Étude et modélisation du codage corrélationnel et application aux réseaux neuronaux

Pascale Giraudet

Rapport de stage
DEA de Sciences Cognitives
Institut de la Communication Parlée
Institut National Polytechnique de Grenoble

Responsable du stage :
Frédéric Berthommier

Jury :
J. Demongeot
C. Abry
J.L. Schwartz

Remerciements

Je tiens à remercier l'Institut de la Communication Parlée, pour m'avoir accueillie au sein de son laboratoire, et permis d'effectuer mon stage de DEA dans les meilleures conditions.

Je remercie tout particulièrement Frédéric Berthommier, responsable du bon déroulement de mon stage, pour son encadrement efficace, ses idées inépuisables, sa disponibilité... et la patience dont il a dû faire preuve pour guider une biologiste à travers la modélisation et le traitement du signal.

Merci aussi à MM. Demongeot, Abry et Schwartz pour s'être penchés sur mon travail et avoir accepté de faire partie du jury.

Je remercie tous les enseignants du DEA de Sciences Cognitives qui ont su tenir compte de nos origines très diverses, et rester intéressants tout en étant accessibles à tous.

Merci enfin à tous ceux qui m'ont aidée et encouragée tout au long de ces trois mois de stage.

Résumé

De nombreuses études électrophysiologiques tendent à révéler l'insuffisance de l'approche du codage neuronal par la seule intensité moyenne de décharge. Neurophysiologistes et modélisateurs s'intéressent donc à la possibilité d'un codage par corrélations temporelles fines entre les activités de populations neuronales. Ce mode de codage fait d'autant plus parler de lui qu'il pourrait donner une explication au problème non résolu du « binding » des composantes perceptuelles d'un objet.

Nous basons le présent travail sur l'hypothèse selon laquelle les neurones sont capables de transmettre des informations selon deux types de codes : un code propre, qui permet à un neurone de communiquer son état propre indépendamment des autres neurones, et un code corrélationnel responsable d'une information de groupe codée par les synchronisations temporelles fines d'un ensemble de neurones.

Nous donnons ici une première approche de ces deux types de codage, de leur caractère complémentaire et partiellement indépendant. Pour cela, nous proposons des moyens de modéliser et de contrôler indépendamment le code intensitif et le code corrélationnel. Nous donnons aussi un aperçu des possibilités offertes par un outil particulièrement adapté à la gestion du code corrélationnel d'un réseau de neurones formels : la phase de la transformée de Fourier de leurs signaux.

Nous présentons ensuite un moyen d'utiliser l'information portée par le code corrélationnel en faisant appel à un réseau classique (Hopfield) que nous réadaptons à cet effet. Pour cela, les poids calculés par la règle de Hebb sont réinterprétés dans le domaine temporel, et le réseau est doté d'une dimension supplémentaire (le temps) sur laquelle l'information est portée et contrôlée.

Table des matières

Chapitre 1

Introduction

Ce stage s'inscrit dans le cadre très large des sciences cognitives. Notre but est de proposer des modèles inspirés par la physiologie, et fondés sur le principe du codage par corrélations temporelles des activités d'une population de neurones (i.e. le codage corrélationnel).

Toute la cognition naturelle reposant sur la capacité des neurones à traiter, transmettre, et donc coder l'information, les tentatives de décodage du « langage » neuronal constituent une des approches privilégiées de l'étude de la cognition. Ces neurones (de l'ordre de 10^{11} dans un cerveau humain) sont largement interconnectés, et la communication entre eux s'effectue le plus souvent selon le même principe : un neurone transmet des potentiels d'actions (ou PA) ponctuels (1 ms) et toujours identiques, à des intervalles de temps variables, vers tous les neurones avec lesquels il est connecté. Suivant la nature du neurone émetteur, le PA sera responsable au niveau des récepteurs d'un potentiel post-synaptique excitateur ou inhibiteur (PPSE et PPSI). Le neurone récepteur (ou post-synaptique) intègre tous les PPS en leur accordant des importances différentes et transmet le message en produisant lui-même un PA lorsque son seuil de dépolarisation est franchi.

La question du principe de ce code ponctuel binaire en tant que moyen de communication essentiel dans le système nerveux est à la base d'une grande partie des travaux de neurobiologie. Les enregistrements unitaires de neurones ont soulevé des questions quant à la signification du message neuronal, et beaucoup de scientifiques se sont lancés dans une tentative de décodage de ce signal. Dans un premier temps avec des enregistrements unitaires, puis en essayant d'acquérir une vision plus globale, notamment par des enregistrements de l'activité de groupes de neurones (enregistrements multi-unitaires, corrélogrammes...). Deux grands types de codage se dégagent de ces études : le codage par fréquence moyenne d'émission de PA par un neurone individuel, et le codage par la corrélation des activités d'un groupe de neurones. Après un survol bibliographique du codage neuronal, nous nous intéresserons à la modélisation et à l'utilisation du codage corrélationnel.

Chapitre 2

Présentation bibliographique du codage neuronal

2.1 Codage spatial et codage temporel

Les premières tentatives de traduction du « langage » neuronal se sont limitées à l'étude de l'intensité moyenne de décharge : dès 1928, Adrian met en évidence dans les systèmes sensoriels une relation précise entre fréquence moyenne de décharge de certains récepteurs et intensité de stimulation [1]. Ces récepteurs semblent obéir à des lois plus ou moins complexes, mais dont la forme reste $f = \phi(I)$ où f est la fréquence moyenne de décharge, ϕ la fonction de couplage et I l'intensité de stimulation.
Le succès de ce premier type d'interprétation de l'activité neuronale, ainsi que la conviction que l'information est codée par la seule intensité moyenne (i.e. que les trains de PA ne comportent aucune structure temporelle organisée) ont longtemps réduit l'attention portée aux enregistrements à leur intensité moyenne (Feldman, 1990 [2]), et les modèles ont été construits sans tenir compte des aspects temporels de la dynamique de décharge (Barlow, 1972 [3] ; Arbib, 1989 [4] ; Feldman, 1990 [2] ; Churchland, 1992 [5]). Ce codage par un processus de Poisson ponctuel d'espérance $f = \phi(I)$ est dit **codage spatial** car l'information est donnée par la distribution topographique des niveaux moyens d'activité.

Si le codage d'un scalaire par une série d'événements à valeurs binaires ne semble pas la stratégie la plus précise ni la plus rapide, il permet de porter une information contenue dans la distribution temporelle de ces événements : c'est le principe du **codage temporel**.

Du point de vue temporel, on peut d'abord s'intéresser au profil d'évolution de la fréquence de décharge dans le temps, mais cela ne prend toujours pas en compte la structure temporelle fine des événements. Outre la fréquence de décharge, un train de potentiels d'action présente une organisation dans le temps aussi potentiellement riche en information. Cariani [6] propose une

liste assez exhaustive des différents types de codage temporel envisageables (*cf.* figure 2.1) :
- codage par simple intervalle entre potentiels d'action,
- codage par intervalles multiples entre potentiels d'action,
- codage par forme temporelle complexe,
- codage par latence (temps d'arrivée du premier PA après un stimulus),
- codage par nombre de potentiels d'action par « burst » (série de courte durée de potentiels d'action à haute fréquence),
- codage par écarts entre « bursts ».

Au niveau de plusieurs neurones, cette organisation temporelle de la décharge individuelle autorise un type d'information supplémentaire porté par les corrélations temporelles des différentes activités : le **codage spatio-temporel**.

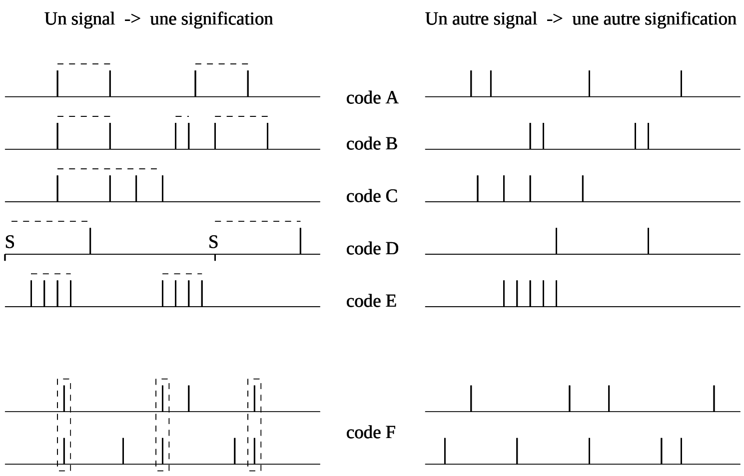

FIG. 2.1 - *(D'après Cariani, 1995) QUELQUES EXEMPLES DE CODES TEMPORELS: A: codage par simple intervalle entre potentiels d'action, B: codage par intervalles multiples, C: codage par forme temporelle complexe, D: Codage par latence, E: codage par nombre de PA par burst, F: codage par corrélations temporelles.*

2.2 Évidences expérimentales d'un code temporel

Des considérations anatomiques et physiologiques permettent déjà de mettre en doute l'aspect purement intensitif du codage de l'information corticale. D'une part l'intensité moyenne de décharge d'un neurone cortical est assez faible (moins de 100 PA par seconde), et une intégration temporelle précise demanderait d' «écouter» les neurones afférents suffisamment longtemps; ce qui est incompatible avec certains transferts rapides d'information corticale [7]. D'autre part, la faiblesse des poids des connexions entre neurones corticaux laisse prévoir l'importance de la synchronisation des décharges neuronales : un grand nombre de PA doivent arriver au même moment pour déclencher l'activité du neurone [8].

Les études électrophysiologiques permettent de visualiser l'organisation temporelle effective (« pattern » de décharge) dans l'activité de neurones corticaux voisins ou éloignés. Pour cela, l'activité doit être enregistrée puis représentée de façon à faire apparaître son organisation temporelle.

2.2.1 Représentations de la structure temporelle des trains de PA

Des enregistrements cellulaires à bonne résolution spatiale et temporelle sont indispensables à l'étude des patterns temporels de décharge. Pour une mise en évidence d'un code corrélationnel, il faut des enregistrements unicellulaires de plusieurs neurones simultanément. Des représentations de la dynamique des neurones permet ensuite de caractériser l'aspect temporel du code (les représentations sont visualisées dans les chapitres suivants).

Pour un seul neurone :

- Le PSTH (Post Stimulus Time Histogram) est une représentation de plusieurs enregistrements d'un même neurone répondant à un même stimulus. Cet histogramme donne le nombre de PA apparus dans chaque tranche de temps (dont la largeur détermine la précision de la représentation) après l'apparition du stimulus, après cumul de tous les enregistrements.
On obtient ainsi une estimation de la probabilité instantanée d'apparition d'un PA après le stimulus : il s'agit toujours d'une fréquence de décharge, mais la sommation de plusieurs trains permet d'acquérir une meilleure précision temporelle que celle donnée par un seul train. Cet histogramme permet d'objectiver le décours temporel de l'activité neuronale après une stimulation.

- L'histogramme d'intervalles représente l'organisation temporelle des PA. L'histogramme d'intervalles d'ordre n donne le nombre d'intervalles dans chaque tranche de longueur entre deux PA séparés par $(n-1)$ PA.

- L' ACH (autocorrélogramme) est la somme des histogrammes d'intervalle de tous les ordres, c'est à dire du nombre de PA qui se trouvent dans une tranche de temps avant ou après n'importe quel PA.

Pour plusieurs neurones :

- Le CCH (cross-corrélogramme) entre 2 trains A et B est basé sur le même principe que l'ACH en représentant le nombre de PA du train B qui se trouvent dans une tranche de temps après n'importe quel PA du train A. Un

pic en 0 représente donc une activité synchronisée tandis qu'un pic décalé par rapport à l'origine révèle une activité séquentielle.
Lorsque les réponses neuronales ne sont pas stationnaires (fréquence moyenne constante), ce qui est souvent le cas lorsque la réponse est induite par un stimulus, on risque d'observer une corrélation due aux co-variations des deux réponses induites par le stimulus, et non à une interaction neuronale... Pour éliminer la composante due au stimulus, Gerstein et Perkel proposent la méthode du «shift predictor» qui consiste à utiliser un stimulus périodique pour retrancher le CCH décalé d'une période au pic correspondant à l'apparition du stimulus . Cette méthode considère qu'il n'y a plus d'interaction entre neurone une période après le début du stimulus [9].

- Le PSCCH permet de visualiser la variation du couplage de deux neurones A et B en fonction du temps écoulé depuis le dernier stimulus. Il s'agit de porter sur un axe l'intervalle de temps entre un PA de A et le début du dernier stimulus, et sur l'autre axe l'intervalle de temps entre un PA de A et un PA de B (CCH) [10].

2.2.2 Résultats

Perkel et son équipe commencent à mettre à jour et à caractériser des patterns temporels de décharge neuronale vers 1967 [11] [12].

Lestienne observe des structures temporelles répétées (triplets) dans les trains d'impulsions évoqués dans le cortex visuel primaire du chat et du singe [13].

Gerstein [14] met en évidence les corrélations temporelles de plusieurs neurones dès les années 1970. Par la suite, des recherches systématiques de patterns spatio-temporels sont effectuées dans le cortex du singe par Abeles et Gerstein [15] [16]. Leurs études statistiques permettent d'éliminer l'hypothèse de corrélations fortuites.

L'organisation temporelle de certains groupes de neurones devient indéniable. Reste à vérifier que ces corrélations sont porteuses d'un message. Pour cela, Krüger et Becker [17] enregistrent les réponses de 24 neurones de l'aire visuelle 17 à la présentation de 16 stimuli différents. Ils cherchent ensuite à retrouver le stimulus présenté à partir des seules réponses enregistrées. Ceci n'est possible que lorsque l'on utilise l'organisation temporelle des enregistrements : l'intensité moyenne est insuffisante.

L'équipe de Vaadia [18] enregistre simultanément des paires de neurones du cortex moteur pré-frontal du singe lors d'une tâche de réponse motrice retardée : de la condition «mouvement» à la condition «pas de mouvement» le taux d'activité globale ne change pas, mais la taux de corrélation entre les paires de neurones varie significativement.

La structure temporelle des PA semble donc contenir une information pertinente.

2.3 Les modèles du codage temporel

Les études neurophysiologiques ont révélé une organisation spatio-temporelle de l'activité des neurones et la possibilité qu'elle soit le support d'un codage neuronal. Les modélisations commencent à intégrer ce codage temporel aux réseaux de neurones formels.

2.3.1 Synchronisation de réseaux d'oscillateurs

Une première approche pour représenter l'aspect temporel de la décharge des neurones est de supposer qu'ils ont une activité oscillante autour de leur intensité moyenne : au paramètre intensité moyenne I, on ajoute une profondeur de modulation m et deux paramètres temporels que sont la période ω et la phase ϕ des oscillations. L'activité du neurone est alors de la forme :

$$S(t) = I\Big(1 + m \cdot \sin(\omega t + \phi)\Big)$$

Ce comportement oscillant est parfois observé par les neurophysiologistes, et permet aux modélisateurs de gérer simplement les corrélations entre neurones en contrôlant les paramètres temporels des oscillations.
Quel que soit le modèle de décharge périodique du neurone isolé (des simples sinusoïdes bruitées aux réalisations plus biomimétiques), plusieurs auteurs reproduisent les corrélations temporelles accompagnées d'oscillations.

2.3.2 Synchronisation de trains de potentiels d'action

Bien que cette activité oscillante ait été enregistrée pour quelques neurones corticaux, ce n'est pas un comportement général. Les neurones à activité apériodique peuvent aussi se synchroniser, et les modèles ont tenté d'intégrer la paramètre temps à l'activité de leurs neurones sans utiliser d'oscillations, mais des neurones plus biomimétiques dotés d'un potentiel de membrane qui évolue au cours du temps, intègre les afférences (entrées), et déclenche un PA au dessus d'un seuil critique.
On peut modéliser la décharge de potentiels d'actions par des neurones formels en simulant tous les flux d'ions transmembranaires qui ont une influence sur le potentiel membranaire (équations de Hodgkin-Huxley). On peut aussi simplifier le comportement membranaire en « integrate and fire ». Ce type de modèle a été proposé par French et Stein [19]. Ils considèrent que sans afférence la différence du potentiel membranaire au potentiel de repos tend à s'annuler au cours du temps selon :

$$h(t) = \frac{e^{-t/\tau}}{\tau}; \tag{2.1}$$

et que le potentiel membranaire est modifié par chaque afférence $a(t)$ selon :

$$\Delta v(t) = a(t) * h(t) \tag{2.2}$$

où $*$ représente le produit de convolution (*cf.* annexe A). L'afférence peut être due à une stimulation extérieure, à une connexion interne au réseau, ou à un bruit (fluctuations aléatoires).
Lorsque le seuil de dépolarisation est franchi, le neurone émet un potentiel d'action puis reprend son potentiel de repos. Il peut alors être contraint à garder ce potentiel pendant une durée fixe après une émission : la période réfractaire.
Lorsque de tels neurones ne sont pas entraînés par des sources périodiques, leur activité reste apériodique. Si on ajoute une propriété de réfractorité, elle peut devenir périodique ; de plus, si on les connectecte par des liens excitateurs ou inhibiteurs, une activité périodique et synchronisée peut apparaître sous certaines conditions.

Par exemple, Bush et Douglas [20] modélisent par des équations très proches de celles de Hodgkin-Huxley la synchronisation des bursts des neurones corticaux dus à un stimulus visuel cohérent. Ils utilisent des interconnexions excitatrices locales pour amplifier les entrées géniculées et produire des décharges fortes et cohérentes des cellules pyramidales ; et des inhibitions globales de type feedback pour introduire un signal de reset et synchroniser le burst suivant.

Tous ces modèles permettent d'observer une organisation spatio-temporelle de l'activité du réseau, mais leurs mécanismes de synchronisation sont basés sur des connexions classiques, c'est-à-dire modulant uniquement l'intensité de leurs afférences, et non leur aspect temporel. Ces connexions n'étant pas forcement bien adaptées à l'utilisation de l'information spatio-temporelle, de nouveaux réseaux cherchent à rendre la notion de temps moins artificielle dans les connexions.

- On peut d'abord introduire un délai dans chaque connexion, qui correspondrait au temps de propagation de l'information d'un neurone à l'autre.
Ainsi, Amit [21] dote un réseau de type Hopfield de liens à retard qui s'ajoutent aux liens classiques pour retarder la transmission de l'information et pour rendre le réseau capable d'enchaîner les formes apprises.
D'autre part, une étude probabiliste de réseaux de Hopfield a été proposée dans le cas où les neurones ont une activité stochastique dépendant du potentiel de membrane évalué selon la règle classique [22].
Nischwitz et Glünder [23] montrent que dans une population de neurones « leaky integrate and fire », les connexions les plus aptes à induire une synchronisation sont les inhibitions latérales retardées.

- On peut aussi représenter les interconnexions par des lignes à retard : les impulsions d'entrée se distribuent sur un arbre dendritique dont chaque branche se comporte comme un câble [24].

- Le «co-reset» est une nouvelle approche des connexions entre neurones qui permet de synchroniser leurs activités : le co-reset consiste à munir une population de neurones de la propriété de réfractorité et de contraindre tous les neurones connectés par des liens de co-reset à entrer en période réfractaire dès que l'un d'entre eux émet un PA. Ce comportement permet de ne laisser passer en sortie que les PA synchronisés (Berthommier, 1995 [25]).

2.4 Utilité du code temporel : le «binding»

2.4.1 Théorie du «binding»

Un objet qui nous apparaît comme une entité est en fait formé de plusieurs composantes perceptives. Il ne peut être caractérisé par aucune de ces composantes prises isolément, mais seulement par leur association en un tout. Par exemple, si plusieurs objets sont présentés dans notre champ visuel, chacun ayant une forme, une couleur, et une direction de mouvement différentes, nous sommes capables d'associer la forme, la couleur, et la direction de chacun des objets. Nous réalisons donc en plus du codage des composantes, un codage de leurs liens : le «binding». L'opération de «binding» (ou liage des composantes) correspond à une association perceptuelle dont les bases physiologiques ne sont pas encore élucidées. La question du mécanisme neuronal soutenant le binding se pose immédiatement, d'autant plus que les différentes caractéristiques perceptives sont souvent codées dans des régions éloignées du cortex. Une des hypothèses actuelles considère que le codage des liens entre composantes d'un même objet perceptif se fait par synchronisation neuronale temporaire, ce qui autorise un même neurone à participer au codage de différents objets (Pour une modélisation du binding, *cf.* von der Malsburg et Schneider [26] ; pour une approche expérimentale, *cf.* Singer [27]).

2.4.2 Confirmations expérimentales

Engel a étudié des neurones du cortex visuel du chat qui répondent à un même stimulus et a mis en évidence des réponses synchronisées sur des intervalles de temps réguliers, que les neurones soient voisins ou non [28].
Gray a mené la même étude sur des neurones à champs récepteurs différents : la synchronisation n'apparaît que si les neurones répondent à un même objet traversant simultanément les deux champs récepteurs de façon cohérente [29]. Plus récemment, une synchronisation a aussi été mise en évidence entre neurones situés dans des aires visuelles différentes (17 et 18) du cortex du chat [30]. Contrairement à toute attente, les pics des CCH ne sont pas décalés, mais symétriques par rapport à l'origine (*cf.* méthodes), ce qui signifie que l'activation des neurones de différentes aires corticales n'est pas séquentielle mais synchrone . Il semble donc que les différentes aires corticales fonc-

tionnent en parallèle pour représenter les attribut d'un même objet, et que leur unité est codée par des corrélations temporelles.

2.5 Présentation du stage

2.5.1 Code propre et code corrélationnel

Dans ce travail, notre approche du traitement de l'information par les neurones au sein d'un réseau est basée sur une hypothèse de «double codage». Nous supposons que le code spatio-temporel présent dans les signaux d'un ensemble de canaux porte deux types d'informations à priori indépendantes:

- le **code propre**, qui délivre une information pouvant être interprétée indépendamment dans chaque canal, codant par exemple une des caractéristiques d'un objet. Nous réduirons ici le code propre à la fréquence moyenne de décharge.

- le **code corrélationnel**, responsable d'une information portée en synergie avec les autres neurones associés par synchronisation temporelle de leurs activités. Il s'agit du «code de groupe» des assemblées neuronales. Le code corrélationnel sera manipulé dans sa dimension temporelle.

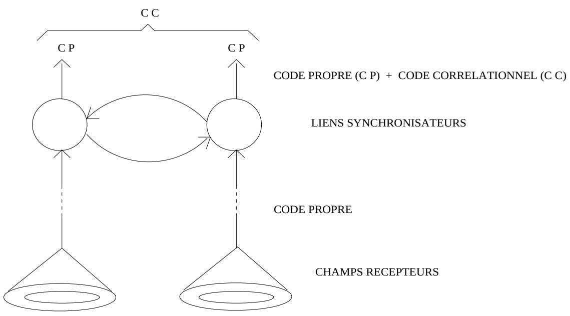

FIG. 2.2 - *CODE PROPRE ET CODE CORRÉLATIONNEL: Le code propre est formé au niveau des neurones récepteurs qui assurent l'analyse de la stimulation en terme de caractéristiques. Le code corrélationnel est ensuite construit grâce à une interaction entre des neurones de plus haut niveau. Ici, nous supposons que cette interaction est basée sur des liens synchronisateurs. Ces liens serviraient à produire une association entre les caractéristiques d'un même objet.*

Pour que ces deux codes soient efficaces, ils doivent être le moins redondants possible, et pouvoir varier indépendamment. Nous donnerons quelques moyens pour manipuler le code corrélationnel d'une population de neurones formels sans altérer le code propre. Nous considérerons le signal émis par un neurone comme la discrétisation dans le temps d'une fonction à valeurs binaires (train de PA) ou d'une fonction à valeurs continues (probabilité d'émission d'un PA au cours du temps). Dans ce deuxième cas, nous approcherons le problème en manipulant des signaux sous forme de bruit blanc.

2.5.2 Utilisation du code corrélationnel

Dans un deuxième temps, nous montrerons comment appliquer la notion de codage corrélationnel. Si beaucoup de modèles proposent des mécanismes de synchronisation de l'activité des neurones, peu d'entre eux utilisent par la suite cette information de façon pertinente. Les réseaux fonctionnent en effet dans leur grande majorité avec la seule information intensitive. Nous partons d'un modèle de réseau classique (réseau de Hopfield), que nous modifions de façon appropriée afin d'exhiber une propriété de contrôle des corrélations temporelles. Sachant que le réseau de Hopfield intègre une information purement spatiale, nous dotons ses unités d'une propriété de manipulation des signaux dans le domaine temporel. Cette propriété repose sur l'introduction de fonctions de transfert temporelles qui se substituent aux poids de connexion classiques. Les valeurs instantanées sont ainsi remplacées par des distributions temporelles d'activité. Biologiquement, ces distributions temporelles correspondent à l'évolution du potentiel membranaire post-synaptique produit par l'arrivée d'un PA.

Chapitre 3

Code propre et code corrélationnel : contrôle indépendant

Cette première partie cherche à mettre en évidence l'indépendance du code propre et du code corrélationnel, et par conséquent la possibilité de les manipuler séparément. Nous présentons ici quelques moyens d'agir sur le code corrélationnel d'une population de neurones formels sans modifier le code propre de chacun d'entre eux. Ceci permet d'ajouter une information de groupe sans perdre les informations individuelles.

3.1 Synchronisation de trains de PA

Les trains de PA étant composés d'une succession dans le temps d'informations binaires, ils se prêtent facilement aux manipulations temporelles (suppressions, ajouts, décalages). On peut imaginer trois principes de réalignement des PA : l'élimination des PA non synchronisés, l'addition de PA synchronisés, et le décalage des PA pour les réaligner. En assimilant code propre et intensité moyenne de décharge sur un intervalle de temps T, le code propre est conservé à partir du moment où l'on ne modifie pas le nombre de PA dans une fenêtre temporelle de longueur inférieure à T. Parmi les différentes procédures de réalignement possibles, seul le décalage conserve le code propre ; nous avons donc choisi pour réaliser les deux procédures suivantes de retarder l'émission de PA d'un train pour les synchroniser sur un autre train (*cf.* figure 3.1).

3.1.1 Synchronisation par réalignements locaux

À partir d'une matrice de couplage $w = (w_{ij})_{(i,j)\in[1,N]^2}$, où w_{ij} représente le lien synchronisateur du neurone J vers le neurone I, N trains de PA

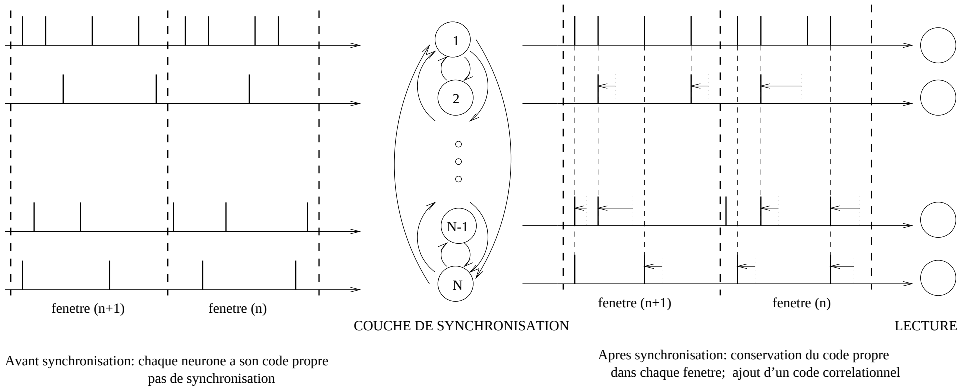

FIG. 3.1 - *MISE EN PLACE D'UN CODE CORRÉLATIONNEL : Exemple de synchronisation de N neurones sur le neurone numéro 1 par retard de PA. Le nombre de PA par neurone par fenêtre (code propre) est conservé, et les émissions sont retardées pour s'aligner sur l'émission du neurone 1. Seuls certains liens sont représentés, mais le réseau peut être connecté selon la matrice des liens de synchronisation.*

sont synchronisés tout en conservant leur code propre sur tout intervalle de longueur T selon la procédure suivante :

```
Choix aléatoire du premier neurone << synchronisateur >> : NS
Dans chaque fenêtre de temps de longueur T :
  recherche de la position du dernier PA de NS dans la fenêtre
  si ce PA existe :
    pour chaque neurone M différent de NS :
      recherche du PA de M précédant le dernier PA de NS
      si ce PA existe :
        réalignement de ce PA sur NS avec une probabilité W(M,NS)
  si au moins un réalignement est effectué :
    choix de la fenêtre et du neurone synchronisateur suivants
  si pas de réalignement effectué :
    choix du neurone synchronisateur suivant sur la même fenêtre
  si pas de réalignement et tous les neurones testés :
    passage à la fenêtre suivante
Jusqu'à la dernière fenêtre de temps
```

3.1.2 Synchronisation par rapprochements itératifs

Le principe est le même, à la différence près que le coefficient w_{ij} de la matrice de couplage représente ici le gain de réalignement : si les deux PA à réaligner sont décalés d'un délai D alors le PA émis par I est retardé de $w_{ij} \cdot D$. Le rapprochement des PA émis par deux neurones est ainsi plus progressif.

3.1.3 Lecture de la synchronisation

Une couche de lecture permettra de visualiser la synchronisation moyenne de chaque neurone avec les neurones d'un voisinage V plus ou moins large (V peut même représenter tout le réseau). Pour cela, chaque neurone somme dans une fenêtre T de temps les PA de ses voisins alignés plus ou moins précisément sur les siens :

$$L_i = \sum_{t \in T} (s'_i(t) \cdot \sum_{j \in V} s_j(t))$$

où $s'_i = s_i * g$ et g est la fenêtre qui représente la précision temporelle désirée : Dirac pour la lecture des alignements parfaits, rectangle ou gaussienne pour une lecture des alignements moins précise dans le temps. Avec un seuillage approprié de la valeur continue L_i, nous retrouvons une représentation spatiale binaire des neurones corrélés.

3.1.4 Simulations du réalignement

La figure 3.2 représente les trains de PA de quatre neurones avant (itération 0) et pendant (itérations 1 et 5) la synchronisation par réalignements. Les liens de la matrice de couplage sont nuls sauf $w_{2,1}$, $w_{3,4}$, et $w_{4,3}$. On observe que les PA du train 2 se réalignent progressivement sur ceux du train 1 (non modifiés puisqu'il n'existe aucun lien synchronisateur vers le neurone 1), tandis que les trains 3 et 4 se réalignent l'un sur l'autre indépendamment de 1 et 2.

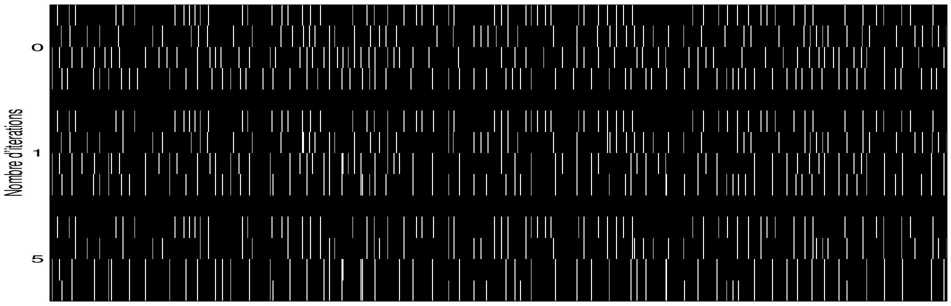

FIG. 3.2 - *SYNCHRONISATION PAR RÉALIGNEMENTS LOCAUX: On part de 4 trains de PA non synchronisés et d'une matrice de synchronisation w nulle sauf pour $w_{2,1}$, $w_{3,4}$, et $w_{4,3}$. On observe un réalignement du neurone 2 sur le neurone 1, et des neurones 3 et 4 l'un sur l'autre.*

Les corrélations temporelles de trains de PA sont plus faciles à lire sur des CCH (*cf.* méthodes). La figure 3.3 représente les corrélations des paires

1x2, 3x4, et 1x3 avant et après synchronisation. Les pics en 0 objectivent la synchronisation des neurones 1 et 2 d'une part et 3 et 4 de l'autre.

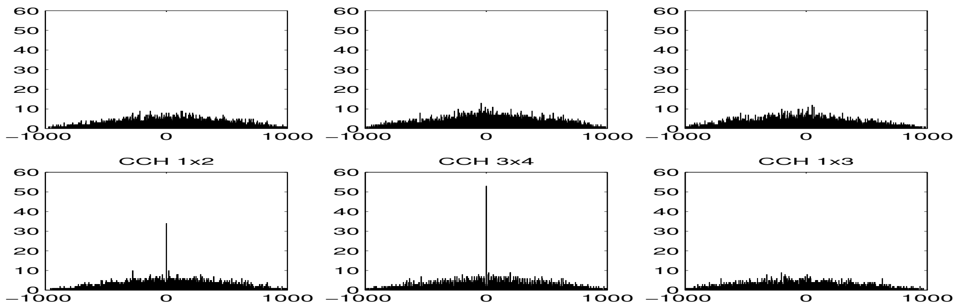

FIG. 3.3 - *CCH AVANT ET APRÈS SYNCHRONISATION : La matrice de poids est la même que celle de la figure 3.2. Les CCH 1*2, 3*4, et pour contrôle 1*3 sont donnés avant (ligne 1) et après (ligne 2) 5 itérations.*

3.1.5 Diffusion de la synchronisation

Lorsque les liens synchronisateurs sont limités à des liens locaux, on peut observer une diffusion de la synchronisation. On remarque alors que l'intensité moyenne de chaque neurone va influer sur sa capacité à transmettre la synchronisation. Un neurone à faible fréquence d'émission va pouvoir être synchronisé par diffusion, mais ne va pas pouvoir transmettre efficacement le message corrélationnel aux neurones voisins, même si leur fréquence est élevée. En effet, ces neurones ne pourront être corrélés à ceux de départ que par le petit nombre de PA du neurone intermédiaire. Nous pouvons par exemple limiter la diffusion de la synchronisation à une région du réseau : la figure 3.4 représente la diffusion de la synchronisation dans un réseau linéaire de 15 neurones, le couplage étant limité aux deux voisins les plus proches. Les cinq neurones centraux sont synchronisés au départ, puis la synchronisation se diffuse jusqu'aux barrières de faible activité au cours des itérations.

Ceci permet d'envisager un contrôle spatial du codage corrélationnel par l'intensité moyenne : une barrière de diffusion de la synchronisation est ici définie en entrée par le code propre de l'assemblée neuronale. Dans un système biologique, des bornes de synchronisation peuvent être construites en modulant la distribution spatiale des entrées par un filtrage spatial (inhibitions latérales par exemple).

3.2 Synchronisation de signaux à valeurs réelles

Si nous ne considérons plus l'activité d'un neurone mais d'une population assez grande, nous pouvons modéliser le signal par l'évolution dans le temps

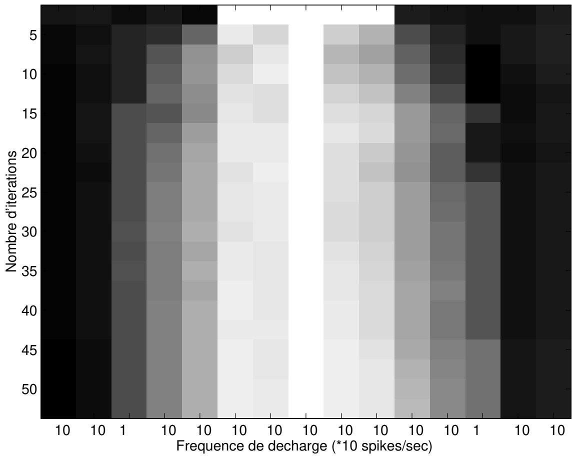

FIG. 3.4 - *DIFFUSION DE LA SYNCHRONISATION LOCALE: Visualisation du pourcentage de corrélation avec le neurone central (100% de corrélation en blanc à 0% de corrélation en noir). Lorsque la synchronisation se fait de proche en proche, sa diffusion est limitée par les neurones les moins actifs (10 PA/sec).*

de la probabilité instantanée d'émission d'un PA. Ce signal à valeurs continues réelles se prête à un autre type de manipulation, qui permet d'obtenir une fonction équivalente au réalignement de trains de PA : le rephasage.

3.2.1 Principe du rephasage

Tout signal périodique de période T (et par conséquent tout signal fini) peut s'écrire sous la forme d'une somme infinie de sinusoïdes de fréquences multiples de 1/T, chacune affectée d'une amplitude et d'une phase appropriées :

$$\forall x \text{ périodique}, \exists!(X_n, \Phi_n)_{n\in\mathbb{N}} / x(t) = \sum_{n=0}^{\infty} X_n.cos(2\pi\frac{n}{T}t - \phi_n)$$

La transformée de Fourier (TF) permet ainsi de décomposer un signal temporel $x(t)$ en ses composantes (X_n) et (Φ_n), qui représentent respectivement l'amplitude et la phase des sinusoïdes de fréquence $2\pi n/T$. Pour plus de détails mathématiques sur la TF, se reporter à l'annexe 1.
Manipuler le spectre de phase de la TF permet de décaler quelques composantes sinusoïdales dans le temps sans modifier leurs amplitudes : la phase de la transformée de Fourier est donc un outil parfaitement adapté à la synchronisation de signaux en conservant leurs intensités moyennes, donc à l'addition d'un code corrélationnel au code propre. Par exemple, en prenant le cas extrême de deux signaux $n1$ et $n2$ sinusoïdaux purs, on a une décorrélation maximale pour $\Phi_{n1} = \Phi_{n2} + \pi$, et on peut obtenir une corrélation maximale en copiant pour $n2$ la phase de $n1$ (*cf.* figure 3.5).

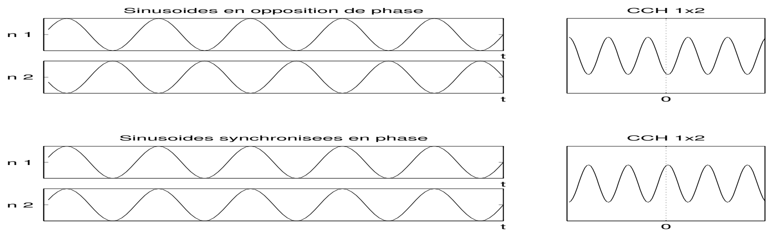

FIG. 3.5 - *MODIFICATION DU CCH DE DEUX SIGNAUX PÉRIODIQUES EN FONCTION DE LEUR DIFFÉRENCE DE PHASE*

3.2.2 Rephasage par copie partielle de phase

Le principe est toujours de partir de N signaux quelconques non corrélés, et d'une matrice de couplage w. Le neurone I se synchronise sur le neurone J en copiant une portion w_{ij} aléatoire de la phase de sa TF (*cf.* figure 3.6).

```
Calcul des TF de tous les signaux
Pour chaque neurone à rephaser NR :
  pour chaque neurone M
    copie d'une proportion w(NR,M) de valeurs du vecteur phase
                                               de M sur celui de N
Calcul des TFI à partir des phases ainsi modifiées
```

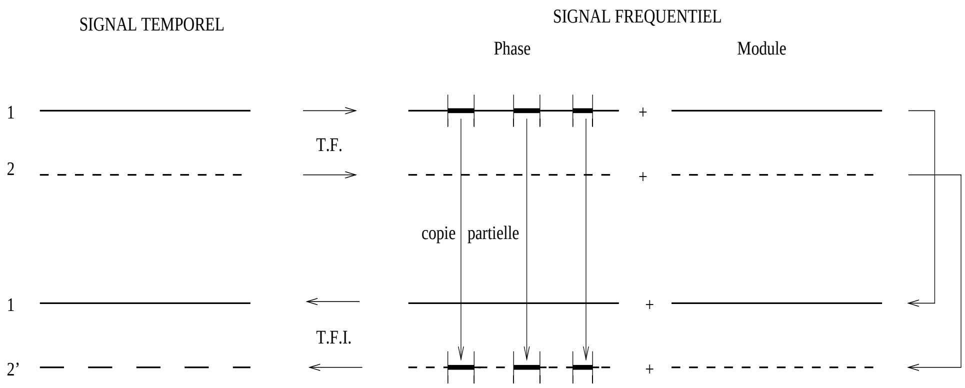

FIG. 3.6 - *SYNCHRONISATION PAR COPIE DE PHASE : exemple de synchronisation du signal 2 sur le signal 1 par copie d'une portion $w_{2,1}$ de la phase de sa TF (TF = Transformée de Fourier; TFI = Transformée de Fourier Inverse).*

3.2.3 Rephasage itératif

En considérant la phase comme un vecteur de dimension la longueur l du signal (donc appartenant à $[-\pi/2, \pi/2]^l$), on peut synchroniser le signal I sur le neurone J en modifiant son vecteur phase d'une proportion w_{ij} de l'angle qui les sépare.

3.2.4 Résultats préliminaires

Les deux méthodes permettent d'observer rapidement une synchronisation des neurones reliés par des liens de synchronisation non nuls (*cf.* figure 3.7). La copie aléatoire partielle de phase donne un résultat moins régulier et un peu moins efficace que le rapprochement car les copies de portions de phase communes sont possibles, et de plus en plus fréquentes : le gain est altéré à cause du recouvrement. Le rapprochement des angles des vecteurs phase assure une convergence continue de la forme : $\Delta\theta_{ij}(n) = (1 - w_{ij})^n.\Delta\theta_{ij}(0)$ où $\Delta\theta_{ij}(n)$ est l'angle entre les vecteurs phase de J et de I à l'itération n, et w_{ij} le lien de synchronisation de I sur J.
Si on interdit la modification des fréquences très faibles, on vérifie aussi que l'intensité moyenne instantanée du signal est conservée. La manipulation de la phase de la TF est donc un outil mathématique simple et bien adapté au problème.

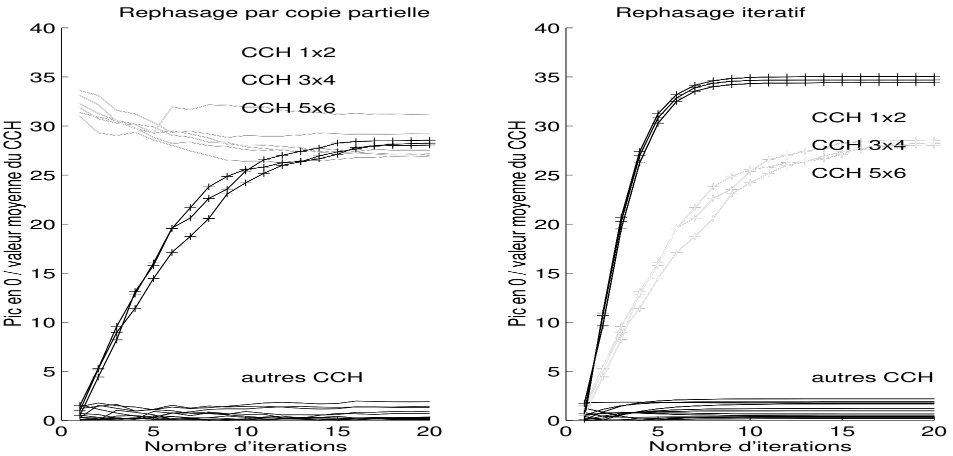

FIG. 3.7 - *ÉVOLUTION DES CCH AU COURS DE LA SYNCHRONISATION : Évolution du rapport pic en 0 sur valeur moyenne du CCH de toutes les paires d'un réseau de 6 neurones synchronisés par une matrice* $w = \begin{pmatrix} w_1 & w_0 & w_0 \\ w_0 & w_1 & w_0 \\ w_0 & w_0 & w_1 \end{pmatrix}$, *où* $w_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, *et* $w_0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

3.2.5 L'intérêt de la phase

Manipuler la phase de la TF d'un ensemble de signaux permet dans un premier temps d'éviter la transitivité de la synchronisation.
Nous avons observé la diffusion de la synchronisation des trains de PA avec des liens locaux : cela signifie que si $w_{ab} > 0$ et $w_{bc} > 0$ alors A et C se synchroniseront malgré $w_{ac} = 0$ lorsque augmente le nombre d'itérations. En travaillant dans le domaine fréquentiel, on peut contrôler la transitivité en limitant le rephasage à certaines bandes de fréquences. Corréler les différentes paires de neurones sur des bandes de fréquences différentes permet en effet d'éviter la diffusion de la synchronisation quel que soit le nombre d'itérations (*cf* figure 3.8).
Cette propriété est indispensable à la participation de quelques neurones à deux groupes de synchronisation indépendants sans que cela puisse synchroniser les deux populations.

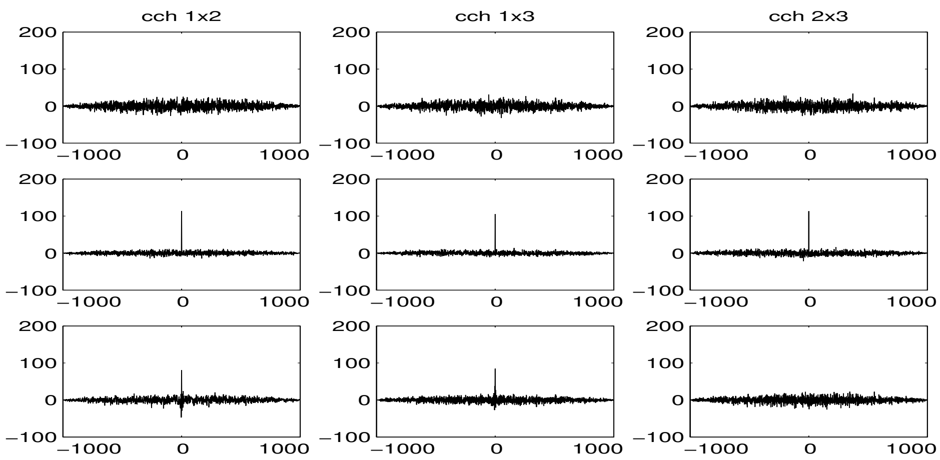

FIG. 3.8 - *INTÉRÊT DE LA SYNCHRONISATION SUR DES BANDES DE FRÉQUENCES : Trois neurones indépendants (ligne 1) sont synchronisés selon* $w = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$*, par 10 itérations sur l'ensemble des fréquences (ligne 2), ou sur des bandes limitées et indépendantes de fréquences (ligne 3). On vérifie la disparition du caractère transitif de la synchronisation dans le deuxième cas.*

La manipulation des transformées de Fourier permet d'autre part d'expliquer différents comportements observables au niveau des ACH et des CCH par les propriétés fréquentielles des signaux.

Ainsi, copier une portion de phase permet, nous l'avons vu, d'obtenir un pic en 0 sur le CCH, et donc de synchroniser les signaux.

Une deuxième manipulation sur le module de la TF permet de jouer sur le caractère oscillant de l'ACH et du CCH. En effet, si on réduit le module de la TF d'un signal à une bande de fréquences étroite, le signal se rapproche d'une sinusoïde, et l'ACH devient oscillant. Le CCH, quant à lui, devient oscillant lorsqu'on réduit la bande de fréquence commune aux deux signaux. Ces deux manipulations sont indépendantes: si on cale deux neurones sur une bande étroite de fréquences communes, le CCH va osciller, et si les fréquences communes sont en phase, il va présenter un pic en 0 (*cf.* figure 3.9). La figure 3.10 donne les corrélogrammes obtenus lors de la simulation : l'oscillation de l'enveloppe dépend directement de la largeur de bande (plus la bande est étroite, plus la fréquence d'oscillation est forte).En agrandissant le signal (*cf.* figure 3.11), on se rend compte que l'enveloppe est en fait la modulation d'un signal déjà périodique de fréquence plus grande : cette fréquence correspond à la moyenne de la bande de fréquences utilisée.

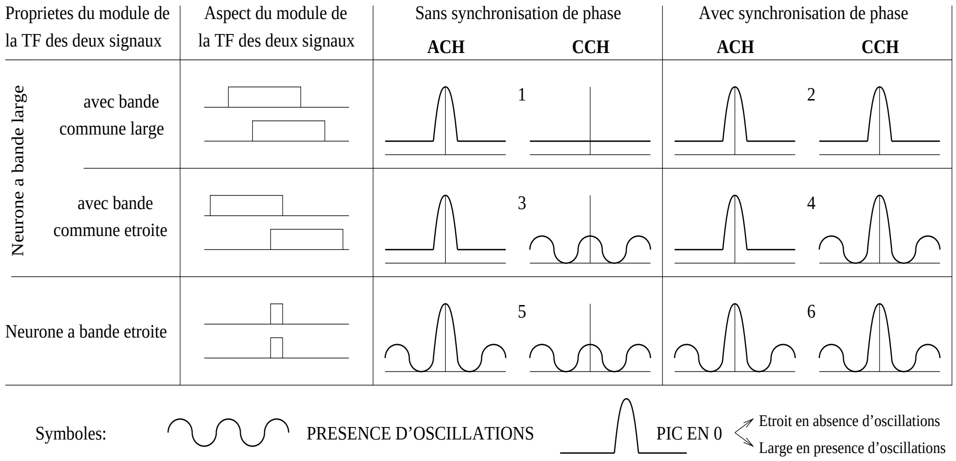

FIG. 3.9 - *RÉSULTATS DE DEUX MANIPULATIONS PAR LA TF : Le passage d'une ligne à l'autre se fait par manipulation du module de la TF, et le passage d'une colonne à l'autre se fait par manipulation de phase. L'étroitesse de la bande de fréquence d'un signal est responsable de l'oscillation de son ACH, l'étroitesse de la bande de fréquence commune à deux signaux est responsable de l'oscillation de leur CCH, et leur synchronisation par remise en phase est responsable du pic en 0 de leur CCH.*

3.3 La synchronisation en terme d'entropie

Il serait intéressant de pouvoir caractériser globalement le « degré de synchronisation » d'une population de neurones sans devoir passer par l'étude de tous les couples de neurones (CCH), ni même de tous les neurones (synchronisation moyenne d'un neurone avec la population). Il s'agirait donc de pouvoir évaluer la distribution temporelle de l'activité globale de la population selon un critère de synchronisation. C'est-à-dire qu'une distribution concentrée dans le temps se verrait attribuer un meilleur indice qu'une distribution étalée dans le temps. Si on considère la probabilité en chaque instant

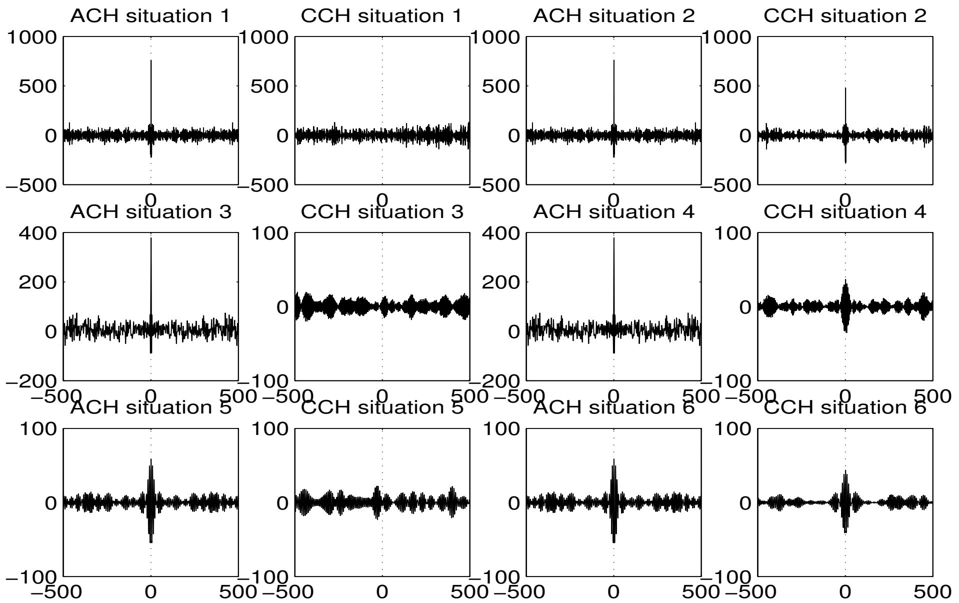

FIG. 3.10 - *VISUALISATION DES CORRÉLOGRAMMES DE LA FIG 3.9 :Résultats obtenus pour des signaux de durée 1 s, et de résolution temporelle 1 ms. La ligne 1 est obtenue pour des neurones à bande de fréquence variant sur 400 Hz avec une bande commune de 300 Hz, la ligne 2 pour des neurones à bandes variant sur 200 Hz et avec une bande commune de 30 Hz,et la ligne 3 pour des neurones à bandes communes de 30 Hz.*

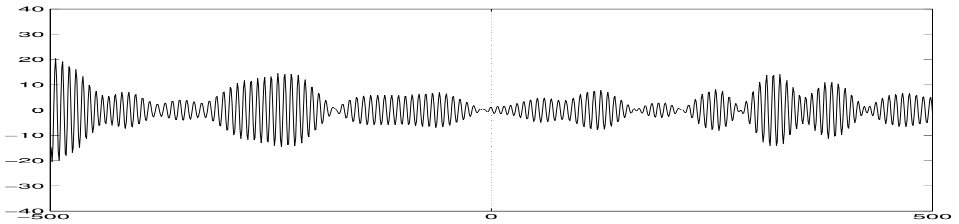

FIG. 3.11 - *AGRANDISSEMENT D'UN CCH (situation 3) : La structure temporelle fine (porteuse) d'un CCH dépend de la fréquence centrale de la bande commune.*

pour un des neurones de la population d'émette un PA, une fonction très adaptée à la mesure de son organisation temporelle est l'entropie H :

$$H = -\sum_i p_i \cdot log(p_i)$$

où p_i est la probabilité d'apparition d'un PA à l'instant i. L'entropie représente en effet une mesure de l'indétermination globale d'une expérience [31]. Plus les probabilités sont uniformes, plus l'indétermination est grande, et plus elles sont concentrées dans le temps, plus l'indétermination est faible. La figure 3.12 donne deux exemples d'évolution de l'« entropie » d'une population de neurones au cours de leur synchronisation par réalignement ou par rephasage. L'entropie diminue de façon plus importante pour la synchronisation de PA car la distribution se rapproche d'une succession de Diracs, tandis que la synchronisation de bruits blancs ne peut pas conduire à une aussi forte concentration temporelle.

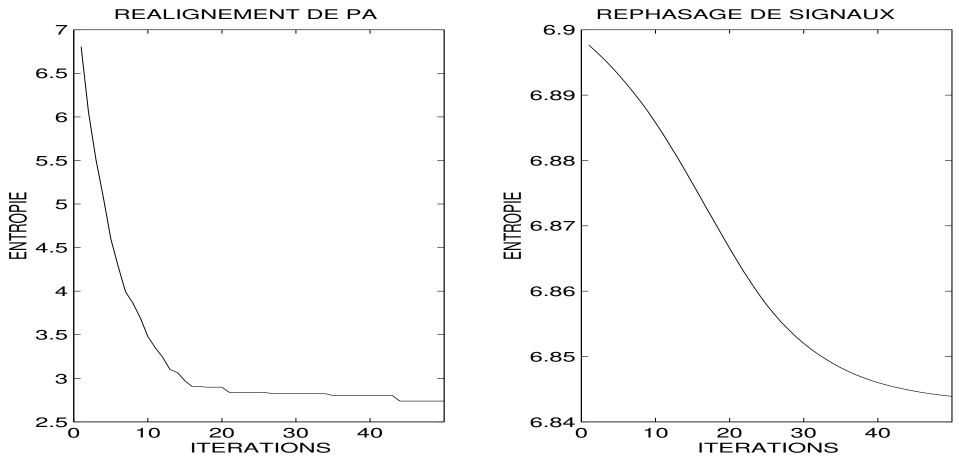

FIG. 3.12 - *EXEMPLE D'ÉVOLUTION DE L'ENTROPIE AU COURS DE LA SYNCHRONISATION : une centaine de neurones sont synchronisés entre eux par réalignement de PA (courbe de gauche) ou par rephasage (courbe de droite).*

Chapitre 4

Utilisation du code corrélationnel par un réseau classique

Nous avons vu qu'il existe des outils qui permettent de synchroniser des neurones formels de manière ad hoc, et d'autres qui permettent de lire cette information. Reste à pouvoir utiliser ce nouveau code, comme les réseaux classiques de catégorisation utilisent le code intensitif. Nous proposons ici une adaptation du réseau de Hopfield pour tester la possibilité d'utiliser le code corrélationnel transmis par l'intermédiaire de trains de PA plus ou moins corrélés. Ici, l'entrée et la sortie sont des trains de PA, et la fonction du réseau est d'agir sur la synchronisation de ces trains.

4.1 Le réseau de Hopfield

4.1.1 Description du réseau

En 1982, Hopfield [32] propose un modèle de mémoire associative basé sur le comportement d'un ensemble de N neurones formels complètement interconnectés par une matrice w de poids symétrique. L'utilisation du réseau s'effectue en deux étapes:
1) La phase d'apprentissage consiste à calculer w pour mémoriser un nombre P d'exemples d'états du réseau $(\vec{\xi}^\mu)_{\mu\in[1,P]}$ où $\vec{\xi}^\mu = (\xi_1^\mu ... \xi_N^\mu)$ et ξ_i^μ est l'état du neurone i dans l'exemple μ. Les poids se calculent par la règle de Hebb selon :

$$w_{ij} = \sum_{\mu=1}^{P} \xi_i^\mu . \xi_j^\mu$$

Nous avons utilisé un algorithme qui calcule N poids supplémentaires (les biais) qui correspondraient aux connexions d'un neurone fictif (toujours

d'état -1) avec les N autres. L'entrée de chacun des neurones étant toujours modifiée d'une même valeur, cela reviendrait à déplacer le seuil de sa fonction neurone $\mathcal{N}$.
2) Durant la phase dynamique, le réseau évolue d'un état donné par l'entrée vers l'état attracteur le plus proche selon la règle de transition :

$$\text{pour tout neurone } i, \;\; s_i(t+1) = \mathcal{N}\Big(\sum_{j=1}^{N+1} w_{ij}.s_j(t)\Big) \tag{4.1}$$

où $\mathcal{N}$ est une fonction croissante non linéaire (seuil, sigmoïde ...).

4.1.2 Propriétés du réseau

Il peut être démontré que dans de bonnes conditions d'apprentissage (exemples orthogonaux...), les états attracteurs sont les états mémorisés (ainsi que les états opposés correspondants). Le réseau converge donc vers l'exemple appris le plus proche, ce qui lui confère des propriétés de classification, de complétion et de débruitage.

4.2 Principe de l'utilisation du code corrélationnel dans un réseau de Hopfield

4.2.1 Réinterprétation des poids de connexion

Les poids de Hopfield sont calculés par un apprentissage classique (règle de Hebb) à partir des exemples à apprendre et de l'état où aucun des neurones d'entrée n'est actif : il faut en effet pouvoir stabiliser l'état le plus fréquent où aucun neurone n'émet de PA. Nous éliminons ensuite les auto-connexions car ayant les valeurs les plus importantes, elles limiteraient très fortement les déplacements de PA. Pour ne pas avoir à baisser les seuils calculés par l'apprentissage, nous redistribuons les auto-connexions sur l'ensemble des autres poids.

Les poids ainsi calculés sont réinterprétés de manière à les appliquer au traitement d'entrées codées en terme de distributions spatio-temporelles de PA. Le réseau de Hopfield reçoit classiquement en entrée des valeurs intensitives et les transmet en les pondérant par les poids de la matrice w selon (4.1). On obtient une succession d'états instantanés du réseau. Les entrées étant ici temporelles, nous donnons aussi aux poids de connexion une dimension temporelle : un PA de J aura une influence sur I plus ou moins étalée dans le temps, que nous représenterons par une gaussienne d'écart type d'autant plus faible que w_{ij} est grand (*cf* figure 4.1). Physiologiquement, cela correspond au potentiel membranaire post-synaptique.

Le potentiel généré par le signal émis par J dans le neurone I sera égal à la convolution de ce signal par une gaussienne g_{ij} définie à partir de w_{ij}. Ici,

la fonction choisie est «non physiologique» car non causale. Les gaussiennes sont définies de façon à vérifier les conditions suivantes :
- l'influence totale de la gaussienne doit être toujours identique en valeur absolue (i.e. $\int_{-\infty}^{\infty} |g_{ij}(t)|dt = 1, \forall(i,j)$), et le signe de la gaussienne doit être celui de w_{ij},
- le maximum de la gaussienne doit être égal à w_{ij} (i.e. $g_{ij}(0) = w_{ij}$).
Pour obéir à ces contraintes, nécessaires à l'utilisation des valeurs des poids obtenus par un apprentissage classique, nous choisissons la gaussienne de formule :
$g_{ij}(t) = \frac{e^{-t^2/2v}}{\sqrt{2\pi v}} \cdot \frac{w_{ij}}{\|w_{ij}\|}$ où $v = \frac{1}{2\pi {w_{ij}}^2}$.

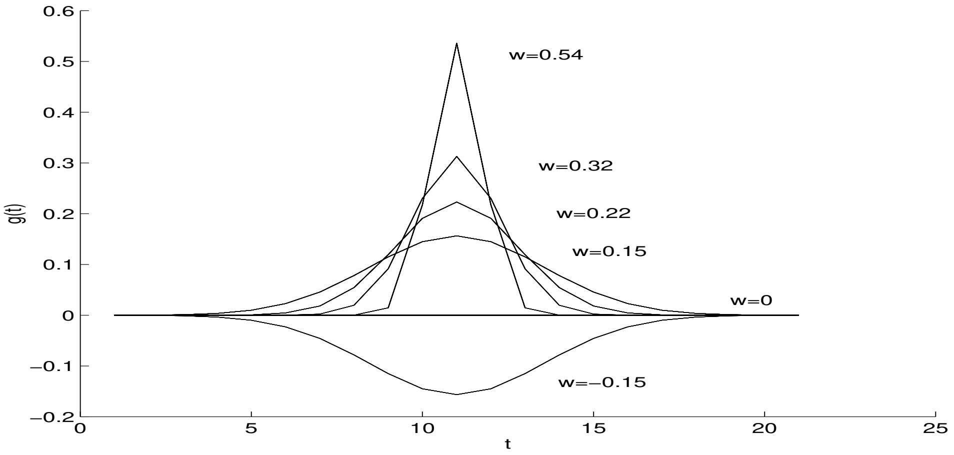

FIG. 4.1 - *DIMENSION TEMPORELLE DES POIDS DE HOPFIELD : Chaque poids de la matrice w résultant de l'apprentissage est représenté par une gaussienne de formule* $g_{ij}(t) = \frac{e^{-t^2/2v}}{\sqrt{2\pi v}} \cdot \frac{w_{ij}}{\|w_{ij}\|}$ *où* $v = \frac{1}{2\pi {w_{ij}}^2}$.

4.2.2 Une adaptation du réseau de Hopfield

Chaque neurone I reçoit donc en entrée l'ensemble des trains de PA des autres neurones. Chaque PA est responsable d'un potentiel post-synaptique excitateur ou inhibiteur, plus ou moins étalé par le temps : sa convolution par la gaussienne correspondant à la connexion entre le neurone émetteur et I. Le neurone I somme ensuite tous les PPSE et PPSI de tous les neurones pour établir l'évolution de son potentiel de membrane au cours du temps :

$$e_i(t) = \sum_{j \neq i} (s_j(t) * g_{ij}(t)) \tag{4.2}$$

Cette information est traitée fenêtre de temps par fenêtre de temps (de longueur T) pour générer les PA de I en conservant leur fréquence sur la fenêtre (i.e. le code propre de I n'est pas altéré). Les PA sont déplacés vers les maxima locaux les plus importants à condition qu'ils soient supérieurs au biais. S'il n'y en a pas ou pas assez, les PA se repositionnent aléatoirement. La réponse dépend donc de la distribution temporelle du potentiel : plus elle est uniforme, plus le neurone aura tendance à décharger aléatoirement, plus elle comporte de pics, plus les PA se synchroniseront sur ces pics. Ceci aura pour effet de synchroniser les neurones connectés par des poids forts (*cf.* figure 4.2).

```
Pour chaque neurone N du réseau :
  convolution des signaux de chaque neurone M par g(N,M)
  calcul de la somme S de ces signaux convolués
Dans chaque fenêtre de longueur T :
  pour chaque neurone N :
    calcul du nombre initial de PA de N dans la fenêtre
    repositionnement de ces PA sur les maxima locaux de S
                                supérieurs au biais de N
    si il n'y a pas assez de maxima locaux sur la fenêtre :
      positionnement aléatoire des PA restants
```

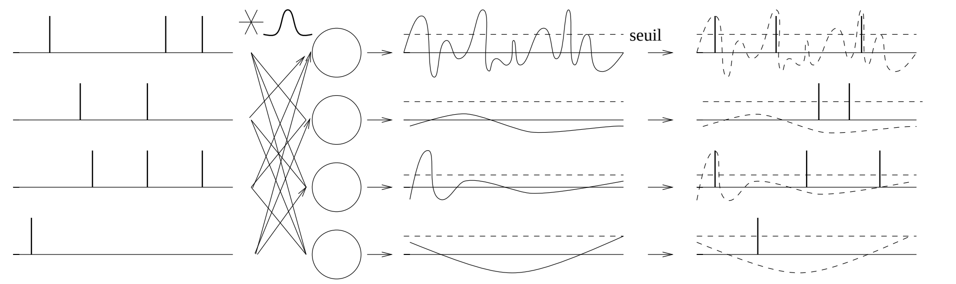

FIG. 4.2 - *ADAPTATION DU RÉSEAU DE HOPFIELD : Chaque neurone reçoit la somme des entrées convoluées par les gaussiennes correspondant aux poids de connexion. Les PA sont ensuite repositionnés sur les maxima locaux (ML) supérieurs au seuil. L'intensité moyenne est conservée. S'il y a moins de PA à placer que de ML, les maxima les plus importants sont retenus (ligne 1). S'il y a plus de PA que de ML (ligne 3) ou aucun ML au dessus du seuil (lignes 2 et 4), les PA restant à placer sont positionnés aléatoirement.*

4.3 Résultats

Le réseau reçoit en entrée N trains de PA où le pattern de départ est donné par une synchronisation totale des neurones concernés. Une couche de lecture permet de mettre en évidence les neurones synchronisés entre eux à la sortie du réseau à tout moment. Après moins d'une dizaine d'itérations, les neurones synchronisés correspondent au pattern appris le plus proche.

4.3.1 Propriétés partagées par le réseau classique et le réseau réadapté

Les propriétés du réseau classique de Hopfield sont conservées : il peut compléter ou débruiter un exemple appris (*cf* figure 4.3).
Le débruitage est rapide et efficace : le réseau décorrèle bien les activités des neurones corrélés au départ mais non connectés, car leurs potentiels sont alors assez uniformes et leurs activités par conséquent aléatoires. La complétion est par contre plus progressive car il s'agit de réaligner des PA qui peuvent différer en nombre dans la majorité des fenêtres de réalignement.

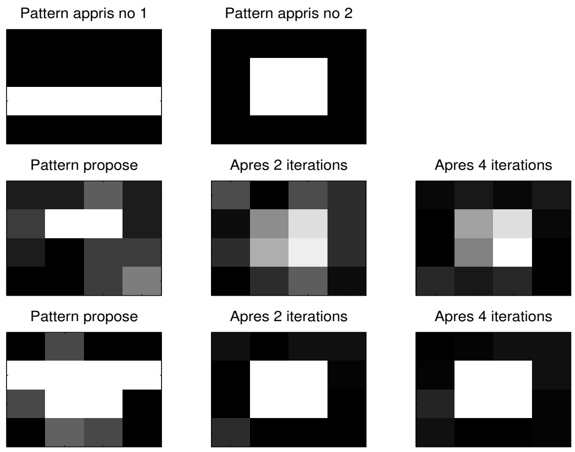

FIG. 4.3 - *PROPRIÉTÉS DE COMPLÉTION ET DÉBRUITAGE : Visualisation de la sortie du lecteur avant seuillage (cf. §3.1.3) pour un voisinage égal au réseau. Les propriétés de complétion (ligne 2) et de débruitage (ligne 3) sont retrouvées avec notre modèle.*

4.3.2 Nouvelle propriété : le débruitage temporel

Le réseau peut débruiter temporellement un signal en réalignant des PA proches dans le temps de neurones corrélés entre eux : cela revient à éliminer le « jitter » dû par exemple aux variations du délai de propagation des PA initialement alignés. Les distributions sont plus resserrées en sortie qu'en entrée : l'effet du réalignement temporel est de concentrer temporellement l'activité neuronale correspondant à la forme reconnue par le réseau (*cf* figure 4.4). Ce débruitage temporel s'effectue en même temps qu'une reconnaissance spatiale de forme apprise.

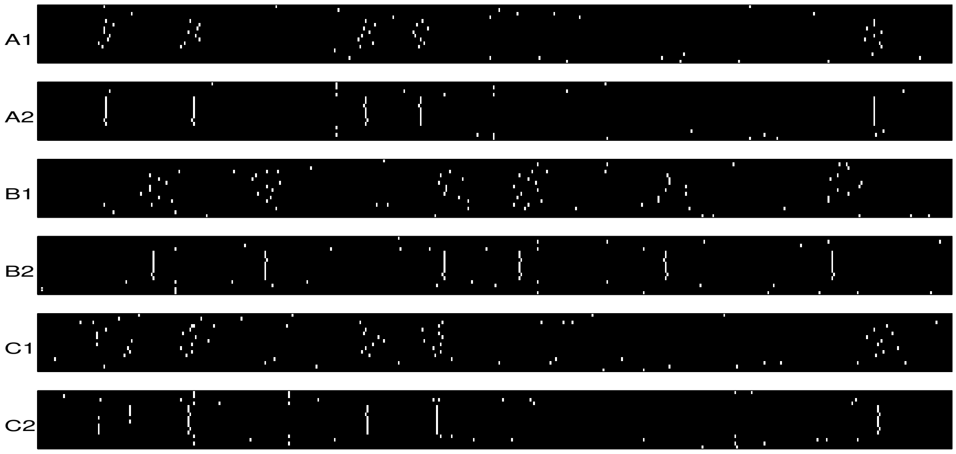

FIG. 4.4 - *DÉBRUITAGE TEMPOREL DU SIGNAL : Le réseau réaligne des PA corrélés temporellement de façon imprécise. En A1, l'imprécision est faible, en A2, le résultat après 5 itérations. En B1 et C1, l'imprécision est plus importante, et le réalignement peut s'effectuer correctement (B2), ou un groupe de PA peut être séparé en deux sous-groupes si ils sont trop éloignés (premier PA bruités de C2).*

4.3.3 Débruitage temporel de deux patterns codés simultanément

Lorsque plusieurs patterns appris sont codés simultanément en entrée, un alignement temporel peu précis peut être responsable d'une mauvaise discrimination des deux objets. Le réseau sépare les deux objets en réalisant le réalignement temporel. La figure 4.5 représente cette séparation pour des objets codés avec une imprécision de 10 ms dont les positions moyennes sont décalées de 30ms (premiers PA) à 2 ms (derniers PA). Les objets sont séparés temporellement lorsque l'intervalle de temps entre les temps moyens d'émission des deux populations est assez grand (il suffit ici de 3 ms).

4.3.4 Débruitage spatial et lecture de plusieurs objets codés simultanément

Si le réseau reçoit en entrée plusieurs objets bruités cette fois dans l'espace, codés par plusieurs groupes de synchronisation (certains neurones pouvant faire partie de deux groupes différents), il pourra débruiter en parallèle les deux patterns, et la sortie correspondra au codage corrélationnel des deux objets appris les plus proches de ceux de départ.

Notre couche de lecture, capable de mettre en évidence l'ensemble des neurones qui présentent une corrélation moyenne importante ne permettra de visualiser que la superposition des deux objets. Pour les séparer, on peut soit éliminer l'intégration temporelle du lecteur et observer alternativement les

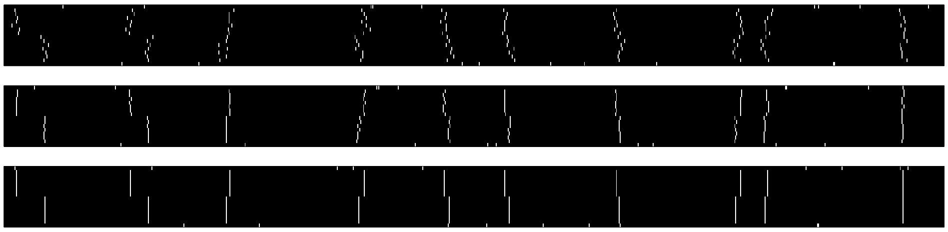

FIG. 4.5 - *SÉPARATION TEMPORELLE DE DEUX PATTERNS : Deux patterns sont codés simultanément avec une imprécision temporelle dans l'alignement de 10 ms. L'intervalle de temps entre les moyennes des positions des PA correspondant aux deux objets vaut respectivement 30, 20, 10, 8, 7, 6, 5, 4, 3, et 2 ms (ligne 1). Séparation des deux objets après une itération du réseau (ligne 2) puis 5 itérations (ligne 3).*

deux objets, soit visualiser en parallèle les neurones corrélés avec un neurone appartenant à un seul des deux patterns appris. On observe alors que les deux patterns sont codés en sortie, et qu'ils ont été débruités (*cf.* figure 4.6).

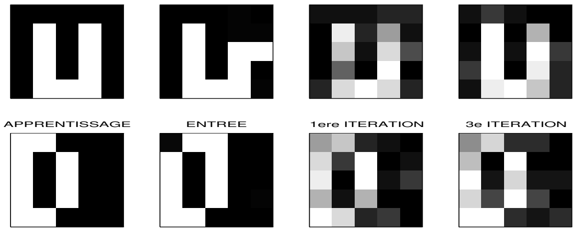

FIG. 4.6 - *DÉBRUITAGE SIMULTANÉE DE DEUX PATTERNS : La première colonne représente les patterns appris : U et D ; la deuxième les patterns proposés simultanément par un codage corrélationnel ; et les troisième et quatrième les neurones corrélés avec un neurone de U en haut et avec un neurone de D en bas.*

4.4 Vers un réseau probabiliste ?

Afin de réduire l'effet lié au choix des maxima d'activité, nous avons modifié l'étape de positionnement des PA en fonction du potentiel de membrane. Les PA ne sont plus positionnés sur les maxima locaux, mais de façon aléatoire avec une probabilité fonction croissante du potentiel de membrane. La probabilité d'émission d'un PA est nulle au dessous du seuil, puis augmente avec le carré du potentiel au dessus du seuil :

$$s_i(t) = \frac{R_+(e_i(t) - seuil_i)^2}{\sum_u R_+(e_i(u) - seuil_i)^2}$$

où $e_i(t)$ est calculé selon (4.2), et R_+ est une fonction de redressement du type : $R_+(t) = t$ si $t \geq 0$ et $R_+(t) = 0$ si $t < 0$.

Le réseau conserve alors ses propriétés de reconnaissance spatiale, mais n'est bien sûr plus capable de réaligner précisément les PA dans le temps. La sortie donne une corrélation temporellement moins précise des patterns. Le lecteur du §3.1.3 doit utiliser une fenêtre temporelle T plus large.

L'intérêt de ce type de réseau est qu'il permettrait de conserver en sortie une image de la distribution temporelle du potentiel de membrane $e_i(t)$ qui ne serait pas limitée à ses maxima locaux.

Chapitre 5

Conclusion

Du point de vue biologique, le transfert d'information par trains de PA autorise au moins deux types de codes. Le code propre est particulièrement utile aux neurones sensoriels et moteurs. Mais aux niveaux intermédiaires, il ne permet pas à lui seul d'expliquer certains traitements rapides d'information corticale. Le code corrélationnel prend alors toute son importance en permettant un apport d'information supplémentaire, et une explication du phénomène de « binding ».

Du pont de vue de la modélisation, code propre et code corrélationnel semblent compatibles, et même complémentaires. Nous montrons qu'ils peuvent être manipulés indépendamment, qu'ils soient portés par un neurone unitaire (signal à valeurs binaires) ou par une population (signal à valeurs continues). La phase de la TF s'avère être un outil parfaitement adapté au contrôle de la corrélation des signaux continus.

Il est de plus possible de modéliser des réseaux fonctionnant avec des trains de PA utilisant ces deux codes, et de modifier l'information donnée par l'un sans altérer celle portée par l'autre. Le code corrélationnel a une valeur fonctionnelle qui peut être ajoutée à celle du code propre. De nouvelles propriétés semblent pouvoir émerger de son utilisation, comme le codage simultané de plusieurs objets.

En ce qui concerne les perspectives de ce travail, il serait intéressant de pouvoir exploiter plus rigoureusement le spectre de phase pour étudier et contrôler la synchronisation de différents signaux. Bien qu'encore assez peu utilisé en général (on manipule plutôt le spectre d'intensité de la TF), et non exploité dans ce domaine à notre connaissance, cet outil mathématique semble prometteur. Développer des réseaux qui utiliseraient à la fois code propre et code corrélationnel pourrait aussi faire émerger des propriétés intéressantes. Nous aimerions finalement pouvoir adapter l'outil de phase à la manipulation par ces réseaux des corrélations de signaux à valeurs continues.

Annexe A

Rappels mathématiques

A.1 Transformée de Fourier

Pour tout signal $g(t)$ périodique donné, on obtient la transformée de Fourier $G(f) = TF[g(t)]$ par la formule :

$$G(f) = \int_{-\infty}^{\infty} g(t) \cdot e^{-i2\pi ft} dt$$

Pour nos signaux nous utilisons la transformée de Fourier discrète :

$$G(f) = \sum_{k=-\infty}^{\infty} g(k\Delta t) \cdot e^{-i2\pi fk\Delta t}$$

Le résultat est un vecteur complexe que l'on peut représenter par :
- son spectre d'amplitude

$$Sa(f) = \parallel G(f) \parallel$$

qui est le module de $G(f)$ et qui nous donne l'énergie contenue dans ce signal pour chaque fréquence, et
- son spectre de phase qui donne le déphasage de chaque fréquence :

$$Sp(f) = \arcsin(\frac{\Im[G(t)]}{\Re[G(t)]})$$

La transformée de Fourier est inversible, c'est à dire qu'à partir de la transformée de Fourier $G(f)$, on peut retrouver le signal $g(t) = TFI[G(f)]$ par la formule :

$$g(t) = \int_{-\infty}^{\infty} G(f) \cdot e^{i2\pi ft} df$$

A.2 Produit de convolution ($*$)

Le produit de convolution $h = f * g$ de deux fonctions f et g est défini par :

$$h(x) = \int_{-\infty}^{\infty} f(t) \cdot g(x-t)dt$$

A.3 Calcul des CCH et des ACH

Dans la section 3.1, lorsque les signaux sont à valeurs binaires (trains de PA), les CCH et ACH sont calculés comme mentionné dans le §2.2.1.

Dans la section 3.2, lorsque les signaux sont à valeurs continues, le CCH $s_1 \mathrm{x} s_2$ est calculé de la façon suivante :

$$CCH(s_1 \mathrm{x} s_2) = s_1' * s_2 \text{ où } \forall t, s_1'(t) = s_1(-t)$$

et $*$ représente le produit de convolution. L'ACH de s_1 se calcule comme le CCH $s_1 \mathrm{x} s_1$. Ce calcul est aussi valable pour obtenir les corrélogrammes des trains de PA.

Lorsque l'on veut observer l'organisation temporelle des corrélogrammes (figure 3.10 et 3.11), il peut être intéressant d'éliminer le biais dû au caractère fini des signaux, responsable de la décroissance du corrélogramme vers ses extrémités. Pour un signal de longueur M (donc un corrélogramme de longueur $2M-1$), le biais est éliminé en multipliant sa valeur en k par $M - |k|$.

Bibliographie

[1] Adrian, E.D.,1928, *The Basis of Sensation*, London, Christophers.

[2] Feldman, J.A., 1990, Computational constraints on higher neural representations, in *Computational Neuroscience*, E.L. Schwartz (ed.), Cambridge : MIT Press, pp. 133-150.

[3] Barlow, H.B., 1972, Single units and sensations : a neuron doctrine for perceptual psychology,*Perception*, vol. 1, pp.371-392.

[4] Arbib, M.A., 1989, *The Metaphoric Brain 2 : Neural Nets and Beyond*, New York, John Wiley.

[5] Churchland, P.S., Sejnowski, T.J., 1992, *The Computational Brain*, Cambridge, MIT Press.

[6] Cariani, P., 1995, As if time really mattered: temporal strategies for neural coding of sensory information, *Communication and Cognition - Artificial Intelligence*, vol. 12, pp. 157-219

[7] Thorpe, S.J., M. Imbert, 1989, Biological constraints on connectionist models, in *Connectionism in Perpective*, R. Pfeiffer et al (eds).

[8] Abeles, M., 1982, Role of the cortical neuron : integrator or coincidence detector, *Israel Journal of Medical Sciences*, vol. 18, pp. 83-92.

[9] Gerstein, G.L., D.H. Perkel, 1972, *Biophys. J.*, vol.12, pp. 453-473.

[10] Nowak, L.G., M.H.J. Munk, J.I. Nelson, J. Bullier, 1995, The structural basis of cortical synchronisation I : Three types of interhemispheric coupling, *J. Neurophysiol.*, vol. 74, pp. 2379-2400.

[11] Perkell, D.H., T.H. Bullock, 1968, *Neural Coding*, Neurosciences Research Program Bulletin, vol. 40, pp. 137-146.

[12] Perkel, D.H., G.L. Gerstein, and G.P. Moore, 1967, Neuronal spike trains and stochastic point processes I : The single spike train. *Biophys. J.*, vol 7, pp. 419-440.

[13] Lestienne, R., E. Gary-Bobo, P. Saillour, J. Przybyslawski, M. Imbert, 1990, Temporal Correlation in Modulated Evoked Responses in the Visual Cortex Cells of the Cat, *Biological Cybernetics*, vol. 62, pp. 425-440.

[14] Gerstein, G.L., Functional Association of Neurons : Detection and interpretation, in *Communication and Coding in the Nervous System*.

[15] Abeles, M., G.L. Gerstein, 1988, Detecting spatiotemporal firing patterns among simultaneously recorded single neurons, *J. Neurophysiol.*, vol. 60, pp. 909-924.

[16] Abeles, M., H. Bergman, E. Margalit, E. Vaadia, 1993, Spatiotemporal firing patterns in the frontal cortex of behaving monkeys, *J. Neurophysiol*, vol. 70, pp. 1629-1638.

[17] Krüger, J., J.D. Becker, 1991, Recognizing the visual stimulus from neuronal discharges, *Trends in Neurosci.*, vol. 14, pp. 282-286.

[18] Vaadia, E., I. Haalman, M. Abeles, H. Bergman, Y. Prut, H. Slovin, A. Aertsen, 1995, Dynamics of neuronal interactions in monkey cortex in relation to behavioural events, *Nature*, vol. 373, pp. 515-518.

[19] French, A.S., R.B. Stein, 1970, A flexible neural analog using integrated circuits, *IEEE Trans. Biomed. Eng.*, vol. 17, pp. 248-253.

[20] Bush, P.C., R.J. Douglas, 1991, Synchronisation of bursting action potential discharge in a model network of neocortical neurons, *Neural Computation*, vol. 3, pp. 19-30.

[21] Amit, D.J., 1988, Neural network counting chimes. *Proc. Nat. Acad. Sci. USA*, vol. 85, pp.2141-2145.

[22] Demongeot, J., T. Hervé, F. Berthommier, O. François, 1989, Neural networks : From neurocomputing to neuromodelling, *Cell to Cell Signalling : From experiments to Theoretical Model*, A. Goldbeter (ed), Academic Press, pp. 109-119.

[23] Nischwitz, A., H. Glünder, 1995, Local lateral inhibition : a key to spike synchronization ?, *Biol. Cybern.*, vol. 73, pp. 389-400.

[24] Klaassen, A.J., 1992, Un modèle de neurone formel à intégration spatio-temporelle distribuée, *actes des* 6^{e} *journées Neurosciences et Sciences de l'Ingénieur 1992*.

[25] Berthommier, F., N. Buonviso, M. Chaput, 1995, A probabilistic model of temporal processing in the olfactory bulb, *Proc. AIDRI*, Hermes, Paris.

[26] von der Malsburg, C., W. Schneider, 1989, A neural cocktail-party processor, *Biol. Cybern.*, vol. 54, pp. 29-40.

[27] Singer, W., 1994, Time as Coding Space In Neocortical Processing, in *Temporal Coding in the Brain*, G. Buzsáki, R. Llinás, W. Singer, A. Berthoz, Y. Christen (eds) Springer-Verlag, pp. 51-79.

[28] Engel, A.K., P. König, A.K. Kreiter, W. Singer, 1991, Interhemispheric synchronisation of oscillatory neuronal responses in cat visual cortex, *Science*, vol. 252, pp. 1177-1179.

[29] Gray, C.M., W. Singer, 1989, Stimilus-specific neuronal oscillations in orientation columns of cat visual cortex, *Proc. Natl. Acad. Sci. USA*, vol. 86, pp. 1698-1702.

[30] Nelson, J.I., P.A. Salin, M.H.J. Munk, M. Arzi, J. Bullier, 1992, Spatial and temporal coherence in corticocortical connections : a cross-correlation study in areas 17 and 18 in the cat, *Visual Neurosci.*, vol. 9, pp. 21-38.

[31] Roubine, E., 1970, *Introduction à la théorie de la communication, Tome III : Théorie de l'information*, M. Ponte, P. Grivet (eds).

[32] Hopfield, J.J., 1982, Neural networks and physical systems with emergent collective computational abilities, *Proceeding of the National Academy of Sciences*, vol. 79, pp. 2554-2578.