui doit surpassé le niveau de bruit, ce qui se révèle favorable pour le calcul des PDC.

3.2 Test sur les données réelles

Les données électroencéphalographiques sont enregistrées dans 3 régions qui correspondent à : 1 - cortex préfrontal (pre-

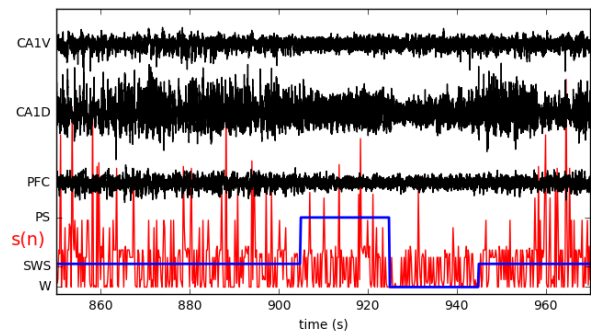


FIGURE 3 – zoom (120 s) de Fig. 1.

frontal cortex - PFC), 2 - l'hippocampe dorsal (cornu ammonis 1 dorsal - CA1D) et 3 - l'hippocampe ventral (cornu ammonis 1 ventral - CA1V). La détection sur 20 min, soit 6000 fenêtres d'analyse de 200 ms non chevauchantes, avec des paramètres identiques à ceux utilisés dans le cas des données simulées, fournit une succession d'états mise en correspondance avec l'hypnogramme réalisé par l'expert, proposée Fig. 1 et Fig. 3. On remarque que la dynamique des états trouvés par notre méthode varie au cours des phases de veille et sommeil et que, comme sur les données simulées, des états semblent être dominants pendant certaines périodes. Comme indiqué pour les données simulées, il est difficile de savoir si les états successifs sont proches l'un de l'autre.

La distribution des états détectés en fonction des phases de veille et sommeil est donnée dans la table de confusion 2. Les classes de l'hypnogramme sont proposées en ligne et les états trouvés par la machine sont en colonne. Seuls les états observés avec une probabilité supérieure à 1 % sont reportés, tous les autres sont sommés dans la colonne 'autres'. On remarque que pour les phases W, SWS et PS, l'état 0 est le plus présent et représente environs 30 % des états trouvés. Pour toutes les phases, les états 4 (00000100), 8 (000001000), 32 (000100000), 36 (000100100) et 40 (000101000) sont les plus représentés et correspondent à des projections de PFC sur CA1D ou CA1V, ou des projections de CA1V sur CA1D. En SWS,

TABLE 1 – Table de confusion pour les données simulées, en pourcentage.

	0	2	4	6	8	10	12	32	34	36	40	42	46	64	66	68	72
0	57,0	3,9	6,1	0,3	5,2	0,6	0,6	5,2	0,3	0,3	0,0	0,3	0,0	6,1	0,6	1,2	0,3
2	33,3	40,0	0,0	6,7	3,3	0,0	0,0	3,3	3,3	0,0	0,0	0,0	0,0	0,0	3,3	0,0	0,0
32	13,3	2,2	2,2	0,0	2,2	0,0	0,0	37,8	6,7	8,9	6,7	0,0	0,0	2,2	0,0	0,0	0,0
36	10,0	3,3	16,7	0,0	0,0	0,0	0,0	26,7	6,7	23,3	0,0	0,0	0,0	0,0	0,0	0,0	0,0
160	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	6,7	0,0	0,0	0,0	0,0
	96	98	102	128	130	132	134	136	160	162	164	168	192	196	224	228	
0	0,6	0,0	0,0	7,3	0,3	0,9	0,3	0,6	0,3	0,0	0,0	0,6	0,3	0,3	0,3	0,3	0,3
2	0,0	0,0	0,0	0,0	0,0	3,3	0,0	0,0	3,3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
32	2,2	2,2	0,0	0,0	0,0	0,0	0,0	0,0	8,9	2,2	0,0	0,0	0,0	0,0	2,2	0,0	0,0
36	0,0	0,0	3,3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	3,3	0,0	0,0	3,3	0,0	3,3	0,0
160	0,0	0,0	0,0	13,3	0,0	0,0	0,0	6,7	66,7	0,0	0,0	6,7	0,0	0,0	0,0	0,0	0,0

TABLE 2 – Table de confusion pour les données du rat en phase de veille et sommeil, en pourcentage. Seuls les états avec un taux supérieur à 1 % sont reportés, les autres sont regroupés dans la colonne 'autres'. Les classes de l'hypnogramme sont en ligne, et les états proposés par la machine sont en colonne.

	0	2	4	8	12	32	36	40
W	34,3	2,3	6,3	11,8	0,9	14,1	6,3	8,0
SWS	28,9	1,5	3,7	14,5	1,0	15,2	3,7	14,1
PS	30,7	0,9	6,7	10,8	1,4	22,3	6,7	9,5
	44	64	72	104	128	192	autres	
W	1,7	4,3	1,8	0,8	0,8	1,2	5,4	
SWS	2,2	2,6	4,2	2,0	1,2	0,4	5,0	
PS	2,4	0,7	1,7	0,8	0,4	0,0	4,9	

CA1D reçoit plus de signaux de CA1V car les états 8 et 40 prennent une proportion plus importante. En PS, on remarque une implication plus forte du réseau PFC vers CA1D mis en évidence par un pourcentage élevé de l'état 32. On remarque que les états supérieurs à 128 sont peu représentés, ce qui indique que PFC reçoit peu de données de la part des autres régions, sauf en SWS ou PFC reçoit un peu plus de signaux de la part des autres régions, comme indiqué par la proportion plus forte des états 72 (001001000), 104 (001101000) et 128 (010000000).

4 Discussion, conclusion

Ces premiers résultats montrent que la méthode de détection de connectivité sur des petites fenêtres permet de mettre en évidence la dynamique des graphes de connectivité au cours de différents modes de fonctionnement du cerveau tels que les phases de veille et sommeil. La méthode proposée semble prometteuse car étant donné la complexité de la détection, les performances obtenues sur les données simulées se révèlent très largement supérieures à celles obtenues de façon aléatoire. Des analyses plus fines sur tous les états et sur d'autres jeux de paramètres pour les modèles multivariés doivent être réalisées pour préciser le cadre d'application de la méthode. La méthode doit aussi être optimisée pour que les calculs soient plus rapides car elle est très coûteuse en calcul. Enfin, la méthode devra être confrontée avec d'autres algorithmes tels que ceux proposés par exemple dans [8] pour conclure sur son intérêt face aux autres méthodes existantes.

Références

- [1] P.-O. Amblard, *A nonparametric efficient evaluation of partial directed coherence*, Biological Cybernetics, 109(2) : 203–214 2015.
- [2] F. Chapotot, G. Becq, *Automated sleep–wake staging combining robust feature extraction, artificial neural network classification, and flexible decision rules*, Int. J. Adapt. Control Signal Process., 24 : 409–423, 2010.
- [3] Comte, J. ; Ravassard, P. & Salin, P. *Sleep dynamics : a self-organized critical system*, Physical Review E, 73(5) : 056127, 2006.
- [4] American Association of Sleep Medicine & others *The international classification of sleep disorders : diagnostic and coding manual* American Acad. of Sleep Medicine, 2005.
- [5] Jouvet, M. ; Michel, F. & Courjon, J. *Sur un stade d'activité électrique cérébrale rapide au cours du sommeil physiologique* CR Soc. Biol.(Paris), 153 : 1024–1028, 1959.
- [6] Kupfer, D. J. ; Foster, F. G. ; Coble, P. ; McPartland, R. J. & Ulrich, R. F. *The application of EEG sleep for the differential diagnosis of affective disorders*, Am. J. Psychiatry, Am. Psychiatric Assoc., 135 : 69–74, 1978.
- [7] Schenck, C. H. ; Milner, D. M. ; Hurwitz, T. D. ; Bundlie, S. R. & Mahowald, M. W. *A polysomnographic and clinical report on sleep-related injury in 100 adult patients* Am J Psychiatry, Am Psychiatric Assoc, 146 : 1166–73, 1989.
- [8] H.E. Wang et al., *A systematic framework for functional connectivity measures*, Frontiers in Neuroscience, 8(405) : 1–22, 2014.