

63461

ORSAY
n° d'ordre : 3502

UNIVERSITE PARIS SUD

Centre d'Orsay

T H E S E

De Doctorat d'Etat Es Sciences Mathématiques

présentée pour obtenir le grade de

DOCTEUR ES-SCIENCES

par

Marie COTTRELL

**Sujet : MODELISATION DE RESEAUX DE NEURONES PAR DES CHAINES DE MARKOV
ET AUTRES APPLICATIONS.**

Soutenu le 23 septembre 1988 devant le Jury composé de :

MM. AZENCOTT	<i>Robert</i>
ANCONA	<i>Alano</i>
BRETAGNOLLE	<i>Jean</i>
DEMONGEOT	<i>Jacques</i>
HERAULT	<i>Jeanny</i>
METIVIER	<i>Michel</i>
RUGET	<i>Gabriel</i>

ABSTRACT

The first part of the thesis consists of a paper published in IEEE Trans. Aut. Control (vol. AC-28, n° 9, 1983), with J.C. Fort and G. Malgouyres. It gives two methods of calculating the exit time of a Markov chain from an attraction domain : this time is extremely long, so we use an exponential change of probability (that of large deviations theory), for a fast simulation and a non-standard approximation by diffusion.

The second part includes two papers published with J.C. Fort in the Annales de l'IHP, Probabilités et Statistiques (vol. 23, n° 1, 1987), and in Biological Cybernetics (n° 53, 1986). In the first one, we prove the convergence of Kohonen's self-organizing algorithm, in dimension 1. In the second one, we define another self-organizing algorithm, which is a simplified variant of Kohonen's, and we prove its convergence in dimensions 1 and 2.

In the third part, published in Biological Cybernetics (n° 58, 1988), we solve the problem of the connection matrix calculus for a Mac-Culloch or Hopfield neural network, so as to get the largest attractivity for the deterministic algorithm and non-orthogonal patterns. Then we calculate the attractivity of each memorized pattern, for a given connection matrix.

The last part is devoted to the study of the role of inhibition in a nearest-neighbours-connected neural network. The model closely resembles the biological reality of the young animal's cerebellar cortex. We prove that, when inhibition is smaller than a certain threshold, the network is ergodic and works in a stationary way. Conversely, when inhibition increases, striped or moiré responses appear, whose form and width depend on the considered neighbourhood size.

KEYWORDS : **Large deviations, Approximation by diffusion, Markov chains, self-organization, Associative memory, Neural networks, Inhibitory coupling, Attractivity.**

A.M.S. CODE : **60 F 10 - 60 J 27 - 60 J 60 - 60 J 20 - 60 k 35 - 94 C 99 - 92 A 90**

Je voudrais tout d'abord remercier Jean-Claude Fort avec lequel j'ai cosigné 3 des articles de cette thèse. Qu'il trouve ici l'expression de toute mon amitié.

Je remercie Jean Bretagnolle, mon directeur de thèse, Robert Azencott et Gabriel Ruget pour m'avoir proposé des sujets de recherche, encouragée et aidée au cours de nombreuses discussions. Je les remercie d'avoir accepté d'être membres du jury et tout particulièrement Robert Azencott d'en assurer la présidence.

Je tiens à remercier Michel Métivier et Jacques Demongeot, rapporteurs pour la commission des thèses et Jeanny Herault de marquer par leur participation au jury la pluridisciplinarité qui caractérise mon travail.

Merci à Alano Ancona pour avoir bien voulu me proposer un sujet de deuxième thèse et m'avoir guidée dans la découverte de l'analyse complexe et de la mesure harmonique.

Merci à Herbert Axelrad qui a suggéré à partir de ses propres études de neurophysiologiste le modèle étudié dans la dernière partie de ma thèse.

Plus globalement je remercie tous les membres de l'équipe de statistique appliquée d'Orsay pour l'amitié qu'ils m'ont témoignée, pour leurs conseils et encouragements de toute sorte.

Je remercie plus particulièrement Didier Dacunha-Castelle qui m'a accueillie dans son équipe et permis ainsi une reconversion réussie après plusieurs années d'enseignement à plein temps, Roger Astier, Gérard Malgouyres avec qui j'ai collaboré au tout début, Bernard Prum, Xavier Guyon, Paul Doukhan, Christian Léonard avec qui j'ai eu de fructueuses discussions, Valentine Genon-Catalot dont les marques d'amitié ont rendu très agréable l'ambiance de notre bureau commun, Yves Misiti et Patrick Jakubowicz qui ont facilité mes rapports quelquefois conflictuels avec les micro-ordinateurs, Patrice Assouad pour les performances et les beautés de son traitement de texte.

Merci d'une façon générale à tous ceux (universitaires ou non), qui pendant ces dernières années ont encouragé et facilité mon travail.

Enfin, un grand merci à Mmes Anne-Marie Baillet et Nicole Parvan pour le soin et le talent avec lesquels elles ont mécanographié tout mon travail, dans ses différentes versions, et Mmes Andrée Zielinsky et Raymonde Fercoq pour la belle impression de l'ensemble.

Que tous et toutes trouvent ici l'expression de ma gratitude et la preuve que la solidarité et l'entraide en matière scientifique sont des atouts essentiels.

MOTS-CLEFS

PREMIERE PARTIE	:	Temps de sortie Approximation-diffusion Système ALOHA Processus de Markov Grandes déviations Evénements rares
DEUXIEME PARTIE	:	Rétinotopie Processus d'auto-organisation Algorithme de Kohonen
TROISIEME PARTIE	:	Réseaux de neurones Mémoire associative Matrice de connexion Réseau d'automates Capacité de mémoire
QUATRIEME PARTIE	:	Cortex cérébelleux Liaisons inhibitrices Réponses à des stimuli Processus de Markov.

TABLE DES MATIERES

I. GRANDES DEVIATIONS ET EVENEMENTS RARES DANS L'ETUDE D'ALGORITHMES STOCHASTIQUES.

1. Grandes déviations
2. Méthode de simulation rapide
3. Approximation par une diffusion
4. Exemples

II. ETUDE D'UN PROCESSUS D'AUTO-ORGANISATION.

1. Algorithme de Kohonen en dimension un
2. Démonstration du théorème 1.1. (Mise en ordre)
3. Démonstration du théorème 1.2. (Existence d'une distribution stationnaire)
4. Quelques propriétés complémentaires.

III. UN MODELE STOCHASTIQUE DE LA RETINOTOPIE : UN PROCESSUS D'AUTO-ORGANISATION.

1. Le modèle
2. Le cas à une dimension
3. Le cas à deux dimensions
4. Effet d'adaptation décroissant avec le temps
5. Conclusion.

IV. STABILITE ET ATTRACTIVITE DANS LES RESEAUX DE MEMOIRE ASSOCIATIVE.

1. Introduction
2. Stabilité - Attractivité
3. Cas des patrons orthogonaux
4. Recherche de la matrice des connexions
5. Taille des domaines d'attraction des patrons
6. Matrice optimale

7. Cas des patrons orthogonaux (bis)
8. Domaines d'attraction de tailles égales
9. Energie
10. Conclusions provisoires.

APPENDICES : Exemples et simulations.

V. **ANALYSE MATHÉMATIQUE D'UN MODÈLE DU CORTEX CÉRÉBELLEUX :**

I. EFFET DE L'ACTION INHIBITRICE DES COLLATÉRALES RÉCURRENTES.

1. Modélisation du cortex cérébelleux
2. Modèle mathématique du réseau
3. Chaîne extraite - Irréductibilité - Apériodicité - Temps de retour
4. Récurrence - Régime stationnaire - $\theta < \theta_0$
5. Divergence : Cas de deux neurones
6. Divergence : Cas général
7. Conclusion.

APPENDICES - Exemples et simulations.

II. EFFET DE L'ACTION INHIBITRICE DES COLLATÉRALES RÉCURRENTES EN PRÉSENCE DE STIMULI.

1. Introduction
2. Modèle mathématique du réseau de neurones avec excitation
3. Conclusion.

APPENDICES - Exemples et simulations.

INTRODUCTION

Cette thèse comporte 5 articles. Les quatre premiers sont publiés. Le dernier est une prépublication d'ORSAY-PARIS-SUD.

PREMIERE PARTIE.

GRANDES DEVIATIONS ET EVENEMENTS RARES DANS L'ETUDE D'ALGORITHMES STOCHASTIQUES.

(A fait l'objet d'une publication dans *IEEE Trans. on Aut. Control*, en collaboration avec J.C. FORT et G. MALGOUYRES) [6].

Pour estimer des probabilités extrêmement petites, une méthode a été fréquemment utilisée : la méthode de l'échantillonnage préférentiel ([8]). Cette méthode est fondée sur la remarque suivante : Si P et \tilde{P} sont deux lois de probabilité telles que P soit absolument continue par rapport à \tilde{P} , de densité $\frac{dP}{d\tilde{P}}$, et si X est une v.a.,

$$E_P(X) = \int_{\Omega} X(\omega) dP(\omega) = \int_{\Omega} X(\omega) \frac{dP}{d\tilde{P}}(\omega) d\tilde{P}(\omega).$$

Si X est la v.a. indicatrice d'un événement A rare sous la loi P , événement presque impossible à observer au cours d'une simulation de loi P , on simule un n -échantillon sous une loi \tilde{P} favorisant l'apparition de l'événement A , et on estime $E_P(X) = P(A)$, par $\frac{1}{n} \sum_{i=1}^n \frac{dP}{d\tilde{P}}(\omega_i) \mathbb{1}_A(\omega_i)$. Cet estimateur est sans biais et convergent mais

sa variance $\frac{1}{n} [E_{\tilde{P}} \left(X \frac{dP}{d\tilde{P}} \right)^2 - (E_P(X))^2]$ n'est pas contrôlée, et peut

être grande et ôter tout intérêt à cet estimateur.

Pour des événements A du type "franchissement" d'une barrière positive a pour un processus de Markov ayant 0 pour attracteur et équilibre asymptotiquement stable, nous avons montré que le changement de probabilité exponentiel utilisé pour démontrer les théorèmes dits de grandes déviations ([2], [4], [11]) fournit un estimateur optimal. Il est en effet de variance minimum parmi tous les changements de probabilité exponentiels, c'est-à-dire obtenus en remplaçant la loi partant de x , $d\mu_x(y)$ par $d\tilde{\mu}_x(y) = C \exp(v(x)y) d\mu_x(y)$.

Le changement de probabilité optimal est défini par $v(x) = \lambda_x$ avec $l_x(\lambda_x) = 0$, pour l_x log-Laplace de μ_x .

Son interprétation est simple : si, par exemple, le processus de Markov est défini par

$$d\mu_x(x+1) = p(x) = 1 - d\mu_x(x-1), \text{ avec } p(x) < \frac{1}{2},$$

le changement de probabilité correspond à l'échange de $p(x)$ et de $1 - p(x)$.

Ce changement de probabilité recentre le processus autour du "chemin de sortie" optimum (celui qui minimise l'intégrale d'action

$$I(\varphi) = \int_0^a \frac{h_x(\dot{x})}{\dot{x}} dx, \text{ avec } h_x \text{ transformée de Cramer de la loi } \mu_x).$$

Nous avons utilisé aussi ce changement de probabilité optimal pour obtenir une formule approchée du temps d'atteinte de la barrière $a > 0$, en utilisant une approximation diffusion non standard.

Soit $0 < y < x_0 \ll n_b < a$.

On évalue d'abord la probabilité, partant de x_0 , d'arriver en n_b avant d'arriver en y . Cette probabilité est très faible, elle est calculée au moyen d'une approximation-diffusion du processus \tilde{X} transformé par le changement de probabilité défini ci-dessus. On écrit ensuite des conditions frontières en x_0 pour évaluer la probabilité, partant de 0, d'arriver en n_b , au moyen d'une approximation-diffusion du processus X lui-même, considéré au voisinage de 0 comme un processus de Orstein-Ulhenbeck. On répète enfin cette opération entre n_b et a .

On associe toujours au processus de Markov étudié $X(t)$, le processus à petits gains $X^\varepsilon(t) = \varepsilon X(\frac{t}{\varepsilon})$ comme par exemple dans [3], et les résultats obtenus sont des résultats asymptotiques quand $\varepsilon \rightarrow 0$.

Nous avons appliqué les deux méthodes (simulation et calcul approché) pour calculer le temps de fonctionnement prévisible d'un canal de communication ALOHA, ([1], [9], [10]) : il s'agit d'un canal partagé entre un très grand nombre d'utilisateurs, géré par une procédure aléatoire.

Lorsque deux utilisateurs cherchent à émettre simultanément, ils sont bloqués. Si l'on note $N(t)$ le nombre d'utilisateurs bloqués à l'instant t , le temps de fonctionnement du canal est le temps d'atteinte d'une valeur critique n_c après laquelle $N(t)$ tend rapidement vers $+\infty$.

Depuis la publication de notre article, plusieurs auteurs ont appliqué ou généralisé ces techniques qui se sont révélées utiles : par exemple DUPUIS et KUSHNER [7] ou de CAMBRY [5].

DEUXIEME PARTIE.**ETUDE D'UN PROCESSUS D'AUTO-ORGANISATION.**

(A fait l'objet d'une publication dans les Annales de l'Institut Poincaré, en collaboration avec J.C. FORT) [1].

et UN MODELE STOCHASTIQUE DE LA RETINOTOPIE.

(A fait l'objet d'une publication dans Biol. Cybern., en collaboration avec J.C. FORT) [2].

Dans de multiples réseaux de neurones réels, on constate certaines propriétés d'ordre porteuses de sens : correspondances ordonnées entre zones réceptrices dans le cortex visuel et orientations spatiales des stimuli, rétinotopie, c'est-à-dire que des zones voisines de la rétine sont en correspondance avec des zones voisines du cortex visuel, mêmes constatations en ce qui concerne le cortex sensoriel ou auditif...

Cet ordre n'est pas entièrement programmé génétiquement. Il se met en place au cours d'une période d'apprentissage, pendant laquelle les synapses se modifient suivant la règle de Hebb : l'efficacité d'une synapse se modifie de façon à augmenter le produit de l'activité pré- et post-synaptique ([4]).

Une disposition spatiale, une auto-organisation, se met en place sous le contrôle de l'information reçue. Il y a "reproduction" conforme des relations d'ordre, de voisinage, de ressemblance entre les stimuli.

T. KOHONEN ([6], [7], [8], [9]) a défini un algorithme d'auto-organisation simple, basé sur la règle de Hebb et spectaculairement efficace :

On dispose de N unités $i = 1, \dots, N$, qui réagissent à un stimulus ω à n coordonnées $(\omega_1, \dots, \omega_n)$, en fournissant une réponse $\sum_{j=1}^n \mu_{ij} \omega_j$. L'unité i est donc caractérisée par son "image", le vecteur $m_i = (\mu_{i1}, \dots, \mu_{in})$. On place m_1, m_2, \dots, m_N dans le même espace que les stimuli et on les normalise. Un système de voisinage $\{V(i) / i = 1, \dots, N\}$ est défini. Au temps $t=0$, les $m_i(0)$ sont aléatoires et indépendants. Si $(m_i(t))$ est l'état au temps t , on présente au temps $(t+1)$ un stimulus ω^{t+1} , tiré suivant une loi p . On détermine alors i_0 , numéro de l'unité appelée par ω (unité de meilleure réponse) par

$$\|\omega^{t+1} - m_{i_0}(t)\| = \min_i \|\omega^{t+1} - m_i(t)\|$$

et on renforce l'ajustement en i_0 et en ses voisins, en posant

$$\begin{cases} \mu_{ij}^{t+1} = \mu_{ij}^t + \varepsilon(t) (\omega_j^{t+1} - \mu_{ij}^t) & \text{si } i \in V(i_0) \cup \{i_0\} \\ \mu_{ij}^{t+1} = \mu_{ij}^t, & \text{sinon} \end{cases}$$

avec ε petit et positif.

(On applique une homothétie de rapport $1-\varepsilon$ et de centre ω^{t+1} au point i_0 et à ses voisins : on les rapproche entre eux et aussi de ω^{t+1}).

Cet algorithme (en version discrète ou continue) définit un processus de Markov, convergent vers un maillage $(m_i(\infty))$, qui reproduit la densité p des stimuli ω et leurs ressemblances mutuelles.

T. KOHONEN a présenté de nombreuses applications de cet algorithme. Mais il n'a pas donné de démonstration explicite de la convergence de l'algorithme, même s'il en a esquissé quelques arguments.

Dans le premier article de cette partie (Etude d'un processus d'auto-organisation), nous avons démontré la convergence en loi du processus $m(t)$, pour des unités alignées, en dimension 1, avec p uniforme et ε fixe, (contrairement à ce qu'annonçait T. KOHONEN, il n'est pas nécessaire que $\varepsilon(t)$ tende vers 0).

On montre que la convergence se fait en deux phases : une phase de mise en ordre correspondant à l'entrée dans une des deux classes absorbantes (dispositions monotones) et une phase de convergence proprement dite. Pour cela, nous montrons que la mise en ordre s'effectue en un temps fini et que la chaîne est Doeblin, en prouvant l'existence d'un entier m et d'une constante $C > 0$, tels que, pour tout x ordonné, et tout ensemble E de mesure de Lebesgue $\lambda(E) > 0$, $P^m(x, E) > C \lambda(E)$.

Il est facile de mettre en évidence une fonctionnelle décroissante au sens large du temps : c'est un indice de désordre, mesuré par le nombre de triplets (m_{i-1}, m_i, m_{i+1}) non ordonnés, c'est-à-dire vérifiant $(m_{i+1} - m_i)(m_i - m_{i-1}) < 0$. Mais il faut démontrer que cet indice décroît strictement en temps fini et avec une probabilité strictement positive.

Il est donc nécessaire de construire "à la main" cette minoration, la chaîne étant non classique et à états dans $[0,1]^N$.

Nous avons enfin étudié le cas où $\varepsilon(t) > 0$: si $\sum \varepsilon(t) = +\infty$ et $\sum \varepsilon^2(t) < +\infty$, (conditions classiques des algorithmes de Robbins

et Monro), la chaîne converge presque sûrement vers une situation constante m^* , solution d'un système linéaire explicite, qui avait été écrit par T. KOHONEN à partir de raisonnements en moyenne. On constate que chaque point m_i^* est le centre de gravité de la densité p sur la région d'influence de i , c'est-à-dire l'ensemble des ω dont le tirage "appelle" i .

Malheureusement, la démonstration des mêmes résultats en dimension 2 s'avère à peu près inextricable bien que la démarche soit semblable. Ce qui cause la complexité des démonstrations (mais aussi l'efficacité de l'algorithme), c'est le fait que le stimulus ω soit tiré au hasard et qu'il appelle un indice i_0 (l'unité la plus proche) sans qu'il soit facile d'explicitier cette v.a. i_0 .

On peut alors penser à tirer des stimuli qui soient eux-mêmes des points du réseau, mais alors l'ensemble des (m_i) se concentre en un seul point, puisqu'à chaque présentation, on contracte les distances mutuelles.

Ceci nous a amené à proposer, dans le 2° article de cette partie, une variante de l'algorithme de KOHONEN : on tire un stimulus ω qui correspond directement à une unité i : si (m_i^t) est la situation à l'instant t et $\omega = i_0$, on définit $m_i^{t+1} = m_i^t$ si i n'est pas voisin de i_0 et

$$m_i^{t+1} = m_i^t (1 - |V|/\epsilon) + \epsilon \sum_{i' \in V(i_0)} m_{i'}^t,$$

si $i \in V(i_0)$.

Pour $i \in V(i_0)$, m_i^{t+1} est le centre de gravité de m_i^t et de $\{m_{i'}^t, i' \in V(i_0)\}$ pondérés respectivement par $1 - |V|/\epsilon$ et ϵ . Cet algorithme obéit lui aussi à la règle de Hebb.

Pour éviter la dégénérescence du réseau vers un seul point, on fixe les bords, c'est-à-dire que par exemple en dimension 1, m_1^t et m_N^t sont constamment égaux à 0 et à 1.

Alors nous démontrons, en dimension 1 ou 2, avec ϵ fixe et suffisamment petit, la convergence en loi de $m(t)$ vers un maillage aléatoire indépendant de $m(0)$, "ordonné". Nous calculons l'espérance de $m(\infty)$ et bien sûr montrons que lorsque $\epsilon = \epsilon(t)$ tend vers 0 (avec $\sum \epsilon(t) = +\infty$, $\sum \epsilon^2(t) < +\infty$), $m(t)$ tend vers $E(m(\infty))$ p.s.

Les démonstrations sont plus simples dans ce cas, car on dispose d'une écriture de $m(t+1)$ en fonction de $m(t)$, de la forme $m(t+1) = Z^{t+1} m(t) + W^{t+1}$, où Z^{t+1} , W^{t+1} sont des matrices aléatoires explicites. On utilise des résultats de KESTEN [5].

L'algorithme de KOHONEN a de nombreuses applications dans tous les domaines où s'effectue une auto-organisation : rétinotopie, reconnaissance et classification de phonèmes, construction de maillages (de dim 1 ou 2 principalement) pour "remplir" des figures de dimension supérieure, etc...

On trouve dans FORT [3] une intéressante application à la solution du problème du voyageur de commerce. Une autre jolie illustration a été donnée par RITTER et SCHULTEN [11] : si après stabilisation d'un maillage de KOHONEN, on change la loi de présentation d'un stimulus ω , il y a déformation et convergence vers un autre maillage adapté à la nouvelle loi. Cela modélise de manière particulièrement simple la réorganisation du cortex sensoriel après un traumatisme affectant une zone externe, comme le décrit par exemple MERZENICH [10].

On peut poursuivre l'étude de cet algorithme dans deux directions la démonstration de la convergence en dimension > 1 , son application dans des domaines variés (tous ceux où l'on utilise des maillages adaptatifs).

TROISIEME PARTIE.**STABILITE ET ATTRACTIVITE DANS DES RESEAUX DE MEMOIRE ASSOCIATIVE.**

(A fait l'objet d'une publication dans *Biological Cybernetics*) [1].

Ce travail porte sur les réseaux de neurones, domaine de recherche très vivant et en pleine expansion en ce moment.

Les automates cellulaires ont été introduits il y a 40 ans (MAC CULLOCH et PITTS, 1943 [5]). Mais depuis les années 1970, les recherches sur les réseaux se sont considérablement développées en informatique d'une part, (calculateurs parallèles) et biologie et physique d'autre part (réseaux de neurones, problèmes d'apprentissage, de mémoire et de reconnaissance, utilisation des outils de la mécanique statistique). La bibliographie en est très importante.

Dans un réseau de neurones, il n'y a pas correspondance bijective entre objet à mémoriser et neurone. Les neurones sont connectés les uns aux autres. Rappelons que dans le cerveau humain, on compte 10^{12} cellules nerveuses et que chacune est reliée à environ mille à dix mille autres cellules. Au cours de la présentation d'un objet à mémoriser (un patron), un certain nombre de cellules sont excitées et chaque cellule du réseau est alors codée par les réponses de toutes les autres. On suppose, en général, dans ces modèles que les connexions sont totales. Elles sont définies pendant la période d'apprentissage suivant la règle de HEBB [2] : une connexion est consolidée chaque fois que les deux cellules qu'elle relie sont dans le même état. L'apprentissage consiste en la présentation séquentielle d'une suite de patrons et aboutit à la formation d'une matrice de connexion qui caractérise alors le réseau.

La mémoire est donc distribuée : c'est tout le réseau qui participe à la reconnaissance d'un patron, et on veut qu'elle soit associative, c'est-à-dire qu'on puisse retrouver un patron à partir d'une présentation d'un objet déformé.

Beaucoup d'auteurs ont contribué à progresser dans la définition de tels réseaux : VON NEUMANN [10], ROSENBLATT et son Perceptron [9], etc... On trouve dans le livre de KOHONEN [4], outre une présentation de ses propres travaux, un bon panorama de ces recherches.

A partir de 1982, l'apport des physiciens (HOPFIELD [3]) a permis des progrès très intéressants. On considère N automates à deux états :

+ 1 (actif), - 1 (inactif), des connexions complètes et symétriques, calculées à partir des patrons à mémoriser ($\in \{-1, +1\}^N$) selon la règle de HEBB [2]. L'utilisation des outils de la mécanique statistique, des notions d'énergie et de température ont permis de grandes avancées et de multiples travaux ont été publiés. Voir la bibliographie de l'article.

En général, les patrons à mémoriser sont aléatoires, les connexions également et l'évolution du réseau est commandée par des règles déterministes ou aléatoires suivant les auteurs, consistant à favoriser l'alignement de l'état du neurone i sur ce qu'il reçoit en provenance de tous les autres soit $V_i = \sum_j C_{ij} \sigma_j$ (C_{ij} représente l'efficacité de la connexion de j vers i , et $\sigma_j = \pm 1$ est l'état du neurone j). Les résultats obtenus sont asymptotiques, quand $N \rightarrow \infty$.

Pendant comme l'ont remarqué PERSONNAZ et al. [8], les patrons ne sont pas nécessairement des points stables de l'algorithme, ce qui explique, entre autres, la capacité assez faible de ces réseaux : on peut stocker de manière fiable environ $0.14 \times N$ patrons.

Ils ont proposé dans [8] un calcul des connexions qui correspond à la règle de Hebb dans le cas orthogonal, et assure la stabilité des patrons quels qu'ils soient (on les suppose seulement linéairement indépendants).

Leur article a été le point de départ de mon travail. J'étudie l'algorithme déterministe et pour améliorer la capacité de mémoire du réseau, je propose une méthode de calcul de la matrice des connexions qui rend les patrons non seulement stables, mais attracteurs et d'une attractivité maximale. Je calcule alors cette attractivité. J'utilise pour cela une interprétation géométrique simple des inégalités qui définissent l'attractivité à distance k .

Bien sûr l'utilisation d'une température décroissante, de patrons tirés au hasard de façon à être pratiquement orthogonaux, de connexions avec effacement progressif des patrons les plus anciens [6], de multi-connexions [7], et d'autres raffinements, permettent d'améliorer les performances de ces réseaux.

QUATRIEME PARTIE.

ANALYSE MATHEMATIQUE D'UN MODELE DU CORTEX CEREBELLEUX.

La dernière partie est consacrée à l'étude d'un réseau de neurones proche de la biologie : nous sommes partis de la modélisation du cortex cérébelleux qui a été particulièrement étudié depuis CAJAL [2], jusqu'à CHAUVET [3] par exemple, en raison de sa structure plus simple que celle d'autres réseaux de neurones.

Le but est d'étudier le rôle des collatérales inhibitrices récurrentes qui relient les cellules de Purkinjia, cellules de base du cervelet. La modélisation proposée est fondée sur des travaux de physiologie (par exemple [4]) et provient du modèle établi par H. AXELRAD [1], modèle qui est trop complexe pour être analysé mathématiquement.

Le modèle étudié consiste à attribuer à chaque neurone à l'instant t un état X_i^t , de sorte que (X_i^t) est un processus de Markov. Au cours du temps X_i^t décroît linéairement, et la décharge du neurone i a lieu quand $X_i^t = 0$; cela alors provoque la surcharge de ses voisins d'une certaine quantité θ qui représente un effet inhibiteur, et la réinitialisation en i suivant une v.a. de loi donnée \mathcal{F} . On peut interpréter l'état du neurone comme le complément au seuil de décharge de son potentiel.

On étudie les propriétés mathématiques du processus de Markov (X^t) . Il est récurrent tant que θ est plus petit qu'un certain seuil θ_0 (inversement proportionnel à la taille des voisinages considérés), il est transient quand $\theta > \theta_0$.

Les difficultés viennent du fait que l'ensemble des états possibles est dense dans $(\mathbb{R}^+)^N$ et que toutes les composantes de (X^t) sont interdépendantes. La démonstration de la récurrence utilise des arguments techniques adaptés à la définition même de la chaîne. On démontre d'abord l'irréductibilité et l'apériodicité du processus. Puis on établit une majoration de l'état en i par l'état qui serait observé si le neurone i était inhibé par des voisins eux-mêmes non inhibés. On associe alors en chaque point voisin de i une marche aléatoire de loi \mathcal{F} correspondant à l'activité "propre" en ce site. On peut utiliser la loi des grands nombres, des inégalités type grandes déviations et des arguments de renouvellement pour ces marches aléatoires. On calcule ensuite quand la chaîne est ergodique, la fréquence moyenne de décharge de chaque neurone. Dans ce cas, l'inhibition provoque simplement un ralentissement général de l'activité.

Dans le cas transient, on observe toujours une récurrence sur un sous-réseau, tandis que sur le complémentaire les états tendent vers $+\infty$. Il apparaît alors une réponse périodique du réseau, une sorte d'échantillonnage spatial dont la période est liée à la taille du voisinage et à la force de l'inhibition.

C'est donc en augmentant l'inhibition, qu'on provoque des réponses en bandes ou en moirures, réponses qui sont caractéristiques des réseaux neuronaux biologiques [7].

Si de plus on superpose au réseau des stimuli (en excitant certains neurones, c'est-à-dire en les réinitialisant après chaque décharge par une v.a. de loi \mathcal{F}' avec $E(\mathcal{F}') < E(\mathcal{F})$), on observe les mêmes phénomènes. Bien sûr, l'échantillonnage spatial s'adapte alors à la forme des zones excitées.

Ces démonstrations sont illustrées par les nombreuses simulations faites et sont en accord avec les résultats expérimentaux obtenus par simulation par J. HERAULT [6] en traitement du signal et avec les résultats de neuro-physiologie : cartes d'orientation préférentielle, de dominance oculaire du cortex visuel par exemple... [7].

Cette étude "de l'intérieur du réseau" apporte une meilleure compréhension du rôle de l'inhibition, et en particulier du fait que l'inhibition latérale (zone inhibée au voisinage immédiat d'un neurone excité) éclate quand l'inhibition augmente.

Les techniques de chaînes de Markov de grandes dimensions utilisées permettent de décrire le comportement du réseau dans sa globalité. Jusqu'à ces dernières années, on ne mesurait l'activité électrique que d'un ou deux neurones. Mais la possibilité d'enregistrer simultanément l'activité de populations nombreuses de neurones donne toute sa place à une modélisation globale [5]. Ces modèles n'en sont probablement qu'à leurs débuts.

De plus il me semble que ce modèle est riche d'interprétations et d'applications dans d'autres contextes que celui du cervelet, en cristallographie par exemple. La forme des réponses divergentes obtenues dans le cas transient peut sans doute être mise en rapport avec la théorie des graphes.

BIBLIOGRAPHIE

PREMIERE PARTIE

- [1] N. ABRAHAMSON : "The ALOHA system-another alternative for computer communications"
Proc. Fall Joint Computer Conf. AFIPS Press, Vol. 37, p. 281-285 (1970).
- [2] R. AZENCOTT et G. RUGET : "Mélanges d'équations différentielles et grands écarts à la loi des grands nombres"
Z. Wahrscheinlichkeitstheorie, vol. 38, p. 1-54 (1977).
- [3] A. BENVENISTE, M. GOURSAT et G. RUGET : "Analysis of stochastic approximations schemes with discontinuous and dependent forcing terms with applications to data communication algorithms"
IEEE Trans. Aut. Control. Vol. AC-25, p. 1042-1058 (1980).
- [4] J. BRETAGNOLLE : "Grandes déviations et applications statistiques"
Astérisque 68, Soc. Math. France, p. 19-25 (1979).
- [5] O. de CAMBRY : "A general method for constructing statistical failure detection test"
A paraître dans IEEE Trans. Aut. Control.
- [6] M. COTTRELL, J.C. FORT et G. MALGOUYRES : "Large deviations and rare events in the study of stochastic algorithms"
IEEE Trans. Aut. Control. Vol. AC-28, n° 9, p. 907-920 (1983).
- [7] P. DUPUIS et H.J. KUSHNER : "Stochastic approximations via large deviations : Asymptotic properties"
SIAM J. on Control and Optimisation, 23, p. 675-696 (1985).
- [8] D. GAUER et G. THOMSON : "Programming and probability models in operations research"
Monterey, CA, Brooks/Cole (1972).
- [9] L. KLEINROCK et S.J. LAM : "Packet switching in a multiaccess broadcast channel : performance evaluation"
IEEE Trans. Comm. Vol. COM-23, n° 4, p. 410-423 (1984).
- [10] S. LAM : "Packet switching in a multiaccess broadcast channel"
SEAS, Univ. Calif. (1964).
- [11] A.D. VENTSEL et M.I. FREIDLIN : "Random perturbations of dynamical systems"
Springer-Verlag, New York (1984)

DEUXIEME PARTIE.

- [1] M. COTTRELL et J.C. FORT : "Etude d'un processus d'auto-organisation"
Ann. Inst. Henri Poincaré, vol. 23, n° 1, p. 1-20 (1987).
- [2] M. COTTRELL et J.C. FORT : "A stochastic model of retinotopy: A self organizing process"
Biol. Cybern. 53, p. 405-411 (1986).
- [3] J.C. FORT : "Solving a combinatorial problem via self-organizing process : an application of the Kohonen algorithm to the traveling salesman problem"
A paraître dans Biol. Cybern. (1988).
- [4] D.O. HEBB : "The organization of behaviour"
Wiley, New York (1949).
- [5] H. KESTEN : "Random difference equations and renewal theory for products of random matrices"
Acta. Math. 131, p. 207-248 (1973).
- [6] T. KOHONEN : "Self-organized formation of topologically correct feature maps"
Biol. Cybern. 43, p. 59-69 (1982 a)

- [7] T. KOHONEN : "Clustering, taxonomy and topological maps of patterns"
Proc. of the 8th Intern. Conf. on Pattern Recognition,
p. 114-128 (1982 b).
- [8] T. KOHONEN : "Analysis of a simple self-organizing process"
Biol. Cybern. 44, p. 135-140 (1982 c).
- [9] T. KOHONEN : "Self-organization and associative memory"
Springer, Berlin-Heidelberg-New York-Tokyo (1984).
- [10] M.M. MERZENICH : "Development and maintenance of cortical somatosensory representations : Functional "maps" and neuro-anatomical repertoires"
Touch, K. Barnard and T.B. Brazelton, eds, Intern. Univ. Press. New York (1985).
- [11] H. RITTER et K. SCHULTEN : "On the stationary state of Kohonen's self-organizing sensory mapping"
Biol. Cybern. 54, p. 99-106 (1986).

TROISIEME PARTIE.

- [1] M. COTTRELL : "Stability and attractivity in associative memory networks"
Biol. Cybern. 58, p. 123-139 (1988).
- [2] D.O. HEBB : "The organization of behavior"
John Wiley & Sons, New York (1949).
- [3] J.J. HOPFIELD : "Neural networks and physical system with emergent collective computational abilities"
Proc. Natl. Acad. Sci. USA, 79, p. 2554-2558 (1982).
- [4] T. KOHONEN : "Self-organization and associative memory"
Springer, Berlin-Heidelberg, New York-Tokyo (1984).
- [5] W.W. MAC CULLOCH et W. PITTS : "A logical calculus of the ideas immanent in nervous activity"
Bull. Math. Biophys. 5, p. 115-133 (1943).
- [6] J.P. NADAL, G. TOULOUSE, J.P. CHANGEUX et S. DEHAENE : "Networks of formal neurons and memory palimpsests"
Europhys. Lett. 1 (10), p. 535-542 (1986).
- [7] P. PERETTO et J.J. NIEZ : "Long term memory storage capacity of multiconnected neural networks"
Biol. Cybern. 54, p. 53-63 (1986).
- [8] L. PERSONNAZ, I. GUYON et G. DREYFUS : "Information storage and retrieval in spinglass like neural networks"
J. Phys. Lett. 46, p. 359-365 (1985).
- [9] F. ROSENBLATT : "Principles on neuro dynamics"
Psych. Rev. 65, p. 386 (1958).
- [10] J. VON NEUMANN : "Theory of self-reproducing a automate"
A. W. Burks, Univ. of Illinois Press. Urbana (1966).

QUATRIEME PARTIE.

- [1] H. AXELRAD, C. BERNARD, M. COTTRELL et B. GIRAUD : "The use of an artificial neural network to analyse the informational transfer properties of a simplified model of the cerebellar cortex"
Proc. of the First ICNN . A paraître (1987).
- [2] S.R. y CAJAL : "Histologie du système nerveux de l'Homme et des Vertébrés" II.
Maloine, Paris (1911).
- [3] G. CHAUVET : "Habituation rules for a theory of the cerebellar cortex"
Biol. Cybern. 55, p. 201-209 (1986).

- [4] F. CREPEL, N. DELHAYE-BOUCHAUD, J.L. DUPONT et C. SOTE LO :
Neuroscience, vol. 5, p. 333-347 (1980)
- [5] G. GERSTEIN, D. PERKEL et J. DAYHOFF : "Cooperative firing
activity in simultaneously recorded populations of neurons :
Detection and measurement"
J. of Neuroscience, vol. 5, n° 4, p. 881-889 (1985).
- [6] J. HERAULT : "Le traitement de l'information dans les structu-
res nerveuses"
Thèse de Doctorat d'Etat. Université de Grenoble (1980).
- [7] D.H. HUBEL et T.N. WIESEL : "Receptive fields, binocular
interaction and functional architecture in the cat's visual
cortex"
J. Physiol. (Lond.), 160, p. 106-154 (1962).

PREMIERE PARTIE

GRANDES DEVIATIONS ET EVENEMENTS RARES DANS
L'ETUDE D'ALGORITHMES STOCHASTIQUES

Large Deviations and Rare Events in the Study of Stochastic Algorithms

MARIE COTTRELL, JEAN-CLAUDE FORT, AND GÉRARD MALGOUYRES

Abstract—New asymptotics formulas for the mean exit time from an almost stable domain of a discrete-time Markov process are obtained.

An original fast simulation method is also proposed. The mathematical background involves the large deviation theorems and approximations by a diffusion process. We are chiefly concerned with the classical Robbins–Monroe algorithm. The validity of the results are tested on examples from the ALOHA system (a satellite type communication algorithm).

INTRODUCTION

IN this paper we study the exit time from an almost stable domain of a discrete-time Markov process. This belongs to the reliability theory and we will develop such applications. But one can encounter the same problems in population dynamics and in the physical theory of metastability (in the work of Penrose and Lebowitz in a much more complex context [19]; for population dynamics, see [18]).

We have chosen to restrict our domain of investigation but to explore it with some accuracy. We will only consider Markov chains with discrete time, making “small” jumps over \mathbb{R}^m . Typically, this is the case of the well-known Robbins–Monroe algorithm with constant gain ϵ : $X_{n+1} = X_n + \epsilon V(X_n, \omega)$, where ϵ is small and $V(x, \cdot)$ is a random increment at x . Our purpose is to describe the behavior of such an algorithm which is claimed convergent after looking at the mean-differential equation (ODE); for example, let $b(x)$ be the mean value of $V(x, \cdot)$ and assume that $b(x)$ has the opposite sign of x . Then 0 is a stable equilibrium point of the ODE $dx/dt = b(x)$. However, in almost all cases ($V(x, \cdot)$ is Gaussian or Poisson distributed, etc.), fixing a barrier $a > 0$, there is a probability 1 that X_n passes through after a sufficiently large time. So it is of interest to know how long it takes to reach a well-chosen barrier, and to know how it happens. This is what we solve in simple cases and give the tools to solve in more general cases. (Note that the same methods could be used when there is dependence between the successive $V(X_n, \cdot)$ (see [20]) or when X is a continuous time process.)

Working on some problems about transmission algorithms we give applications in an ALOHA system. We will describe ALOHA later on. But let us show another possible application: the phase-locked loop [3].

Messages are emitted from a source; each message is a sequence $(a_n)_{n \in \mathbb{Z}}$ with $a_n = e^{i\phi_n}$, $\phi_n \in ((2\pi/N) + (2k\pi/N), k \in \mathbb{Z})$. This means that the phase ϕ_n contains all the information. The received messages are sequences $(y_n)_{n \in \mathbb{Z}}$ with $y_n = \sum_{k \in \mathbb{Z}} s_k a_{n-k} + v_n$ where the channel characteristics $S = (s_k)_{k \in \mathbb{Z}}$ and the noise sequence $(v_n)_{n \in \mathbb{Z}}$ take complex values; that is to say that y_n is a

mixing of the signal $a_n, a_{n-1}, a_{n-2}, \dots$. The problem is to estimate $(\phi_n)_{n \in \mathbb{Z}}$, the phases of $(a_n)_{n \in \mathbb{Z}}$. A classical estimator (Boucle de Costas [3]) is $(\text{Im } z)$ is the imaginary part of z)

$$\hat{\phi}_{n+1} = \hat{\phi}_n - \epsilon V_n(\hat{\phi}_n)$$

where

$$V_n(\phi) = \mathfrak{I}_m(y_n^N e^{-iN\phi})$$

the corresponding ODE is (assuming stationarity) $d\phi/dt = -b(\phi)$ where $b(\phi)$ is the mean value of $\mathfrak{I}_m(y_n^N e^{-iN\phi})$.

The right-hand side of the ODE is derived from a periodic potential with a minimum value in each interval $[(2k-1)\pi/N, (2k+1)\pi/N]$. So when the phase ϕ_n is in one of the darkened domains of Fig. 1 there is a large probability of erroneous interpretation and we observe a packet of errors.

We are interested in the mean exit time from ϕ_1^* to the next attraction domain. An immediate application of this work gives an explicit formula (when $s_k = 0$ except s_0) closer to reality than those given by a standard diffusion approximation.

A possible extension would give such a formula in more general cases, assuming good properties of ergodicity.

The content of each section is as follows. Section I contains some asymptotic well-known statements describing at different degree of precision the behavior of the Robbins–Monroe algorithm as $\epsilon \rightarrow 0$. The two main results are exposed in Sections II and III. In Section II we give the principle of a probability change used for quick simulation (importance sampling) and prove its best optimality among a large class. This solves an important problem which is encountered in simulation practice. Section III provides an analytic formula for the mean exit time from a quasi-stable domain using an approximation by a diffusion process, following the previously mentioned probability change. Finally, in Section IV we apply the two preceding results to sharing time algorithms issued from the study of the ALOHA system. We submit that the numerical results are quite conclusive. Two appendixes complete this paper. In Appendix I we present the definitions and the basic mathematical theory useful to understand the text. In Appendix II we give indications about the proofs of the theorem of Section II.

I. LARGE DEVIATIONS

We examine algorithms with discrete-time small constant gain defined by

$$X_0^s = x_0 \in \mathbb{R}^m$$

$$X_{n+1}^s = X_n^s + \epsilon V_n(X_n^s, \omega)$$

$X_n^s \in \mathbb{R}^m$, that is to say X_n^s is a time homogeneous Markov chain, with small increments (see [4]).

Manuscript received June 15, 1981; revised March 29, 1982 and June 7, 1982. Paper recommended by S. I. Marcus, Past Chairman of the Stochastic Control Committee.

M. Cottrell and G. Malgouyres are with the Université Paris-Sud, Orsay, France.

J.-C. Fort is with the Université de Nancy, Nancy Cedex, France.

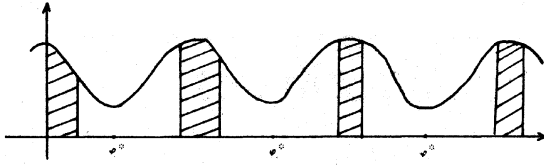


Fig. 1.

The mathematical results will be asymptotic when $\epsilon \rightarrow 0$.

Let μ_x be the distribution of V_n if $X_n^* = x$, and $b(x) = EV_n(x, \omega) = E(\mu_x) = \int y d\mu_x(y)$ be the expectation of μ_x (Note that $b(x)$ does not depend upon n because of the homogeneity of the process.)

A. General Hypotheses

In all this work, we need some general hypotheses about the process (X_n^*) ; the distribution μ_x must vary slowly between the instant n and the instant $n+1$. More precisely, we need the following.

- 1) $E(\exp(t, y) d\mu_x(y)) < +\infty$ for each x , i.e., μ_x has a finite Laplace transform (see Appendix I-A-1).
- 2) $d(\mu_y, \mu_x) \leq C\|y - x\|$ where d is the Prohorov distance [5], i.e., μ_x is a Lipschitz function of x .

These hypotheses are currently true in many usual recursive algorithms (see [3]).

The space Ω is the space of the right-continuous functions with left-hand limits (cad-lag) $\mathbb{R}^+ \rightarrow \mathbb{R}^m$, endowed with the Borel σ -algebra for the Skorohod topology [5] and the probability P . We denote by P^ϵ the probability associated with the process (X_n^*) . See [4] for a more precise definition.

The first result is a law of large numbers.

1) *Law of Large Numbers:* Let $x^\circ(t)$ be the deterministic trajectory solution of the ordinary (or mean) differential equation (ODE)

$$\begin{aligned} \frac{dx^\circ(t)}{dt} &= b(x^\circ) \\ x^\circ(0) &= x_0 \end{aligned}$$

if $X_0^* = x_0$, then

$$\forall \eta > 0, \forall T < +\infty, P\left(\max_{0 \leq n \leq T} |X_n^* - x^\circ(\epsilon n)| > \eta\right) \rightarrow 0$$

when $\epsilon \rightarrow 0$ (see [7] for technical assumptions).

Since the space scale is small (the increments are multiplied by ϵ), we get a result where two scales of time appear: n for X_n^* , and $n\epsilon$ for $x^\circ(t)$. So let us contract the time, considering that the process jumps at instants $t = n\epsilon$, and denote $X^\epsilon(t) = X_{t/\epsilon}^*$. Then this first result means that the process $X^\epsilon(t)$ converges uniformly in each interval $[0, T]$ to the deterministic process $x^\circ(t)$. The equivalence $t \leftrightarrow n\epsilon$ defines the change of time, " n corresponds to t " means that X_n^* is the point near $x^\circ(t)$.

This result allows the application of the ODE method [4].

A first evaluation of the deviation between the (contracted in time) process $X^\epsilon(t)$ and its deterministic limit is given by the next theorem.

2) *Central Limit Theorem:* Let us define $Y^\epsilon(t) = (X^\epsilon(t) - x^\circ(t))/\sqrt{\epsilon}$. Then, with some technical assumptions (see [7]), the process $Y^\epsilon(t)$ converges weakly when $\epsilon \rightarrow 0$, on $[0, T]$ to $Y^\circ(t)$, the solution of the stochastic differential equation

$$dY^\circ(t) = \frac{\partial b}{\partial x}(x^\circ(t))Y^\circ(t) dt + \Gamma(x^\circ(t)) dW(t)$$

where $W(t)$ is the Brownian motion and $\Gamma(x^\circ(t))$ is the variance-covariance matrix at instant t of the distribution $\mu_{x^\circ(t)}$ (see [15]).

Even if these theorems indicate the convergence of $X^\epsilon(t)$ to $x^\circ(t)$, it is possible to observe rare trajectories where $X^\epsilon(t)$ is far from $x^\circ(t)$. These are very small probability events, explained by the distribution tails, which are systematically replaced by normal tails in the central limit theorem.

3) *Large Deviation Results:* We recall in Appendix I the main definitions and results we need.

We denote by $\hat{\mu}_x$ the Laplace transform of the law μ_x , l_x its logarithm, and h_x its Cramer transform, $h_x(u) = \sup_s \langle \langle s, u \rangle - l_x(s) \rangle$.

Let \mathcal{C}_T (resp. \mathcal{C}) be the set of the continuously piecewise differentiable applications $\varphi: [0, T]$ (resp. \mathbb{R}^+) $\rightarrow \mathbb{R}^m$, such that $\varphi(0) = x_0$ fixed. The elements of \mathcal{C}_T (and \mathcal{C}) are called *paths*.

If $\varphi \in \mathcal{C}_T$, the action integral along φ is

$$I(\varphi) = \int_0^T h_{\varphi(t)}(\dot{\varphi}(t)) dt$$

where $\dot{\varphi}(t)$ is the derivative of φ at the point t .

First we give an equivalent (when $\epsilon \rightarrow 0$) of the probability to stay inside a tube along a path φ .

Theorem 1 [21], [22]: Let $\delta > 0$, φ be a path of \mathcal{C}_T . Let $T_\delta^2(\varphi)$ be a tube around φ , with diameter δ ; that is, the set of the trajectories of $X^\epsilon(t)$, issued from x_0 , such that

$$\forall t \in [0, T], |X^\epsilon(t) - \varphi(t)| < \delta.$$

Then, there exists δ_0 such that for $0 < \delta < \delta_0$, we have

$$\lim_{\epsilon \rightarrow 0} (-\epsilon \log P^\epsilon(T_\delta^2(\varphi))) = I(\varphi) + \alpha(\delta)$$

with

$$\lim_{\delta \rightarrow 0} \alpha(\delta) = 0.$$

This result does not permit us to calculate $P^\epsilon(T_\delta^2(\varphi))$, even for a small ϵ , but it permits us to discriminate the tubes: as soon as $I(\varphi)$ is different for two tubes, their probabilities are very different for small ϵ . The value $\exp(-I(\varphi)/\epsilon)$ is a *measure of the resistance of the process to follow the path φ* ; if φ is the solution of the mean differential equation (ODE), then $\dot{\varphi}(t) = E(\mu_x) = b(x)$ and $h_{\varphi(t)}(\dot{\varphi}(t)) = 0$ for each t , so $I(\varphi) = 0$.

For establishing Theorem 1, some assumptions are necessary. One can find them in [7], but the following are of interest.

The Cramer transform must slowly vary between two close points x_1 and x_2 (it is nearly the same as the second one of the general hypotheses).

The Cramer transform $h_x(u)$ increases at infinity faster than $\|u\|$. This expresses that the tail of the distribution μ_x is not too big.

A complete proof can be found in [21], [2].

Corollary 1 [21], [2]: With the same assumptions, let B be a measurable set of trajectories such that $\inf\{I(\varphi)/\varphi \in B \cap \mathcal{C}_T\} = \inf\{I(\varphi)/\varphi \in \bar{B} \cap \mathcal{C}_T\}$, then

$$\lim_{\epsilon \rightarrow 0} (-\epsilon \log P^\epsilon(B)) = \inf_{\varphi \in B \cap \mathcal{C}_T} I(\varphi).$$

Let us comment on this corollary. Generally, we apply it to events B such as exists from an attraction pool, reaching an exterior region \mathcal{E} . (See Fig. 2.)

The exit probability is the sum of the probability of different exit tubes, approximately

$$\sum_k \exp(-I(\varphi_k)/\epsilon)$$

but only the smallest coefficient $I(\varphi_k)$ is important (because of

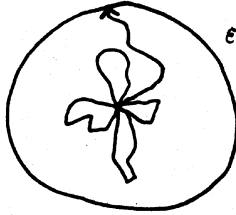


Fig. 2.

the exponential). So the exit will follow approximately some optimal path φ_{opt} , whose action integral is minimum.

Optimal Path and Probability Change: To determine the optimal exit path φ_{opt} (starting from x_0) is therefore to minimize the action integral $\int_0^T h_{\varphi(t)}(\dot{\varphi}(t)) dt$ among the paths φ such that $\varphi(0) = x_0$, $\varphi(T) \in \mathcal{E}$ (T is free). It is a dynamical programming problem to solve.

The main tool of these demonstrations is a *probability change*: if we want to reach \mathcal{E} with the greatest probability, we need to follow φ_{opt} , and for that, transform the probability measure μ_x into $\tilde{\mu}_x$ in such a way that the transformed mean field will be tangent to φ_{opt} .

More precisely, $\tilde{\mu}_x$ is defined by $d\tilde{\mu}_x(y) = (e^{\lambda_x y} d\mu_x(y)) / (e^{\lambda_x \varphi_{opt}(t)})$, λ_x being chosen in such a way that if $x = \varphi_{opt}(t)$, $\tilde{b}(x) = E(\tilde{\mu}_x) = \dot{\varphi}_{opt}(t)$; that is to say

$$\frac{\int y e^{\lambda_x y} d\mu_x(y)}{\tilde{\mu}_x(\lambda_x)} = \frac{\tilde{\mu}'_x(\lambda_x)}{\tilde{\mu}_x(\lambda_x)} = l'(\lambda_x) = \dot{\varphi}_{opt}(t).$$

The path φ is a solution of the new ODE $dx/dt = \tilde{b}(x)$. Further, we will denote by $\tilde{\mu}_x$, $\tilde{X}^x(t)$, \tilde{P}^x , and \tilde{P} the transformed form of μ_x , $X^x(t)$, P^x , and P , respectively.

4) *The One-Dimensional Case:* In the one-dimensional case it is very easy to find the optimal path φ_{opt} .

Let be $m = 1$, $X^x(0) = 0$, and $E(\mu_x) = b(x)$ with the sign of $(-x)$ for each x . Thus, in Fig. 3 the point 0 is an attractor. Let $a > 0$, and B be the set of trajectories ω starting from 0, reaching a before coming back to 0.

In this case, the exit point is necessarily a . And by Corollary 1, we have

$$\lim_{\epsilon \rightarrow 0} (-\epsilon \log P^\epsilon(B)) = \inf_{T, \varphi} \{ I(\varphi) / \varphi(0) = 0, \varphi(T) = a, \varphi \in B \}.$$

To find this minimum, we can restrict ourselves to injective φ ($\dot{h} \geq 0$). Then the variable change $x = \varphi(t)$ leads to

$$I(\varphi) = \int_0^T h_{\varphi(t)}(\dot{\varphi}(t)) dt = \int_0^a \frac{h_x(\dot{x})}{\dot{x}} dx$$

where $\dot{x} = \dot{\varphi}(\varphi^{-1}(x))$.

The minimum of $I(\varphi)$ is obtained by a point-by-point minimization of $h_x(\dot{x})/\dot{x}$. Hence, the optimal path φ_{opt} is defined by

$$\frac{h_{\varphi_{opt}(t)}(\dot{\varphi}_{opt}(t))}{\dot{\varphi}_{opt}(t)} = \inf_{u > 0} \frac{h_{\varphi(t)}(u)}{u}$$

which is the slope of the tangent Δ issued from 0 to the graph of $u \rightarrow h_{\varphi(t)}(u)$. (See Appendix I-A-3-a.)

If $\chi(\varphi(t))$ is the point where $h_{\varphi(t)}(\chi(\varphi(t)))/\chi(\varphi(t))$ is minimum, and equal to $\lambda_{\varphi(t)}$, the optimal path is hence the

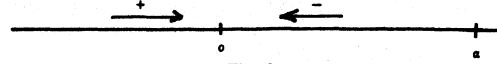


Fig. 3.

solution of

$$\frac{d\varphi}{dt} = \chi(\varphi), \quad \varphi(0) = 0.$$

Then $I(\varphi_{opt}) = \int_0^a \lambda_x dx$ and λ_x satisfies the equation $h'_x(\chi(x)) = \lambda_x$, thus

$$l_x(\lambda_x) = h'_x(\chi(x)) \cdot \chi(x) - h_x(\chi(x))$$

(because of the reciprocity of the Cramer transformation (Appendix I-A-3-a), l_x is h_x transformed), hence $l_x(\lambda_x) = 0$, by definition of $\chi(x)$ because the tangent passes through 0.

So we get

$$\lim_{\epsilon \rightarrow 0} -\epsilon \log P^\epsilon(B) = \int_0^a \lambda_x dx$$

where λ_x is a solution of $l_x(\lambda_x) = 0$ and $\lambda_x > 0$.

We actually have $l'_x(\lambda_x) = l'_x(h'_x(\chi(x))) = \chi(x)$ (property of the convex duality: Appendix I-A-3-d), and $l'_x(\lambda_x) = \dot{\varphi}_{opt}(t)$.

The probability change $\mu_x \rightarrow \tilde{\mu}_x$, defined by the coefficient λ_x is therefore that one which centers the measure "around" the optimal path.

II. QUICK SIMULATION METHOD

If A is a rare event in relation with some Markov chain, it is quite impossible to estimate $P(A)$ by a direct simulation, for the length of the trajectories would require a frightfully large time of computation. Moreover, the time of simulation would be larger than the period of the pseudorandom numbers generator (generally $\approx 2^{32} - 5 \cdot 10^9$). For example, if $P(A)$ is estimated from an n -sample (X_1, \dots, X_n) of Bernoulli distributed variables (X_i are computed from Markov chain trajectories) by

$$\widehat{P(A)} = \frac{1}{n} \sum_{i=1}^n X_i,$$

its standard deviation is $\sqrt{P(A)(1-P(A))/n}$. To have a 20 percent error on $P(A)$ with 95 percent of confidence, we must take $\sqrt{1-P(A)}/\sqrt{nP(A)} \leq 0,1$ and then for $P(A) = 10^{-4}$, we need a 10^6 sample.

So our idea is to use the probability change exhibited in the proof of Theorem 1 to realize a fast simulation. First, we recall the principle of the importance sampling method of simulation [13]: if P is absolutely continuous with respect to a probability P^* , we observe that $P(A) = \int_A dP/dP^*(\omega) dP^*(\omega)$. Thus, we can carry out an n -sample $\omega_1, \dots, \omega_n$ of distribution P^* such that $P^*(A) \gg P(A)$ and take the estimator (unbiased and convergent)

$$\widehat{P(A)} = \frac{1}{n} \left(\sum_{i=1}^n \frac{dP}{dP^*}(\omega_i) \mathbf{1}_A(\omega_i) \right)$$

whose standard deviation is $(\int_A (dP/dP^*(\omega))^2 dP^*(\omega) - P(A)^2)^{1/2}$ which is smaller than the preceding if $dP/dP^*(\omega) < 1$ on A . In fact, the problem is how to choose the new probability P^* . If it is badly chosen it can give a very bad estimator, that means with a great variance.

We solved this problem in a particular situation (we keep the notations of the previous section): $X^\epsilon(t)$ is a Markov chain over \mathbb{R} and $\epsilon\mu_x$ is the jump distribution at x , $b(x)$ is the mean value of μ_x . We suppose that $b(x)$ has the sign of $(-x)$, so that $x=0$ is a stable equilibrium point of the ODE. We define A by (a is a real positive barrier)

$$A = \{\omega \in \Omega | \omega(0) = 0, (\forall t \in \mathbb{R}_+^*) \omega(t) > 0, \\ (\exists t \in \mathbb{R}_+^*) \omega(t) \geq a\}.$$

A is the event (X passes through a before coming back to 0).

Our result (Theorem 2) is that the probability change used in the large deviation theory is asymptotically ($\epsilon \rightarrow 0$) the best. To prove it we apply the results of Section I. So let us summarize the natural conditions required for a good approximation of $\epsilon \rightarrow 0$ by ϵ fixed as follows.

" μ_x is a Lipschitz function of x ": the jump distribution has slow variations in the time. For example, the variation is small from x to $x + \epsilon b(x) \pm \sqrt{\epsilon} \sigma(x)$. ($\sigma(x)$ is the standard deviation of μ_x .)

ϵ is "small": the ODE is a good first approximation of $X^\epsilon(t)$: the ratio $\sqrt{\epsilon} \sigma(x)/a$ is sufficiently small ($\sigma(x)$ has slow variations).

Note that the choice of the probability change of the large deviations $P \rightarrow \tilde{P}$ is quite natural, for we know that $P^\epsilon(A)$ is exponentially (as $\epsilon \rightarrow 0$) concentrated near φ_{opt} which gives the minimum of $I(\varphi)$, $\varphi \in A$, and that $P \rightarrow \tilde{P}$ turns φ_{opt} into the solution of the new (for \tilde{P}) ODE (starting from any $x > 0$).

Thus, we are sure that $\tilde{P}(A) \gg P(A)$.

This remark is still valid in any case, and the change $P \rightarrow \tilde{P}$ is always asymptotically optimal (even if we do not write it down in a more general situation) (see another use in [7]). We give a definition and Theorem 2.

Definition: An exponential change of probability of a Markov chain over \mathbb{R} (we can easily extend to \mathbb{R}^m) $X(t)$ is defined by a bounded continuous function $x \rightarrow v_x$ from \mathbb{R} to \mathbb{R} . The transformed chain $X^*(t)$ is given by its jump distribution at x , μ_x^*

$$d\mu_x^*(y) = \frac{e^{v_x \cdot y}}{\hat{\mu}_x(v_x)} d\mu_x(y).$$

P^* is the corresponding probability on Ω .

In that sense the transformation $P \rightarrow \tilde{P}$ is obtained for $v_x = \lambda_x$, where $I_x(\lambda_x) = 0$, and $\lambda_x \neq 0$ for $x \in]0, a[$.

Theorem 2: Among all the exponential changes of probability the transformation $P \rightarrow \tilde{P}$ is asymptotically optimal in the sense of the variance:

$$\lim_{\epsilon \rightarrow 0} \int_A \left[\frac{dP^\epsilon}{d\tilde{P}^\epsilon}(\omega) \right]^2 d\tilde{P}^\epsilon(\omega)$$

is minimum. (A is the event we defined above.)

This result solves the difficult problem of the right choice for the change of probability in the fast simulation practice.

Concerning the optimality of this transformation, we want to emphasize the following points.

1) One only needs to compute the λ_x 's, the solutions of $I_x(\lambda_x) = 0$. In general, this is not a big problem, even if $\hat{\mu}_x$ is not known. In the example of a process which jumps of $+1$ or -1 with probability $p(x)$, $1-p(x)$, $\lambda_x = \log |1-p(x)/p(x)|$, thus we just have to evaluate the parameter of a Bernoulli distribution which is a very simple task and not too expensive.

2) Even if the computation of λ_x is not very precise, we hope the simulation will be greatly accelerated (but be careful: a large λ_x gives an illusory feeling of acceleration).

3) Finally, one last remark: in the actual simulation we must not forget that the tails of the jump distributions are very important.

We put the proof of Theorem 2 in Appendix II.

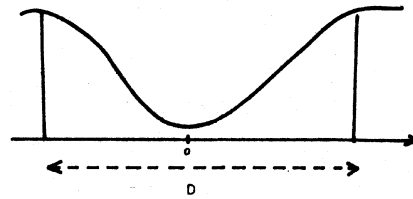


Fig. 4.

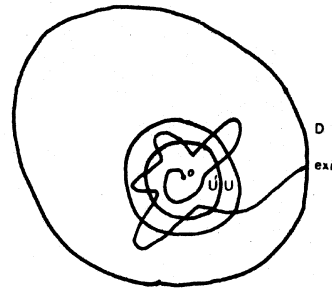


Fig. 5.

III. APPROXIMATION BY DIFFUSION

A. Introduction

The simulation methods are of much interest because of their accuracy.

However, if an approximate value is sufficient (for example, with an error < 20 percent), we are tempted to use simple and quickly computable formulas.

First we expose the general concept of our approximation by diffusion in a general context too. Then we derive analytical formulas in the one-dimensional case (as in Section II).

Thus, suppose that the ODE $dx/dt = b(x)$ derives from a potential V , with $V(0)$ minimum (see Fig. 4). Let D be the attraction domain of 0. Its boundary ∂D is supposed sufficiently regular. The large deviation theorems tell us that to exit from D starting from 0 is almost the same as to follow the path minimizing $I(\varphi)$ with $\varphi(0) = 0$ and $\varphi(T) \in \partial D$, T is free. At this step, we only get an evaluation of the logarithm of the probability to exit from D without returning near 0 (see Section I). We want to overcome this difficulty using a good diffusion approximation, in contrast to what we will call a standard one. Recall that our goal is to give an analytical formula for the mean exit time from D . We propose to proceed as in the simulation, except that we must take into account that we use continuous approximations. So let U and U' be small neighborhoods of 0, and ∂U and $\partial U'$ their boundaries, respectively, with $U' \subset U$. (See Fig. 5.)

First we estimate the probability $p(x)$, starting from $x \in \partial U$ to pass through ∂D before $\partial U'$ (as we do by simulation for the probability of passing through a without returning to 0). Here in fact we obtain equivalence when $\partial U' \rightarrow \partial U$.

Second, in U we approximate our process by a simple diffusion process, such as (OU) the Ornstein-Uhlenbeck process (see [11]), which is very close to the real process.

Then, noting that the time spent in U is very large compared to the time spent in $D \setminus U$, we neglect the sojourn time in $D \setminus U$.

We can now give a condition on the behavior of the OU process at the boundary ∂U ; at $x \in \partial U$ the process is reflected with a probability $1-p(x)$ and absorbed with a probability $p(x)$. The mean exit time from D is thus approximated by the mean time before extinction (or absorption).

Let us now give actual calculations in the one-dimensional case.

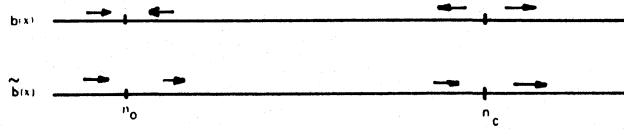


Fig. 6.

The notations are still the same except for small modifications.

We write $\tilde{\mu}(x, s)$ for $\hat{\mu}_s(x, s)$ and the subscripts indicate partial derivatives. We recall that

$$b(x) = E(\mu_x) = \hat{\mu}_s(x, 0) = \frac{\partial}{\partial s} \hat{\mu}(x, 0)$$

$$\sigma^2(x) = \text{var}(\mu_x) = \hat{\mu}_{ss}(x, 0) - [\hat{\mu}_s(x, 0)]^2.$$

Let us put n_0 for 0 (stable equilibrium point of the ODE) and let us suppose that n_c is an unstable equilibrium point, and that $b(x)$ has the same sign as $(x - n_0)(x - n_c)$. (These notations will be used in Section IV.)

The problem we examine is very close to the one of Section II, so we introduce the probability change of large deviations and the corresponding process \tilde{X}^ϵ with jump distributions $\tilde{\mu}_x$

$$\tilde{b}(x) = E(\tilde{\mu}_x) = \tilde{\mu}_s(x, 0) = \tilde{\mu}_s(x, \lambda(x))$$

$$\tilde{\sigma}^2(x) = \text{var}(\tilde{\mu}_x) = \tilde{\mu}_{ss}(x, \lambda(x)) - [\tilde{\mu}_s(x, \lambda(x))]^2$$

where $\lambda(x)$, for $x \in [n_0, n_c]$ is the nonnull solution of $\hat{\mu}(x, \lambda(x)) = 1$ and is 0 for $x \notin [n_0, n_c]$. (See Fig. 6.) We use the Landau notations o and O . First we introduce the principle of a diffusion approximation of the process $X^\epsilon(t)$ in a direct way.

B. Approximation of a Process by a Diffusion

If the process $X^\epsilon(t)$ verifies the assumptions of the law of large numbers, it converges, when $\epsilon \rightarrow 0$, to the deterministic process x^o which is the solution of the ODE. Also if $x \mapsto b(x)$ is C^2 and $x \mapsto \sigma^2$ is Lipschitz

$$\frac{X^\epsilon(t) - x^o(t)}{\sqrt{\epsilon}} \xrightarrow[\epsilon \rightarrow 0]{\text{weakly}} Y_o(t)$$

on each interval $[0, T]$, where $Y_o(t)$ is the solution of the stochastic differential equation

$$dY_o(t) = b'(x^o(t))Y_o(t) dt + \sigma(x^o(t)) dW(t).$$

(W is the Brownian motion on \mathbb{R} .) (For a proof see [15], which proves a more general result.)

Thus, for ϵ sufficiently small, we can consider the process $X^\epsilon(t)$ as an approximate solution of the differential equation

$$dX^\epsilon(t) = b'(x^o(t))(X^\epsilon(t) - x^o(t)) dt + b(x^o(t)) dt + \sqrt{\epsilon} \sigma(x^o(t)) dW(t). \quad (1)$$

Note that many people approximate the process X^ϵ by diffusion with the same drift and variance (2)

$$dX^\epsilon(t) = b(X^\epsilon(t)) dt + \sqrt{\epsilon} \sigma(X^\epsilon(t)) dW(t). \quad (2)$$

This is not exactly (1), but the law of large numbers allows us to replace $b'(x^o(t))(X^\epsilon(t) - x^o(t))$ by $b(X^\epsilon(t))$, and the difference between the solutions of (1) and (2) converges weakly to 0 as $\epsilon \rightarrow 0$. Further, the calculations are much easier with (2) and we use (as many people do) that kind of approximation when we want computable formulas: for the accuracy we need, this is quite sufficient.

To approximate $X^\epsilon(t)$ by the solution of (1) or (2) is what we call a standard *diffusion approximation*, in this context: if we compute the mean exit time from n_0 to n_c , $T_\epsilon(n_0, n_c)$, with these approximations we find a wrong value (as we show with the numerical results).

So our idea is to replace our process by a diffusion process only when we evaluate "sufficiently" large probabilities, that is to say, when the event we consider occurs in a time interval of order $[0, 1/\sqrt{\epsilon}]$. Actually, the diffusion approximation is valid only for these times [22].

C. Calculations

First we put $n_0 = 0$ for simplicity and assert an elementary lemma, useful in all the following computations.

Lemma: In the neighborhood of 0, if $b'(0) \neq 0$, we have

$$\lambda(x) = -\frac{2b(x)}{\sigma^2(x)} + O(x^3)$$

$$\tilde{b}(x) = -b(x) + O(x^2)$$

$$\tilde{\sigma}^2(x) = \sigma^2(0) + O(x).$$

(For a proof, see [7].)

We will follow the way mentioned in the introduction of this section. In this one-dimensional case we distinguish four steps to calculate $T_\epsilon(0, n_c)$, the mean exit time from 0 to n_c .

1) An evaluation of the direct exit probability, starting from x_o close to 0: x_o plays the role of ∂U .

2) Boundary conditions at x_o for the OU approximation near 0. For this purpose we introduce y which plays the role of $\partial U'$.

3) Classical calculation of the mean exit time with the OU approximation, including the boundary condition. But, at first, we do not find $T_\epsilon(0, n_c)$, but $T_\epsilon(0, n_b)$ for some $n_b < n_c$ because the speed $b(n_c)$ is null.

4) Repeating, these operations near n_c we find $T_\epsilon(0, n_c)$.

a) *The First Step:* We choose x_o and y close to $n_0 = 0$, depending upon ϵ , such that

$$x_o = a(\epsilon)\sqrt{\epsilon} \text{ and } a(\epsilon) \rightarrow +\infty \text{ when } \epsilon \rightarrow 0$$

$$\frac{x_o^2}{\sqrt{\epsilon}} \rightarrow 0 \text{ when } \epsilon \rightarrow 0$$

$$x_o - y = O(\epsilon).$$

(It is sufficient to take $a(\epsilon) = \epsilon^{-(1/4)+\delta}$ with $0 < \delta < \frac{1}{4}$, and thus $x_o = \epsilon^{(1/4)+\delta}$.)

In Fig. 7 let n_b be fixed such that $n_0 < n_b < n_c$. Although x_o and y depend upon ϵ , we do not indicate it. These conditions mean that x_o is far from 0, with respect to the jump law at 0, but not too much.

Let A be the set of the trajectories which, starting from n_0 , reach n_b before y ; R^ϵ and \tilde{R}^ϵ are, respectively, the probabilities that $X^\epsilon(t)$ and $\tilde{X}^\epsilon(t)$ belong to A .

To estimate the probability R^ϵ , we begin by calculating \tilde{R}^ϵ , using the diffusion approximation for \tilde{X}^ϵ . Then we evaluate $dP/d\tilde{P}(\tilde{X}^\epsilon)$; here we show that $dP/d\tilde{P}(\tilde{X}^\epsilon)$ is almost constant on the event A . More precisely, it is constituted of one term which does not depend upon the form of the trajectories of \tilde{X}^ϵ : this is

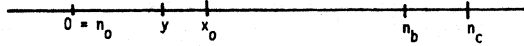


Fig. 7.

the large deviation term. The other term only depends upon the form of the trajectories of A , but is convergent when $\epsilon \rightarrow 0$ by the law of large numbers.

We have approximately the following (this is the same as [2]):

$$d\tilde{X}^\epsilon(t) = \tilde{b}(\tilde{X}^\epsilon(t)) dt + \sqrt{\epsilon} \tilde{\sigma}(\tilde{X}^\epsilon(t)) dW(t). \quad (3)$$

Identifying $\tilde{X}^\epsilon(t)$ as the solution of (3) we have, putting

$$\psi(z) = \int_y^z \frac{2\tilde{b}(u)}{\tilde{\sigma}^2(u)} du$$

the evaluation

$$\tilde{R}^\epsilon = \frac{\int_y^{x_0} \exp\left(-\frac{1}{\epsilon} \psi(z)\right) dz}{\int_y^{n_0} \exp\left(-\frac{1}{\epsilon} \psi(z)\right) dz}$$

(see, for example, [14]).

We find

$$\tilde{R}^\epsilon \sim \frac{2\tilde{b}(x_0)(x_0 - y)}{\epsilon \tilde{\sigma}^2(x_0)}.$$

From the lemma, we deduce, putting $b_0 = -b'(0) = \tilde{b}'(0)$, and $\sigma_0^2 = \sigma^2(0)$,

$$\tilde{R}^\epsilon \sim \frac{2b_0 x_0 (x_0 - y)}{\epsilon \sigma_0^2}. \quad (4)$$

From this equivalent of \tilde{R}^ϵ (probability that the diffusion which approximates $\tilde{X}^\epsilon(t)$ reaches n_b before y , starting from x_0) we try to calculate an equivalent of R^ϵ .

Let $I(z) = \int_0^z \lambda(u) du$, applying the Taylor formula to $I(z)$, we get

$$\begin{aligned} I(\tilde{X}^\epsilon(t)) - I(x_0) &= \int_0^{\tilde{X}^\epsilon(t)} \lambda(u) d\tilde{X}^\epsilon(u) \\ &+ \frac{\epsilon}{2} \int_0^{\tilde{X}^\epsilon(t)} \lambda'(\tilde{X}^\epsilon(u)) \tilde{m}^2(\tilde{X}^\epsilon(u)) du + o(\epsilon^2) \end{aligned} \quad (5)$$

where $\tilde{m}^2(x) = \int y^2 d\tilde{\mu}(x, y)$.

We define T_{n_b} as the reaching time of n_b by the deterministic motion \tilde{x}° , the solution of the ODE:

$$\frac{d\tilde{x}^\circ}{dt} = \tilde{b}(\tilde{x}^\circ).$$

By definition of the probability change, given at each point by $\lambda(x)$, we have along a trajectory ω which reaches n_b before y

$$\epsilon \log \frac{d\tilde{P}^\epsilon}{dP^\epsilon}(\omega) = \int_0^{T_{n_b}} \lambda(\tilde{X}^\epsilon(u)) d\tilde{X}^\epsilon(u) \quad (6)$$

where τ_{n_b} is the reaching time of n_b by ω .

When $\epsilon \rightarrow 0$, τ_{n_b} , $\lambda(\tilde{X}^\epsilon(u))$, and $\tilde{m}^2(\tilde{X}^\epsilon(u))$ converge, respectively, to T_{n_b} , $\lambda(\tilde{x}^\circ(u))$, and $\tilde{m}^2(\tilde{x}^\circ(u))$ in virtue of the law of large numbers.

Hence, using (5) and (6) we have

$$\begin{aligned} \frac{dP^\epsilon}{d\tilde{P}^\epsilon}(\omega) &\underset{\epsilon \rightarrow 0}{\sim} \exp\left(\frac{I(x_0) - I(n_b)}{\epsilon}\right) \\ &\cdot \exp\left(\frac{1}{2} \int_0^{T_{n_b}} \lambda'(\tilde{x}^\circ(u)) \tilde{m}^2(\tilde{x}^\circ(u)) du\right). \end{aligned} \quad (7)$$

Using the variable change $x = \tilde{x}^\circ(u)$, replacing $\tilde{m}^2(x)$ and $\tilde{b}(x)$ by their expressions as functions of $\hat{\mu}_x$, and observing that

$$\lambda'(x) = -(\hat{\mu}_x / \hat{\mu}_s)(x, \lambda(x))$$

we get that

$$\int_0^{T_{n_b}} \lambda'(\tilde{x}^\circ(u)) \tilde{m}^2(\tilde{x}^\circ(u)) du = \int_{x_0}^{n_b} \left[\hat{\mu}_x \cdot \frac{\hat{\mu}_{ss}}{\hat{\mu}_s^2} \right] (x, \lambda(x)) dx \quad (8)$$

but

$$\frac{b_0 x_0}{\tilde{b}(n_b)} \underset{\epsilon \rightarrow 0}{\sim} \exp\left(-\int_{x_0}^{n_b} \frac{\tilde{b}'(x)}{\tilde{b}(x)} dx\right) \quad (9)$$

and

$$\begin{aligned} \frac{\tilde{b}'(x)}{\tilde{b}(x)} &= \frac{1 \times (\hat{\mu}_{sx}(x, \lambda(x)) + \hat{\mu}_{ss}(x, \lambda(x)) \lambda'(x))}{\hat{\mu}_s(x, \lambda(x))} \\ &= \left(\frac{\hat{\mu}_{sx}}{\hat{\mu}_s} - \frac{\hat{\mu}_{ss} \times \hat{\mu}_x}{\hat{\mu}_s^2} \right) (x, \lambda(x)). \end{aligned} \quad (10)$$

Finally,

$$R^\epsilon = P^\epsilon(A) = \int_A \frac{dP^\epsilon}{d\tilde{P}^\epsilon}(\omega) d\tilde{P}^\epsilon(\omega).$$

The restriction to A of $dP^\epsilon/d\tilde{P}^\epsilon$ does not depend (asymptotically) upon $\tilde{X}^\epsilon(t)$, but only on \tilde{x}° , thus using (4), we obtain

$$\begin{aligned} R^\epsilon &\underset{\epsilon \rightarrow 0}{\sim} \frac{2b_0 x_0 (x_0 - y)}{\epsilon \sigma_0^2} \exp\left(\frac{I(x_0) - I(n_b)}{\epsilon}\right) \\ &\cdot \exp\left(\frac{1}{2} \int_0^{n_b} \left[\hat{\mu}_x \cdot \frac{\hat{\mu}_{ss}}{\hat{\mu}_s^2} \right] dx\right) \end{aligned}$$

and by (9) and (10),

$$R^\epsilon \underset{\epsilon \rightarrow 0}{\sim} (x_0 - y) q_\epsilon \quad (11)$$

where

$$q_\epsilon \underset{\epsilon \rightarrow 0}{\sim} \frac{2\tilde{b}(n_b) \exp\left(\frac{I(x_0)}{\epsilon}\right)}{\epsilon \sigma_0^2 \exp\left(\frac{I(n_b)}{\epsilon}\right)} \exp K(n_b)$$

with

$$K(n_b) = \int_0^{n_b} \left[-\frac{\hat{\mu}_{sx}}{\hat{\mu}_s} + \frac{1}{2} \hat{\mu}_x \cdot \frac{\hat{\mu}_{ss}}{\hat{\mu}_s^2} \right] (x, \lambda(x)) dx. \quad (12)$$

The derivatives are calculated at the point $(x, \lambda(x))$.

b) The Second Step: Let $T_i(a, b)$ be the expectation of the time necessary to reach b , starting from a .

Putting $F_i(x_0, n_b) =$ mean time to reach n_b , starting from x_0 ,

given that it does not pass through y , $H_\epsilon(x_o, y)$ = mean time to reach y , starting from x_o , given that it does not pass through n_b , we have

$$T_\epsilon(x_o, n_b) = (1 - R^\epsilon)(H_\epsilon(x_o, y) + T_\epsilon(y, n_b)) + R^\epsilon F_\epsilon(x_o, y). \quad (13)$$

Observing that, when $\epsilon \rightarrow 0$, $H_\epsilon(x_o, y)$ is negligible relative to $T_\epsilon(y, x_o)$ and that $F_\epsilon(x_o, n_b)$ is negligible relative to $T_\epsilon(y, n_b)$, we get

$$T'_\epsilon(x_o, n_b) \underset{\epsilon \rightarrow 0}{\sim} -q_\epsilon T_\epsilon(x_o, n_b). \quad (14)$$

($T'_\epsilon(x_o, n_b)$ is the derivative of $T_\epsilon(x, n_b)$ with respect to x .)

This is the reflection-extinction condition on the "boundary" of U .

c) *The Third Step:* The approximation of $X^\epsilon(t)$ by a diffusion is simpler in a neighborhood of $n_o = 0$ (we take $[-x_o, x_o]$); if we start from 0, $x^\epsilon(t) = 0$ for each t , and the equation (1) allows us to write

$$dX^\epsilon(t) = -b_o X^\epsilon(t) dt + \sqrt{\epsilon} \sigma_o dW(t). \quad (15)$$

Then the time $T_\epsilon(x, n_b)$ ($x \in [-x_o, x_o]$) verifies the differential equation

$$\frac{\epsilon}{2} \sigma_o^2 u''_\epsilon - b_o \times u'_\epsilon = -1$$

with the boundary condition

$$u'_\epsilon(x_o) \sim -q_\epsilon u_\epsilon(x_o)$$

and $u'_\epsilon(-x_o) = 0$ (reflection in $-x_o$). (The choice of $(-x_o)$ is not very important because $-x_o/\sqrt{\epsilon} \rightarrow -\infty$ as $\epsilon \rightarrow 0$.)

Integrating one time we have

$$u'_\epsilon(x) = \frac{-2}{\epsilon \sigma_o^2} \exp \frac{b_o x^2}{\epsilon \sigma_o^2} \times \int_{-x_o}^x \exp \frac{-b_o v^2}{\epsilon \sigma_o^2} dv$$

$$\frac{u'_\epsilon(x_o)}{u_\epsilon(x_o)} \sim -q_\epsilon. \quad (16)$$

We easily see that, when $\epsilon \rightarrow 0$, $u_\epsilon(x_o)$ and $u_\epsilon(0)$ are equivalents. So we have

$$T_\epsilon(0, n_b) = u_\epsilon(0) \underset{\epsilon \rightarrow 0}{\sim} \frac{u'_\epsilon(x_o)}{q_\epsilon}$$

$$\underset{\epsilon \rightarrow 0}{\sim} \frac{2\sqrt{\pi}}{\sqrt{\epsilon} \sigma_o \sqrt{b_o}} \exp \frac{b_o x_o^2}{\epsilon \sigma_o^2}$$

$$\frac{\epsilon \sigma_o^2 \exp \frac{I(n_b)}{\epsilon} \exp -K(n_b)}{2\tilde{b}(n_b) \exp \frac{I(x_o)}{\epsilon}} \quad (17)$$

by (12) and (15).

Noting that

$$\frac{I(x_o)}{\epsilon} = \frac{b_o x_o^2}{\epsilon \sigma_o^2} + \frac{1}{\epsilon} 0(x_o^2)$$

and choosing x_o such that $x_o^3/\epsilon \rightarrow 0$ when $\epsilon \rightarrow 0$, ($x_o = \epsilon^{(1/4)+\delta}$ and $\frac{1}{2} < \delta < \frac{1}{4}$), it holds that

$$\exp \frac{b_o x_o^2}{\epsilon \sigma_o^2} \exp -\frac{I(x_o)}{\epsilon} \rightarrow 1$$

when $\epsilon \rightarrow 0$.

Finally, we get an equivalent of $T_\epsilon(0, n_b)$ (the expression does not depend upon x_o)¹

$$T_\epsilon(0, n_b) \underset{\epsilon \rightarrow 0}{\sim} \frac{\sqrt{\pi} \sqrt{\epsilon} \sigma_o}{\sqrt{b_o}} \frac{\exp \frac{I(n_b)}{\epsilon}}{\tilde{b}(n_b)} \exp -K(n_b) \quad (18)$$

where $K(n_b)$ is defined by (12).

d) *Crossing n_c :* As in the neighborhood of n_o , we can repeat the same calculation, without probability change in the neighborhood of n_c . Using (18), and supposing $b_c = b'(n_c) > 0$ and $\sigma_c^2 = \sigma^2(n_c)$, we get

$$T_\epsilon(n_b, n_c) \underset{\epsilon \rightarrow 0}{\sim} \frac{\pi}{\sqrt{b_o b_c}} \frac{\sigma_o}{\sigma_c} \exp \frac{I(n_c)}{\epsilon} \exp -K(n_c)$$

where

$$-K(n_c) = \int_0^{n_c} \left[\frac{\dot{\mu}_{xx}}{\dot{\mu}_x} - \frac{1}{2} \dot{\mu}_x \cdot \frac{\dot{\mu}_{xx}}{\dot{\mu}_x^2} \right] (x, \lambda(x)) dx \quad (19)$$

and

$$I(n_c) = \int_0^{n_c} \lambda(x) dx.$$

The time $T_\epsilon(n_o, n_c)$ is the sum of $T_\epsilon(n_o, n_b)$ and of $T_\epsilon(n_b, n_c)$, since the process $X^\epsilon(t)$ is Markovian, but

$$T_\epsilon(n_o, n_b) = 0 \left(\frac{\sqrt{\epsilon}}{b_c(n_b - n_c)} \exp \frac{I(n_b)}{\epsilon} \right).$$

So, the final result is

$$T_\epsilon(n_o, n_c) \underset{\epsilon \rightarrow 0}{\sim} \frac{\pi}{\sqrt{b_o b_c}} \frac{\sigma_o}{\sigma_c} \exp \frac{I(n_c)}{\epsilon} \exp -K(n_c) \quad (20)$$

where $I(n_c)$ and $K(n_c)$ are as in (12).

This formula gives an accuracy of 10 percent when we apply it with $\epsilon = 1$ to some examples.

D. Some Numerical Results

Here we test the validity of formulas (18) for $\tau(0, n_b)$ and (20) for $\tau(0, n_c)$ on some examples.

This is a very simple example where we can compare the result to the analytical one. It is a Markov chain $N(t)$, with transition probabilities.

$$p_{i, i-1} = r(i), \quad p_{i, i+1} = 1 - r(i) \quad \text{if } -10 < i < n_c$$

$$p_{-10, -9} = p_{n_c, n_c} = 1.$$

We took $r(i) = 0.5 + \rho i(n_c - i)$ for different values of ρ and n_c . Table I shows the results for $n_c = 50$. The first number is the exact result, the second is provided by the standard diffusion approximation, and the third is derived from formula (20), by our diffusion approximation.

¹D. Ludwig in [18] has obtained, for diffusions only, a similar expression for this time, using quite a different method.

TABLE I

$n_b \backslash p$	10^{-4}	$2 \cdot 10^{-4}$	$3 \cdot 10^{-4}$	$4 \cdot 10^{-4}$	$5 \cdot 10^{-4}$
20	$2.8 \cdot 10^3$	$2 \cdot 10^4$	$2.3 \cdot 10^5$	$3.8 \cdot 10^6$	$0.9 \cdot 10^8$
	$2.8 \cdot 10^3$	$2.3 \cdot 10^4$	$3.7 \cdot 10^5$	$13 \cdot 10^6$	$17 \cdot 10^8$
	$2.8 \cdot 10^3$	$2 \cdot 10^4$	$2.4 \cdot 10^5$	$4.3 \cdot 10^6$	$1.1 \cdot 10^8$
25	$9.1 \cdot 10^3$	$2.3 \cdot 10^5$	$1 \cdot 10^7$	$8 \cdot 10^8$	$1.1 \cdot 10^{11}$
	$9.4 \cdot 10^3$	$2.9 \cdot 10^5$	$2.4 \cdot 10^7$	$8.5 \cdot 10^9$	$337 \cdot 10^{11}$
	$9.3 \cdot 10^3$	$2.4 \cdot 10^5$	$1.1 \cdot 10^7$	$9.1 \cdot 10^9$	$1.4 \cdot 10^{11}$
30	$3.1 \cdot 10^4$	$2.9 \cdot 10^6$	$5 \cdot 10^8$	$1.8 \cdot 10^{11}$	$1.6 \cdot 10^{14}$
	$3.2 \cdot 10^4$	$4.0 \cdot 10^6$	$1.8 \cdot 10^9$	$60 \cdot 10^{11}$	$7 \cdot 10^{17}$
	$3.3 \cdot 10^4$	$3.0 \cdot 10^6$	$5.7 \cdot 10^8$	$2.1 \cdot 10^{11}$	$2 \cdot 10^{14}$
$n_c = 50$	$1.4 \cdot 10^6$	$3.7 \cdot 10^8$	$1.9 \cdot 10^{13}$	$2.2 \cdot 10^{17}$	$8.9 \cdot 10^{21}$
	$1.5 \cdot 10^6$	$60 \cdot 10^8$	$12 \cdot 10^{13}$	$323 \cdot 10^{17}$	$11 \cdot 10^{26}$
	$1.3 \cdot 10^6$	$3.5 \cdot 10^8$	$1.8 \cdot 10^{13}$	$2.1 \cdot 10^{17}$	$8.3 \cdot 10^{21}$

IV. EXAMPLES

A. Aloha System

The ALOHA system is a radio channel shared by many users who want to transmit messages called packets. The channel time is divided into slots equal to the packet duration (the packets are supposed to be of a constant size about 10^{-4} s). The beginning of the packet transmission and of the slots are coincident. Whenever two packets are simultaneously emitted, there is a conflict, they are lost, and the corresponding users are blocked. They must emit again at random to avoid another conflict.

Let us present the mathematical model.

1) *Not Controlled System*: When a conflict occurs, each involved user flips a coin (with a fixed common bias p) to see if it is going to retransmit or not.

We define the process $N(t)$ as the number of blocked users, at time t . Assuming the global source is a Poisson process with parameter α , $N(t)$ is a transient homogeneous Markov chain over \mathbf{N} for every value of α and p (see [1] and [17]).

The transition probabilities are (for $i \geq 0, j \geq 0$),

$$p_{ij} = P(N(t+1) = j | N(t) = i) = \begin{cases} 0 & \text{for } j \leq i-2 \\ ip(1-p)^{i-1}e^{-\alpha} & \text{for } j = i-1 \\ (1-p)^i \alpha e^{-\alpha} \\ + [1-ip(1-p)^{i-1}]e^{-\alpha} & \text{for } j = i \\ \alpha e^{-\alpha} [1-(1-p)^i] & \text{for } j = i+1 \\ \frac{\alpha^{j-1}}{(j-i)!} e^{-\alpha} & \text{for } j \geq i+2. \end{cases}$$

Let μ_i be the jump distribution at i , $b(i) = E(\mu_i)$ and $\sigma^2(i) = \text{var}(\mu_i)$. If $\alpha < 1/e$, we can observe that $N(t)$ remains for a very long time period near a value n_o , and as soon as it reaches (or overtakes) a value n_c (critical point corresponding to the channel saturation), it increases quickly to infinity.

The points n_o and n_c are mean equilibrium points, one stable (n_o), the other unstable (n_c), defined by the fact that the sign of $E(N(t+1) - N(t) | N(t) = i)$ changes between i and $i+1$.

So in Fig. 8 we have $b(i) > 0$ for $0 \leq i < n_o$ or $n_c < i$, and $b(i) < 0$ for $n_o < i < n_c$.

A good measure of the system's "stability" is in the mean time of reaching n_c , starting from n_o .

For a good choice of the parameters α and p , this time can be very long, and it is during this time that the channel works. We call it *exit time*, denoted by τ .

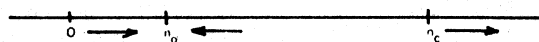


Fig. 8.

2) *Controlled System*: For stabilizing the system, we choose a retransmission probability $p(i)$, common to every user, but depending upon the number i of blocked users, for example $p(i) = (1-\alpha)/(i-\alpha)$. Then $N(t)$ is a positive recurrent chain iff $0 < \alpha < 1/e$ (see [10]).

The probability that a given user remains blocked during a time greater than a fixed bound is a good evaluation test of different retransmission politics (for the same source rate).

As the event "a given user is still in the system" has measure 0, to keep up with a given user it is necessary to consider a conditional distribution, and to define a new process $M(t)$, which is the number of blocked users, observed by a marked one, that is conditionally on the fact that he is still waiting. It is also a Markov chain, with a death probability different from 0: the death is the event "the marked user can emit successfully."

When the number of blocked users is i , the death probability is

$$e^{-\alpha} p(i) (1-p(i))^{i-1} = 1 - \int d\nu_i$$

if ν_i is the jump distribution at i for the new process $M(t)$.

We are interested in studying the distribution of the survival time of the marked user, i.e., the time before the death of the process $M(t)$.

B. Estimating Exit Time for the Not Controlled Aloha System

The jump distribution μ_i does not vary too much as a function of i , and the jumps are small, so it is interesting to consider the process $N(t)$ as an element of a family of processes $X^\epsilon(t)$.

We write $N(t+1) = N(t) + V(N(t))$ where $V(N(t))$ has the distribution $\mu_{N(t)}$. We define the family $X^\epsilon(t)$ by

$$X^\epsilon(t+1) = X^\epsilon(t) + \epsilon V(X(t))$$

and we have $N(t) = X^\epsilon(t)$ for $\epsilon = 1$.

So we can apply the results of Sections II and III.

1) *Application of Theorem 2 of Section II*: Let β be the probability $P(A)$, where A is the event "starting from n_o , the process comes directly to n_c , without returning to n_o ."

Every trajectory of $N(t)$, starting from n_o , can be cut into sections. A great number of sections come back to n_o , we call them sections of the first kind. The last section reaches n_c and goes out. It is a section of second kind.

So the number of sections in a trajectory before the exit is a geometric random variable, with parameter β and mean value $1/\beta$.

As the durations of the sections of the first kind are random variables iid (by strong Markov property), calling τ_o their common expectation, and neglecting the duration of the last part of the trajectory, we have an evaluation of the mean exit time τ as

$$\tau \approx \tau_o / \beta.$$

For estimating τ_o , we merely simulate the chain $N(t)$, with transition probabilities p_{ij} , and take the average as the estimator $\hat{\tau}_o$.

Applying Theorem 2, of Section II we use the quick simulation method to estimate the very small probability β , considering in the sections space the event A = "the section ω belongs to the second kind."

So we simulate n sections $\omega_1, \omega_2, \dots, \omega_n$ using the transformed

transition probabilities

$$\tilde{p}_{ij} = e^{\lambda_i(j-i)}$$

with $\tilde{p}_i(\lambda_i) = 1$ and $\lambda_i > 0$ for $n_o < i < n_c$ and $\tilde{p}_{ij} = p_{ij}$ for $i \leq n_o$ or $i \geq n_c$.

And we take the following unbiased and convergent estimator of β :

$$\hat{\beta} = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_A(\omega_i) \frac{dP}{d\tilde{P}}(\omega_i).$$

The Radon-Nikodym derivative is calculated along a section by

$$\frac{dP}{d\tilde{P}}(\omega_i) = \exp\left(-\sum_t \lambda_{\omega_i(t)} (\omega_i(t+1) - \omega_i(t))\right).$$

On the other hand, we can use the diffusion approximation formula presented in Section III to calculate the mean exit time starting from n_o .

So we present below numerical results obtained by the two methods.

Here we take as examples $\alpha = 0.27$ and 0.30 for different values of p .

The simulation results are calculated with an accuracy of 10 percent and are compared to both results given by the formula (20) of Section III, and those obtained by a standard diffusion approximation. The expected exit time is the solution of the boundary value problem

$$\frac{1}{2} \sigma^2(x) u'' + b(x) u' = -1, \quad u(n_c) = 0, \quad u'(0) = 0.$$

In Table II it clearly appears that the standard diffusion approximation greatly overestimates the exit time, and that the formula (20) is much closer to the real values.

Note About the Acceleration of the Simulation: In Table III are some examples of the acceleration of the simulation, achieved by our method.

C. Controlled Aloha System

Here we present a second example of *best asymptotically change of probability for fast simulation*.

In many cases, one can apply the large deviation results to obtain the functional whose minimization gives the best change of probability, and the problem is then to determine the solution minimizing this functional.

1) *The Large Deviations Result:* Let $X^\epsilon(t)$ be a Markov chain with jump distribution at $x, \epsilon \cdot \nu_x$, where $\int d\nu_x(y) < 1$. That is to say that our chain has positive probability to die at each point x .

Let \mathcal{T} be the death time of this process. For T and x_o , we denote by A_{T, x_o} the event "starting from x_o the process is still alive at time T ".

$$A_{T, x_o} = \{\omega \in \Omega / \omega(o) = x_o, \mathcal{T}(\omega) \geq T\} \quad (1)$$

We have the large deviation results (with always the same hypotheses on the measure ν_x). Set $\rho_{T, x_o}^\epsilon = P^\epsilon(A_{T, x_o})$

$$\lim_{\epsilon \rightarrow 0} -\epsilon \log \rho_{T, x_o}^\epsilon = \inf_{\varphi} \{I(x_o, T, \varphi), \varphi(o) = x_o, \varphi \in \mathcal{C}_T\} \quad (2)$$

where $I(x_o, T, \varphi)$ is the action integral along φ defined by

$$I(x_o, T, \varphi) = \int_0^T h_{\varphi(t)}(\dot{\varphi}(t)) dt$$

(h is the Cramer transformed of ν_x) and changing of notations in the continuation

TABLE II

	p	n_o	n_c	$\hat{\tau}_o$	β	$\hat{\tau}$	Diffusion	Standard diff.
$\alpha=0.30$	0.049	4	30	8.75	$1.8 \cdot 10^{-4}$	$4.8 \cdot 10^4$	$4.6 \cdot 10^4$	$5.3 \cdot 10^4$
	0.039	5	37	9.0	$6.0 \cdot 10^{-5}$	$1.5 \cdot 10^5$	$1.6 \cdot 10^5$	$1.9 \cdot 10^5$
	0.032	6	45	10.5	$1.9 \cdot 10^{-5}$	$5.5 \cdot 10^5$	$5.8 \cdot 10^5$	$7.5 \cdot 10^5$
	0.028	7	52	11.0	$6.0 \cdot 10^{-6}$	$1.8 \cdot 10^6$	$1.8 \cdot 10^6$	$2.6 \cdot 10^6$
	0.024	8	60	11.3	$1.9 \cdot 10^{-6}$	$5.9 \cdot 10^6$	$6.2 \cdot 10^6$	$9.2 \cdot 10^6$
	0.022	9	67	11.8	$6.7 \cdot 10^{-7}$	$1.8 \cdot 10^7$	$2 \cdot 10^7$	$3 \cdot 10^7$
	0.019	10	75	13.1	$2.2 \cdot 10^{-7}$	$5.9 \cdot 10^7$	$6.4 \cdot 10^7$	$10 \cdot 10^7$
$\alpha=0.27$	0.064	2	26	5.6	$3.3 \cdot 10^{-5}$	$1.7 \cdot 10^5$	$1.6 \cdot 10^5$	$2.2 \cdot 10^5$
	0.049	3	35	7.1	$4.3 \cdot 10^{-6}$	$1.6 \cdot 10^6$	$1.7 \cdot 10^6$	$2.6 \cdot 10^6$
	0.039	4	44	8.2	$5.5 \cdot 10^{-7}$	$1.6 \cdot 10^7$	$1.6 \cdot 10^7$	$2.8 \cdot 10^7$
	0.032	4	52	8.6	$6.8 \cdot 10^{-8}$	$1.2 \cdot 10^8$	$1.1 \cdot 10^8$	$2.5 \cdot 10^8$
	0.028	5	61	9.4	$9.5 \cdot 10^{-9}$	$1.0 \cdot 10^9$	$1.0 \cdot 10^9$	$2.5 \cdot 10^9$
	0.024	6	70	10.7	$1.3 \cdot 10^{-9}$	$8.2 \cdot 10^9$	$8.9 \cdot 10^9$	$23 \cdot 10^9$

TABLE III

	$\hat{\tau}$	Trajectory duration in simulation	Acceleration
$\alpha = 0.30$			
$p = 0.049$	$4.8 \cdot 10^4$	682	18 times
$p = 0.039$	$1.5 \cdot 10^5$	787	48
$p = 0.028$	$1.8 \cdot 10^6$	115	403
$\alpha = 0.27$			
$p = 0.064$	$1.7 \cdot 10^5$	372	114
$p = 0.049$	$1.6 \cdot 10^6$	539	744

$$= \int_0^T h(\varphi(t), \dot{\varphi}(t)) dt.$$

First we determine the path φ minimizing $I(x_o, T, \varphi)$ and then we apply to a fast simulation method.

Let us introduce

$$\psi(x, s) = \inf \left\{ \int_s^T h(\varphi(t), \dot{\varphi}(t)) dt / \varphi \in \mathcal{C}_T, \varphi(s) = x \right\} \quad (3)$$

and let φ_{opt} be the path which gives the minimum. Assuming that ψ is C^2 , ψ verifies the dynamical programming equation (see [9]), i.e., for $\varphi \in \mathcal{C}, x = \varphi(s), dx/ds = \dot{\varphi}(s) = v$,

$$\begin{aligned} \frac{\partial \psi}{\partial s}(x, s) &= - \inf_v \left(h(x, v) + \frac{\partial \psi}{\partial x}(x, s) v \right) \\ &= \sup_v \left(-h(x, v) - \frac{\partial \psi}{\partial x}(x, s) v \right). \end{aligned} \quad (4)$$

We recognize the transform of $h(x, \cdot)$ by the convex duality at the point $-\partial\psi/\partial x(x, s)$ (see Appendix I-A-3).

We put

$$\lambda(x, s) = - \frac{\partial \psi}{\partial x}(x, s)$$

$v(x, s)$ is the point providing the sup in (4), i.e., the speed at time s , point x , along the optimal path φ_{opt} .

Then $\partial\psi/\partial s(x, s) = I(x, \lambda(x, s))$ by reciprocity of the convex duality, and $v(x, s) = \partial I/\partial s(x, \lambda(x, s))$ (reciprocity of the h 's and I 's derivatives).

Thus, we want to determine φ_{opt} and λ , along φ_{opt} . We have

$$\frac{\partial \lambda}{\partial s}(x, s) = -\frac{\partial}{\partial x} l(x, \lambda(x, s)) - \left(\frac{\partial l}{\partial s}(x, \lambda(x, s)) \frac{\partial \lambda}{\partial x}(x, s) \right)$$

and

$$\begin{aligned} d\lambda &= v(x, s) \frac{\partial \lambda}{\partial x} \left(-\frac{\partial l}{\partial x}(x, \lambda(x, s)) \right. \\ &\quad \left. - v(x, s) \frac{\partial \lambda}{\partial x}(x, s) \right) ds \\ &= -\frac{\partial l}{\partial x}(x, \lambda(x, s)) ds. \end{aligned}$$

Along φ_{opt} , it holds that ($x = \varphi_{\text{opt}}(s)$)

$$\begin{aligned} d\lambda &= -\frac{\partial l}{\partial x}(x, \lambda(x, s)) ds \\ dx &= \frac{\partial l}{\partial s}(x, \lambda(x, s)) ds. \end{aligned} \quad (5)$$

Furthermore, φ_{opt} is determined by the boundary conditions $\varphi_{\text{opt}}(0) = x_0$, $\partial \psi / \partial x(x, s) = 0$ if $s > T$, in such a way that

$$\psi(x, 0) = I(x_0, T, \varphi_{\text{opt}})$$

is minimum.

Hence the pair ($x = \varphi_{\text{opt}}(s)$, $\lambda = \lambda(x, s)$) is the solution of the system

$$\begin{aligned} l(x, \lambda) &= C = \text{Cste} \quad (\text{level curves of } l) \\ x(0) &= x_0 \\ \lambda(x(T), T) &= 0. \end{aligned} \quad (6)$$

Along the trajectory φ_{opt} , it holds that

$$\frac{\partial l}{\partial s}(x, \lambda(x, s)) = \frac{dx}{ds}.$$

So the coefficient $\lambda(x, s)$ is the rate of probability change which centers the process around the path φ_{opt} . We use this remark in the next section.

2) *Simulation of the Controlled ALOHA System:* For the Markov chain $M(t)$ defined in Section IV-A-2, we denote \mathcal{T} the survival time, and we want to estimate for each value of T ,

$$\rho_T = P(\mathcal{T} \geq T) = \sum_{x_0=0}^{\infty} \rho_{T, x_0} P_{x_0}$$

where P_{x_0} is the probability to find x_0 blocked users, at equilibrium.

As in the uncontrolled ALOHA case we apply the previous results of large deviations when $\epsilon = 1$. We hope the conditions for application are still good. However, since we only simulate, the approximation $\epsilon = 1 - \epsilon = 0$ is not dangerous: if this approximation is not valid, the simulation is not greatly accelerated, and there is no other trouble.

For each x_0 , we simulate n trajectories $\omega_1, \dots, \omega_n$ of a transformed process with distribution \bar{P} to determine and we take

$$\hat{\rho}_{T, x_0} = \frac{1}{n} \sum_{k=1}^n \mathbb{1}_{A_{T, x_0}}(\omega_k) \frac{dP}{d\bar{P}}(\omega_k).$$

Choice of the Transformed Process:

a) *Asymptotically Optimal Probability Change:* For each time t and initial point x , an optimal trajectory φ_{opt} , denoted by $\varphi_{t, x}$,

is associated. Putting $\Lambda(x, t) = \lambda(\varphi_{t, x}(0), 0)$, when a time t is spent, we choose the exponential probability change defined by $\Lambda(x, T-t)$.

b) *A Simpler Choice:* Instead of taking care of the real evolution of the trajectory to choose the probability change, we can take, as an approximation, the optimal path starting from x_0 for the time T . This trajectory is determined by an equation

$$l(x, \lambda) = C. \quad (1)$$

For each x , the new probability distribution $\bar{v}(x, \cdot)$ is defined by the exponential change with parameter $\lambda(x)$ solution of (1). So the total weight of $\bar{v}(x, \cdot)$ is $\bar{v}(x, \lambda) = e^C$; therefore we simulate trajectories with constant death probability.

For $T = 40, 60, 80, 100$, we estimate $\rho_{T, 0}$ and compare the relative standard error σ_R obtained after 5000 simulations by both a) and b) methods. (See Table IV.) The two last columns represent the Bernoulli parameter for which we obtain the same accuracy after 5000 simulations and the acceleration due to the probability change.

The results a) and b) are actually almost the same: here the simplified method is as efficient as the optimal method. Below we present some results.

c) *Approximation by Diffusion:* As in Section III we have a corresponding asymptotical formula using approximations by diffusions. We proceed as follows. We note that

$$\bar{P}(A_{T, x_0}) = e^{CT/\epsilon}$$

and we evaluate $dP^\epsilon/d\bar{P}^\epsilon$ on A_{T, x_0} by Ito's formula. We obtain two terms. The first is the large deviation term

$$\exp - \frac{1}{\epsilon} \int_0^T \lambda(\varphi_{\text{opt}}(t), t) dt.$$

The second depends upon the most likely trajectory, but not on ϵ

$$\exp \frac{1}{2} \int_0^T \lambda_x(\varphi_{\text{opt}}(t), t) \bar{m}^2(\varphi_{\text{opt}}(t), t) dt$$

where λ_x is the derivative of λ with respect to the first variable and $\bar{m}^2(\varphi_{\text{opt}}(t))$ is the 2^d moment of the probability measure $\bar{v}(\varphi_{\text{opt}}(t), \cdot) / e^C$.

Fig. 9 shows the form of the curve $l(x, \lambda) = C$. The equation $l(x, \lambda) = C$ has two solutions in general, we put $\lambda_1(x)$ [resp. $\lambda_2(x)$] the largest one (resp. the smallest).

Thus we get

$$\begin{aligned} P^\epsilon(A_{T, x_0}) &= \exp \frac{CT}{\epsilon} \\ &\cdot \exp - \frac{1}{\epsilon} \left[\int_{x_0}^{x_R} \lambda_1(x) dx + \int_{x_R}^{x_F} \lambda_2(x) dx \right] \\ &\cdot \exp \frac{1}{2} \left[\int_{x_0}^{x_R} - \left[\hat{\mu}_x \cdot \frac{\hat{\mu}_{ss}}{\hat{\mu}_s^2} \right] (x, \lambda_1(x)) dx \right] \\ &\cdot \exp \frac{1}{2} \left[\int_{x_R}^{x_F} - \left[\hat{\mu}_x \cdot \frac{\hat{\mu}_{ss}}{\hat{\mu}_s^2} \right] (x, \lambda_2(x)) dx \right]. \end{aligned}$$

(The two last integrals are divergent, but the sum is convergent.)

For instance, we studied numerically the distribution of the survival time of the controlled ALOHA system, and we get

$$\rho_{40} = P(\mathcal{T} \geq 40) = 40 \cdot 10^{-3} \pm 5 \text{ percent}$$

$$\rho_{60} = P(\mathcal{T} \geq 60) = 20 \cdot 10^{-3} \pm 5 \text{ percent}$$

$$\rho_{80} = P(\mathcal{T} \geq 80) = 10 \cdot 10^{-3} \pm 5 \text{ percent.}$$

TABLE IV

T	$P_{T,0}$	σ_R a)	σ_R b)	P	$P/P_{T,0}$
40	$41 \cdot 10^{-4}$	5.4 %	4.6 %	$86 \cdot 10^{-3}$	21
60	$15 \cdot 10^{-4}$	6.5 %	7.0 %	$40 \cdot 10^{-3}$	26
80	$6 \cdot 10^{-4}$	7.9 %	7.9 %	$30 \cdot 10^{-3}$	50
100	$3 \cdot 10^{-4}$	11.3 %	11.1 %	$16 \cdot 10^{-3}$	55

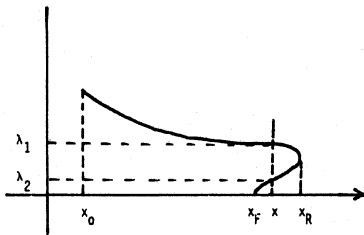


Fig. 9.

APPENDIX I

A. Laplace Transform-Cramer Transform

Let μ be a finite positive measure on \mathbb{R}^n (it is not necessarily a probability) and $b = \int x d\mu(x)$ its mean value.

1) Laplace Transform: $\hat{\mu}$ of the measure μ is the application

$$s \rightarrow \hat{\mu}(s) = \int e^{s \cdot x} d\mu(x)$$

from \mathbb{R}^m to \mathbb{R} .

Let U be the interior of the set $\{s \in \mathbb{R}^m / \hat{\mu}(s) < +\infty\}$. We assume that U is an open set containing 0.

One easily checks that

$$\hat{\mu}(0) = \mu(\mathbb{R}^m) \quad \hat{\mu}'(0) = b$$

and that $\hat{\mu}$ is a convex infinitely derivable function.

2) Log-Laplace Transform: l of the measure μ is the application $s \rightarrow l(s) = \log \hat{\mu}(s)$. It is a convex real-valued infinitely derivable function, and $l'(0) = b/\mu(\mathbb{R}^m)$, i.e., the barycenter of the measure μ .

If μ is a probability, $\hat{\mu}(0) = 1$ and $l(0) = 0$.

3) Cramer Transform: The function l being convex real-valued, we can define a function h , from \mathbb{R}^m to $[0, +\infty]$ by

$$h(u) = \sup_s (\langle s, u \rangle - l(s)).$$

The function h is called Cramer transform of the measure μ . Let V be the interior of the set $\{u / h(u) < +\infty\}$.

Let us give some classical properties of h .

a) The function h is convex. The transformation $l \rightarrow h$ is called convex duality and is involutive

$$l(s) = \sup_u (\langle s, u \rangle - h(u)).$$

b) For each $u \in V$, the point of U where $\sup_s (\langle s, u \rangle - l(s))$ is reached is denoted $s(u)$ and

$$l'(s(u)) = \frac{\hat{\mu}'(s(u))}{\hat{\mu}(s(u))} = u.$$

Since l is convex, the point $s(u)$ is unique and $s(u) = l'^{-1}(u)$.

c) The closed set \bar{V} is the closure of the convex envelope of the support of the measure μ : in fact, if $s \in U$

$$l'(s) = \frac{\hat{\mu}'(s)}{\hat{\mu}(s)} = \frac{\int x e^{s \cdot x} d\mu(x)}{\int e^{s \cdot x} d\mu(x)}$$

is the barycenter of the probability measure ν defined by

$$d\nu(x) = \frac{e^{s \cdot x} d\mu(x)}{\int e^{s \cdot x} d\mu(x)}$$

which has the same support as the measure μ .

d) The derivative h' and l' are reciprocal, i.e.,

$$h'(l'(s)) = s \quad \text{for } s \in U$$

and

$$l'(h'(u)) = u \quad \text{for } u \in V.$$

e) The function h is minimum for $b/\mu(\mathbb{R}^m) = l'(0)$. If μ is a probability, this minimum is 0.

f) We have

$$h_{\mu+v}(u) = h_{\mu}(u-v)$$

$$h_{\mu \circledast n}(u) = nh_{\mu}(u)$$

for $u, v \in \mathbb{R}$.

g) The ratio $h(u)/\|u\| \rightarrow +\infty$ when $\|u\| \rightarrow \infty$ iff $\hat{\mu}(s) < \infty$ for every $s \in \mathbb{R}^m$; particular case $m = 1$.

h) Minimizing $h(u)/u$: Let

$$b = E(\mu) = \int x d\mu(x) = 0.$$

Then we have

$$\lambda = \min_{u > 0} \frac{h(u)}{u} \Leftrightarrow l(\lambda) = 0, \quad \lambda > 0. \quad (A1)$$

Actually, $h(u)/u$ is the slope of the straight line joining 0 to the point of abscissa u , and $\lambda = \min_{u > 0} h(u)/u = h'(n_0)$ is the slope of the tangent to the graph of h , issued from 0. Hence, (A1) holds because of the properties a) and i).

Moreover,

$$l'(\lambda) = U_0 = \frac{\hat{\mu}'(\lambda)}{\hat{\mu}(\lambda)}.$$

B. Kullbach Information and Probability Change

1) Exponential Change: Let μ be a probability measure, on \mathbb{R} and let us define $\hat{\mu}, l, h$ as in Section I.

If $s \in U$, let us define a new probability measure μ^s by

$$d\mu^s(x) = \frac{e^{sx} d\mu(x)}{\hat{\mu}(s)}.$$

We say that μ^s is transformed of μ by the exponential change of parameter s .

We see that

$$E(\mu^s) = \int x d\mu^s(x) = \frac{\hat{\mu}'(s)}{\hat{\mu}(s)} = l'(s).$$

So for every $u \in U$, it is possible to find $s = s(u)$ (with $l'(s(u)) = u$) in such a way that the mean of the probability $\mu^{s(u)}$ becomes u .

2) Kullbach Information: Let μ_1, μ_2 be two probabilities such that μ_1 is absolutely continuous with respect to μ_2 , i.e., μ_1 has a

density $d\mu_1/d\mu_2$ with respect to μ_2 . In this case, we denote $\mu_1 \ll \mu_2$.

The Kullbach information (see [16]) of μ_1 with respect to μ_2 is

$$K(\mu_1, \mu_2) = \int \log \left(\frac{d\mu_1}{d\mu_2}(x) \right) d\mu_1(x).$$

Note that $K(\mu_1, \mu_2) = 0$ iff $\mu_1 = \mu_2$. We have the following result which links the Cramer transform and the Kullbach information (see [6], for instance).

Theorem: If $u \in U$ is a probability measure, h its Cramer transformation, then

$$\inf \left\{ K(\nu, \mu) / \nu \ll \mu, \int x d\nu(x) = u \right\} = K(\mu^{s(u)}, \mu) = h(u).$$

It means that the measure which minimizes the Kullbach information with respect to μ of all the probabilities with fixed mean u , is the probability $\mu^{s(u)}$ obtained from μ by the exponential change of parameter $s(u)$ [defined by $u = I'(s(u))$].

C. Chernoff's Theorems

We give some propositions in \mathbb{R} , but it is not difficult to generalize them in \mathbb{R}^m . The more classical Chernoff's theorem is as follows.

Let (X_i) be a sequence of real-valued iid random variable with distribution μ , such that its Laplace transform is finite in an open interval containing 0. Let h be its Cramer transform.

Let $a > EX_1$, and $S_n = X_1 + \dots + X_n$ the associated random walk, with $S_0 = 0$.

Then

$$\lim_{n \rightarrow +\infty} -\frac{1}{n} \log P \left(\frac{S_n}{n} > a \right) = h(a).$$

The reader can find a complete proof, for instance, in [8].

Empirical Laws, Chernoff's Theorem: Let (X_i) be a sequence of real-valued iid random variables with distribution μ .

Let L_n be the empirical measure associated with X_1, \dots, X_n defined by

$$L_n = \frac{1}{n} (\delta_{X_1} + \dots + \delta_{X_n})$$

where δ_{X_i} is the Dirac measure at point X_i ; if A is a Borelian in \mathbb{R} , $L_n(A)$ is the number of observed values in the n . Sample X_1, \dots, X_n which belong to A .

Let \mathcal{C} be an open set of the set of probability measures on \mathbb{R} (endowed with weak topology), then

$$\liminf \frac{1}{n} \log P_\mu (L_n \in \mathcal{C}) \geq -\inf \{ K(\nu, \mu) / \nu \in \mathcal{C} \} \quad (1)$$

$$\limsup \frac{1}{n} \log P_\mu (L_n \in \mathcal{C}) \leq -\inf \{ K(\nu, \mu) / \nu \in \mathcal{C} \} \quad (2)$$

where $K(\nu, \mu)$ is the Kullbach information of ν with respect to μ (see [6]).

APPENDIX II
PROOF OF THEOREM 2

First we prove an extension of Corollary 1 in Section I.

Theorem 1 BIS: The assumptions are those of Theorem 1. In addition let ϕ be a linear functional of the increments of X^ϵ : there exist two bounded and continuous functions u and r such that

$$\begin{aligned} \phi(X^\epsilon) &= \sum_{t \in [0, T], t \in \epsilon\mathbb{N}} [(u(X^\epsilon(t)))(X^\epsilon(t+\epsilon) - X^\epsilon(t)) \\ &\quad + \epsilon r(X^\epsilon(t))] \\ &= \int_0^T u(X^\epsilon(t)) dX^\epsilon(t) + \int_0^T r(X^\epsilon(t)) dt. \end{aligned}$$

Remark: Here ϕ replaces

$$\left\{ \sum_{t \in [0, T]} [v_{X^\epsilon(t)} [X^\epsilon(t+\epsilon) - X^\epsilon(t)] - \epsilon I_{X^\epsilon(t)}(v_{X^\epsilon(t)})] \right\} = \epsilon \log \left[\frac{dP^\epsilon}{d\bar{P}^\epsilon}(X^\epsilon) \right]$$

or some other functional of this kind.

Let B be a measurable set in Ω , satisfying

$$\begin{aligned} S(B) &= \inf \{ I(\varphi) - \phi(\varphi) / \varphi \in \bar{B} \cap \mathcal{C}_T \} \\ &= \inf \{ I(\varphi) - \phi(\varphi) / \varphi \in \bar{B} \cap \mathcal{C}_T \}. \end{aligned}$$

Then

$$\lim_{\epsilon \rightarrow 0} -\epsilon \log \int_B \exp(\phi(\omega)/\epsilon) dP^\epsilon(\omega) = S(B).$$

The proof can be found in [12]. We can now carry out the proof of Theorem 2.

Proof of Theorem 2: (See Section II for the notations.) Let P^* be the new probability defined by the mapping $x \mapsto v_x$, bounded and continuous. We denote $X^\epsilon(t)$, $X^{*\epsilon}(t)$ the processes with distributions P^ϵ and $P^{*\epsilon}$.

As $\epsilon \rightarrow 0$, we evaluate

$$\log \int_A \left(\frac{dP^\epsilon}{dP^{*\epsilon}} \right)^2 dP^{*\epsilon} = \log \int_A \frac{dP^\epsilon}{dP^{*\epsilon}} dP^\epsilon.$$

We apply Theorem 1 bis to the functional

$$\phi(\omega) = \epsilon \log \left(\frac{dP^\epsilon}{dP^{*\epsilon}}(\omega) \right)$$

with

$$\phi(\omega) = \sum_{\substack{t \in \epsilon\mathbb{N} \\ t \in [0, T]}} (v_{\omega(t)} [\omega(t+\epsilon) - \omega(t)] - \epsilon I_{\omega(t)}(v_{\omega(t)})).$$

If φ is a differentiable path, we set

$$\phi(\varphi) = \int_0^T v_{\varphi(t)} \dot{\varphi}(t) dt - \int_0^T I_{\varphi(t)}(v_{\varphi(t)}) dt.$$

Since A is a measurable set of trajectories, satisfying the conditions of Theorem 1 bis; we have

$$\log \int_A \left(\frac{dP^\epsilon}{dP^{*\epsilon}} \right) dP^\epsilon \sim -\frac{1}{\epsilon} \inf_{\tau, \varphi \in A \cap \mathcal{C}_T} \{ I(\varphi) - \phi(\varphi) \}$$

with

$$I(\varphi) = \int_0^T h_{\varphi(t)}(\dot{\varphi}(t)) dt.$$

Hence, we get

$$\begin{aligned} \log \int_A \left(\frac{dP^\epsilon}{dP^{*\epsilon}} \right)^2 dP^{*\epsilon} &\sim -\frac{1}{\epsilon} \inf_{\tau, \varphi \in A \cap \mathcal{C}_T} \left[\int h_\varphi(\dot{\varphi}) dt \right. \\ &\quad \left. - \int I_\varphi(v_\varphi) dt + \int v_\varphi \dot{\varphi} dt \right]. \end{aligned}$$

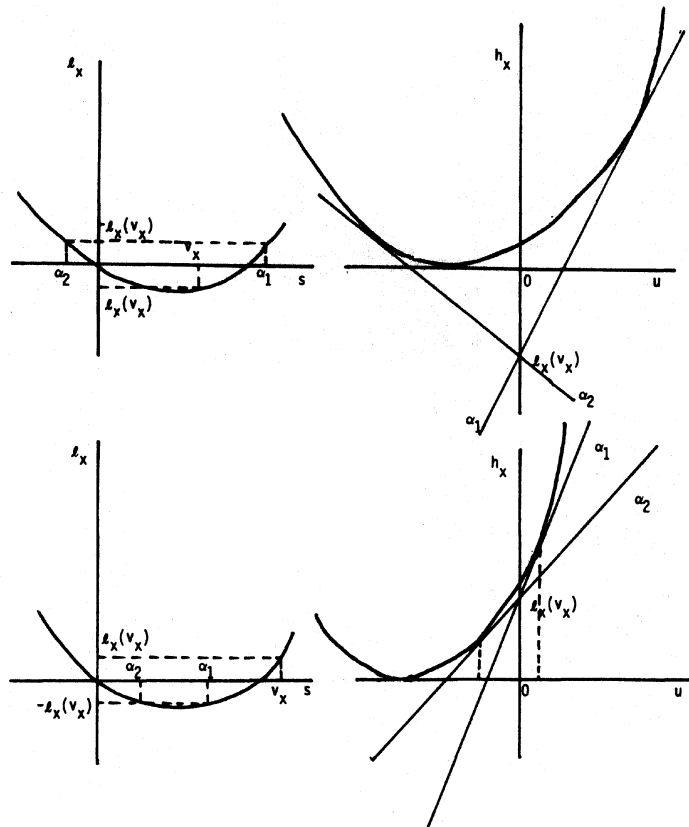


Fig. 10.

We put

$$J_v(A) = \inf_{T, \varphi \in \mathcal{A} \cap \mathcal{C}_T} \left[\int h_\varphi(\dot{\varphi}) dt - \int l_\varphi(v_\varphi) dt + \int v_\varphi \dot{\varphi} dt \right]$$

and we seek the function $x \rightarrow v_x$, which minimizes $J_v(\varphi)$.

We have

$$J_v(\varphi) = \inf_{\varphi \in \mathcal{A} \cap \mathcal{C}_{x_0}} \int_{x_0}^a \left(\frac{h_\varphi(\dot{\varphi}) - l_\varphi(v_\varphi)}{\dot{\varphi}} + v_\varphi \right) d\varphi.$$

Let $x = \varphi(t)$ be fixed, then

$$U(x) = \inf_u \left[\frac{h_x(u) - l_x(v_x)}{u} + v_x \right]$$

is achieved for \bar{u} such that $\bar{u}h'_x(\bar{u}) - l_x(v_x) + h_x(\bar{u}) = \bar{u}h'_x(\bar{u}) - h_x(\bar{u})$.

Because of the reciprocity of the Cramer transformation, we have

$$l_x(v_x) = \sup_w (v_x w - h_x(w)) = v_x \bar{w} - h_x(\bar{w})$$

where \bar{w} is unique, and satisfies $h'_x(\bar{w}) = v_x$, hence

$$l_x(v_x) = h'_x(\bar{w})\bar{w} - h_x(\bar{w}). \tag{1}$$

Comparing to the relation (1), we deduce that \bar{u} is the point satisfying $h'_x(\bar{u}) = v_x^1$, where v_x^1 is such that $l'_x(v_x^1) = -l_x(v_x)$. Generally we obtain two solutions described in Fig. 10.

In each case, α_1 corresponds to $\bar{u} > 0$, and α_2 to $\bar{u} < 0$. Thus the path $\varphi_v \in \mathcal{C}$, which gives the minimum $\mathfrak{J}_v(\varphi)$, satisfies $h'_{\varphi(t)}(\dot{\varphi}(t)) = v_x^1, v_x^1$ being α_1 or α_2 . In all the cases above, we have

$$J_v(\varphi_v) = \int (v_x^1 v_{\varphi_v(t)} + v_{\varphi_v(t)}) d\varphi_v(t)$$

and thus

$$\log \int_A \left(\frac{dP^\epsilon}{dP^{*\epsilon}} \right) dP^{*\epsilon} \underset{\epsilon \rightarrow 0}{\sim} -\frac{1}{\epsilon} \int (v_x^1 v_{\varphi_v(t)} + v_{\varphi_v(t)}) d\varphi_v(t).$$

But the function l_x is convex, and for each $x, l_x(v_x^1) + l_x(v_x) = 0$ implies that $(v_x^1 + v_x)/2 \leq \lambda_x$ with $l_x(\lambda_x) = 0$. So for each $x, v_x^1 + v_x$ is maximum where $v_x^1 = v_x = \lambda_x$. This achieves the proof.

ACKNOWLEDGMENT

The authors would like to thank Pr. Ruget who has suggested this work, and has guided them by stimulating discussions.

REFERENCES

- [1] N. Abramson, "The throughput of packets broadcasting channels," *IEEE Trans. Commun.*, vol. COM-25, Jan. 1977.
- [2] R. Azencott and G. Ruget, "Mélanges d'équations différentielles, grands écarts à la loi des grands nombres," *Z. Wahrscheinlichkeitstheorie verm. Gebiete*, vol. 38, pp. 1-54, 1977.
- [3] A. Benveniste, "Outils et modèles mathématiques pour l'analyse des systèmes," presented at the Séminaire des Techniques Probabilistes

- en automatique et télécommunications, Paris, France, Mar. 20-21, 1980, RCP 567, pp. 13-51.
- [4] A. Benveniste, M. Goursat, and G. Ruget, "Analysis of stochastic approximation schemes with discontinuous and dependent forcing terms with applications to data communication algorithms," *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 1042-1058, Dec. 1980.
 - [5] P. Billingsley, *Convergence of Probability Measures*. New York: Wiley, 1968.
 - [6] J. Bretagnolle, *Astérisque* 68, Soc. Math. France, 1979, pp. 33-53.
 - [7] M. Cottrell, J. C. Fort, and G. Malgouyres, "Evénements rares pour l'étude de certains algorithmes stochastiques," *l'Univ. Paris-Sud, Orsay, France, Prepubl.* 80 T 35.
 - [8] D. Dacunha-Castelle, *Astérisque* 68, Soc. Math. France, 1979, pp. 19-25.
 - [9] S. Dreyfus, *Dynamic Programming and the Calculus of Variation*. New York: Academic, 1965.
 - [10] G. Fayolle, E. Gelenbe, and J. Labetoulle, "Stability and optimal control of the packet switching broadcast channels," *JACM*.
 - [11] W. Feller, *An Introduction to Probability and its Applications, Vol. II*. New York: Wiley, 1971.
 - [12] J. C. Fort, These de 3ème cycle, Univ. Paris-Sud, Orsay, France, 1980.
 - [13] D. Gaver and G. Thompson, *Programming and Probability Models in Operations Research*. Monterey, CA: Brooks/Cole, 1972.
 - [14] I. I. Gihman and A. V. Skorohod, *Stochastic Differential Equations*. New York: Springer-Verlag, 1972.
 - [15] H. J. Kushner and Haihuang, "On the weak convergence of a sequence of general stochastic difference equations to a diffusion," *Lévesque Center for Dynam. Syst., Division of Appl. Math., Brown Univ., Providence, RI*, 1979.
 - [16] S. Kullback, *Information Theory and Statistics*. New York: Dover, 1968.
 - [17] S. Lam, "Packet switching in a multiaccess broadcast channel," S.E.A.S., Univ. Calif., 1964.
 - [18] D. Ludwig, "Persistence of dynamical system under random perturbation," *SIAM Rev.*, vol. 17, pp. 605-640, 1975.
 - [19] O. Penrose and L. Lebowitz, *Fluctuations Phenomena: Studies in Statistical Mechanics*, J. L. Lebowitz and E. W. Montroll, Eds. Amsterdam, The Netherlands: North-Holland.
 - [20] D. Perriot-Mathonna, "Stopping times for cumulative sums of mixtures and approach by the large deviations theory," presented at the 20th CDC, 1981.
 - [21] A. D. Ventsel, "Rough limit theorems on large deviations for Markov stochastic processes II," *Theory Prob. Appl. (USSR)*, vol. 21, pp. 499-512, 1976.
 - [22] A. D. Ventsel and Freidlin, *Fluktnacii v dinamicheskikh sistemakh mal'kikh sluchainykh vozmushchenii*. Moscow, USSR: Nauka, 1979.



Marie Cottrell was born in Bethune, France, on December 12, 1943. She was a student at the Ecole Normale Supérieure de Jeunes Filles from 1961 to 1964 and received the Agrégation de Mathématiques degree in 1964, and the Pure Mathematics DEA in 1969.

From 1964 to 1967, she was a High School Teacher, and from 1968 to 1969 she was Assistant at the University of Paris, Paris, France, and the Université Paris-Sud, Orsay, France. From 1970 to 1973, she was Professor at the University of Habana, Cuba. Currently, she is Maitre-Assistante at the Université Paris-Sud. Her research interests include stochastic algorithms, rare events, large deviations, biometry, and biomathematics.



Jean-Claude Fort was born in Lyon, France, in November 1953. He received the "Agrégation de Mathématiques" degree in 1976 and the "thèse de 3ème cycle" degree in 1980, from the Université Paris-Sud, Orsay, France.

He is currently Assistant at the Université de Nancy, Nancy Cedex, France. His present research interests are stochastic population dynamic and stochastic algorithms.



Gérard Malgouyres was born in Toulouse, France, in 1942. He studied in the University of Toulouse, Toulouse, France.

Presently, he is Maitre-Assistant at the Université Paris-Sud, Orsay, France, where his research includes large deviation and the Ising-model.

DEUXIEME PARTIE

ETUDE D'UN PROCESSUS D'AUTO-ORGANISATION

ET

UN MODELE STOCHASTIQUE DE LA RETINOTOPIE :
UN PROCESSUS D'AUTO-ORGANISATION

Étude d'un processus d'auto-organisation

par

Marie COTTRELL
Université Paris-Sud
U. A. 743 C. N. R. S.
Mathématiques, Statistique Appliquée
Bât. 425
F-91405 Orsay (France)

Jean-Claude FORT
Université Paris V
Laboratoire de Statistiques
Médicales
45, rue des Saints-Pères
F-75270 Paris Cedex 06 (France)

RÉSUMÉ. — T. Kohonen a défini en [2] un algorithme pour modéliser le phénomène biologique appelé rétinotopie : il s'agit de l'établissement de relations structurées entre la rétine et le cortex, qui préservent la topologie de la rétine. Il a observé la convergence de l'algorithme par simulation et proposé en [4] un schéma de démonstration avec quelques éléments de preuves. Dans cet article, nous démontrons la convergence vers une distribution stationnaire, dans le cas uni-dimensionnel.

Mots-clés : Chaînes de Markov, distribution invariante, rétinotopie.

ABSTRACT. — T. Kohonen defined in [2] a mathematical algorithm in order to modelize the biological phenomena called retinotopy: that is the self-organized formation of maps between retina and cortex preserving the topological structure of the retina. He observed the convergence of the algorithm by simulation and presented in [4] a sketch of proof. In this paper we prove the convergence toward an invariant distribution in the one dimensional case.

Key-words: Markov chains, invariant distribution, retinotopy.

INTRODUCTION

T. Kohonen (1982) a proposé un modèle qui décrit l'établissement de relations structurées entre les cellules du cortex et de la rétine.

Initialement, la liaison rétine-cortex est anarchique. Au cours de la vie embryonnaire, et par l'activité visuelle, s'établit une relation qui conserve la topologie de la rétine.

Le principe de Hebb affirme que la variation de l'efficacité de la liaison synaptique entre un neurone de la rétine et un neurone du cortex est proportionnelle au produit de l'activité pré-synaptique (stimulus) et post-synaptique (excitation d'une cellule du cortex et de ses voisines).

En conséquence, si une cellule de la rétine est excitée, elle transmet cette excitation à une zone du cortex et les cellules excitées dans le cortex renforcent alors leur liaison avec la cellule « émettrice ».

Kohonen (dans [2] [3]) a proposé l'algorithme suivant :

La rétine est représentée par le carré $[0, 1]^2$. Le cortex est un réseau discret bidimensionnel : $\{1, \dots, n\}^2$. Ses cellules sont notées (i, j) . Il est clair qu'une cellule du cortex est contactée par plusieurs cellules de la rétine, et réciproquement.

Par souci de simplification, on suppose qu'une cellule du cortex (i, j) est reliée à une cellule de la rétine notée x_{ij} , centre de gravité des cellules de la rétine reliées à (i, j) pondérées par les efficacités synaptiques.

A l'instant t , on note $X^t = (X_{ij}^t, (i, j) \in [1, n]^2)$ l'état des liaisons. Initialement X^0 correspond au tirage de n^2 variables aléatoires uniformes indépendantes.

A chaque instant $t = 1, \dots, n$, on excite la rétine en un point ω^{t+1} (tiré suivant une loi uniforme). On détermine ainsi, presque sûrement, l'indice « appelé » (i_0, j_0) tel que X_{i_0, j_0}^t soit le point le plus proche de ω^{t+1} .

L'excitation est alors transmise à la cellule (i_0, j_0) du cortex, et à une zone voisine.

Le renforcement des liaisons se traduit par une homothétie de centre ω^{t+1} et de rapport $(1 - \varepsilon)$, ε étant petit, compris entre 0 et 1, qui rapproche de ω^{t+1} et entre elles les liaisons concernées.

Cet algorithme se généralise immédiatement en dimension quelconque finie.

Ce type d'algorithme d'auto-organisation est utilisé pour des applica-

tions variées : reconnaissance de la parole, choix de la dimension d'une représentation réduite d'un ensemble de points de \mathbb{R}^n , construction de représentations hiérarchisées (Kohonen (1984) [4]).

Les simulations.

Kohonen dans [2] a exposé les résultats des simulations faites à partir de cet algorithme, pour $(i, j) \in [1, 8]^2$ et 8 voisins. On constate effectivement la convergence de l'algorithme lorsque $t \rightarrow +\infty$, non pas vers une disposition constante, mais vers une configuration aléatoire reproduisant la disposition du carré des (i, j) .

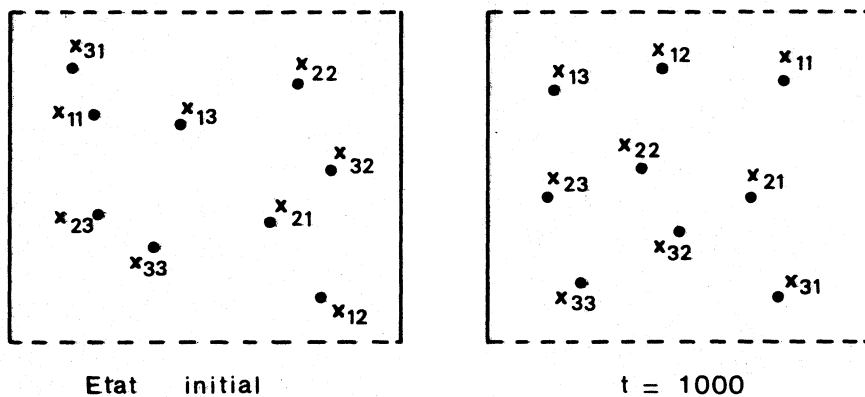


FIG. 1. — $(i, j) \in \{1, 2, 3\}^2$, $\varepsilon = 0,05$.

En dimension 1, on obtient une configuration (aléatoire) monotone, application de $\{1, 2, \dots, n\}$ dans $[0, 1]$.

Dans cet article, nous démontrons la convergence de la chaîne X^t dans le cas unidimensionnel.

1. ALGORITHME DE KOHONEN EN DIMENSION 1

En dimension 1, l'algorithme est défini par :

- a) Un scalaire ε compris entre 0 et 1.
- b) Une suite $\omega^1, \omega^2, \dots, \omega^t, \dots$ de v. a. iid, de loi uniforme sur $[0, 1]$.

c) La chaîne de Markov (X^t) à valeurs dans $[0, 1]^n$, de loi initiale la loi uniforme, vérifiant

$$X_i^{t+1} = X_i^t + \varepsilon(\omega^{t+1} - X_i^t) \quad \text{pour } i \in \{i_0 - 1, i_0, i_0 + 1\} \cap [1, n]$$

$$X_i^{t+1} = X_i^t \quad \text{pour } i \notin \{i_0 - 1, i_0, i_0 + 1\} \cap [1, n],$$

où $i_0 = i(X^t, \omega^{t+1})$ est défini presque sûrement par

$$|\omega^{t+1} - X_{i_0}^t| = \inf_i |\omega^{t+1} - X_i^t|.$$

On note $X^{t+1} = H(X^t, \omega^{t+1})$, on remarque que H est définie sur l'ensemble E où

$$E = \{ (x, \omega) / x \in [0, 1]^n, \omega \in [0, 1], x_i \neq x_j \quad \forall i \neq j, i(x, \omega) \text{ unique} \}$$

Il est commode de représenter une réalisation de X^t par un graphe

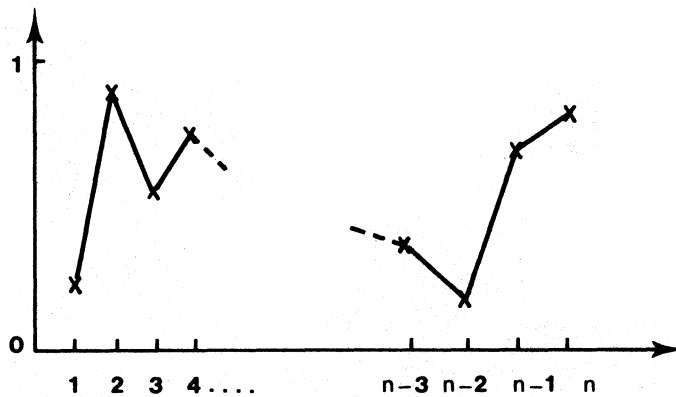


FIG. 2. — Graphe de (X^t) à l'instant t .

Pour bien faire comprendre le fonctionnement de l'algorithme, nous commençons par faire les observations suivantes :

DÉFINITION 1.1. — *Le nombre d'inversions de X^t est*

$$N_t = \text{Card} \{ i \in \{2, \dots, n-1\} / (X_{i+1}^t - X_i^t)(X_i^t - X_{i-1}^t) < 0 \}.$$

PROPOSITION 1.1. — *Le nombre d'inversions du graphe X^t est une fonction décroissante (au sens large) du temps.*

Démonstration. — Supposons $X^t = (x_1, \dots, x_n)$ et $X^{t+1} = (x'_1, \dots, x'_n)$

et pour envisager le cas général $i_0 = i(\omega^{t+1}, X^t)$ avec $i_0 \in [4, n - 3]$. Posons pour plus de commodité $\omega^{t+1} = z$. Si

$$S'(i_0 - 1, i_0, i_0 + 1) = \text{Signe}(x_{i_0} - x_{i_0-1})(x_{i_0+1} - x_{i_0})$$

il est bien clair que $S^{t+1}(i_0 - 1, i_0, i_0 + 1) = S^t(i_0 - 1, i_0, i_0 + 1)$.

Le nombre des inversions n'est jamais modifié en i_0 . Donc seules nous intéressent les inversions possibles en $i_0 - 2, i_0 - 1$ (ou $i_0 + 2, i_0 + 1$, ce qui se traite de la même manière).

En tenant compte des symétries, seuls 4 cas se présentent :

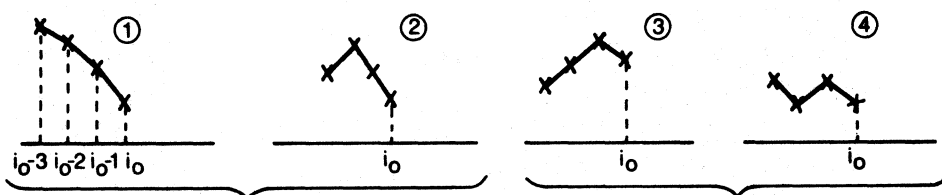


FIG. 3. — Pas d'inversion en $i_0 - 1$.

FIG. 4. — Une inversion en $i_0 - 1$.

Dans tous les cas $x'_{i_0-3} = x_{i_0-3}$ et $x'_{i_0-2} = x_{i_0-2}$.

⊙ 1 et 2 :

$x'_{i_0-2} - x'_{i_0-1} = (x_{i_0-2} - x_{i_0-1}) - \varepsilon(z - x_{i_0-1})$ reste > 0 , car dans ce cas $x_{i_0-1} > z$.

Donc $N^{t+1} = N^t$.

⊙ 3 et 4 :

a) Si $x_{i_0-1} - x_{i_0-2} < \varepsilon(x_{i_0-1} - z)$ alors $x'_{i_0-2} - x'_{i_0-1} > 0$, alors dans le cas 3, $S^{t+1}(i_0 - 3, i_0 - 2, i_0 - 1) < 0$ mais $S^{t+1}(i_0 - 2, i_0 - 1, i_0) > 0$ donc $N^{t+1} = N^t$: la pointe se déplace de $i_0 - 1$ à $i_0 - 2$.

Dans le cas 4, $S^{t+1}(i_0 - 3, i_0 - 2, i_0 - 1) > 0$ et également $S^{t+1}(i_0 - 2, i_0 - 1, i_0) > 0$ donc $N^{t+1} = N^t - 2$: 2 pointes s'effacent simultanément.

b) Si $x_{i_0-1} - x_{i_0-2} \geq \varepsilon(x_{i_0-1} - z)$ comme précédemment $N^{t+1} = N^t$.

Les cas particuliers $i_0 = 1, 2, 3, n - 2, n - 1, n$ se traitent de la même manière avec cependant la possibilité d'avoir $N^{t+1} = N^t - 1$: une pointe « tombe » au bord du graphe. □

La fonction N^t n'est pas strictement décroissante et de ce fait ne peut pas être utilisée directement comme fonction de Lyapounov associée au problème.

Toutefois on en déduit que l'ensemble des graphes strictement mono-

tones est absorbant pour X^t . Il est constitué de deux classes absorbantes qui ne communiquent pas : A , les graphes strictement croissants, A' , les graphes strictement décroissants.

Dans la suite, nous démontrons deux résultats essentiels.

THÉORÈME 1.1. — Soit $\tau_{A \cup A'}$ le temps d'entrée dans $A \cup A'$. Soit \mathbb{P}_x la loi de la chaîne de Markov X^t , partant de l'état x à coordonnées toutes distinctes, alors

$$\mathbb{P}_x(\tau_{A \cup A'} < +\infty) = 1.$$

La démonstration de ce théorème fait l'objet du § 2.

THÉORÈME 1.2. — Conditionnellement à l'entrée dans A (resp. A'), la chaîne (X^t) converge en loi vers une distribution stationnaire unique ${}^s\pi$ qui possède une partie absolument continue par rapport à la mesure de Lebesgue.

Ce résultat est démontré au § 3.

Le § 4 est consacré à quelques propriétés complémentaires.

2. DÉMONSTRATION DU THÉORÈME 1.1. (MISE EN ORDRE)

Tout d'abord, on peut mettre en évidence une conséquence de la proposition 1.1 :

LEMME 2.1. — Si les deux inégalités suivantes sont réalisées

$$\begin{aligned} X_{i_0}^t &< X_{i_0-2}^t < X_{i_0-1}^t \\ X_{i_0-1}^t - X_{i_0-2}^t &< \varepsilon(X_{i_0-1}^t - X_{i_0}^t) \end{aligned}$$

alors un tirage de ω^{t+1} très proche de $X_{i_0}^t$ réordonne le triplet dans l'ordre décroissant.

$$\begin{aligned} \text{On a } A &= \{x = (x_1, \dots, x_n) / 0 < x_1 < \dots < x_n < 1\} \\ A' &= \{x = (x_1, \dots, x_n) / 0 < x_n < \dots < x_1 < 1\} \end{aligned}$$

On note $x_0 = 0, x_{n+1} = 1$.

Le théorème 1.1 est un corollaire immédiat du résultat suivant.

THÉORÈME 2.1. — Soit $\tau_{A \cup A'}$ le temps d'entrée dans $A \cup A'$. Il existe $T < +\infty$ et $\alpha(n, \varepsilon) > 0$ tels que

$$\forall x^0 \in [0, 1]^n, \text{ à coordonnées toutes distinctes } \quad \mathbb{P}_{x^0}(\tau_{A \cup A'} \leq T) \geq \alpha(n, \varepsilon).$$

Nous décomposons la preuve du Théorème 2.1 en deux lemmes dans les démonstrations desquels nous laisserons quelques détails techniques au lecteur.

DÉFINITION 2.1. — On dit que $x \in [0, 1]^n$ contient une (k, η) chaîne de début p et de fin q si $p < q$, $q - p + 1 = k$, et $\eta > 0$ avec

i) $x_p < x_{p+1} < \dots < x_q$ (1)

ou $x_p > x_{p+1} > \dots > x_q$ (2)

ii) $\inf \{ |x_{i+1} - x_i| / i = p, p+1, \dots, q-1 \} \geq \eta$

iii) $\forall j = 1, \dots, p-1, q+1, \dots, \eta$
 $x_j < x_p - \eta$ ou $x_j > x_q + \eta$ dans le cas (1)
 $x_j > x_p + \eta$ ou $x_j < x_q - \eta$ dans le cas (2)

η est le module de la (k, η) chaîne.

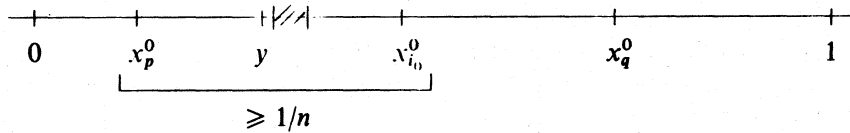
En notant $B_{k,\eta}$ l'ensemble des x contenant une (k, η) chaîne, il est clair que

$$A \cup A' = \bigcup_{\eta > 0} B_{k,\eta}.$$

Soit $\tau_{B_{k,\eta}}$ le temps d'entrée dans $B_{k,\eta}$.

LEMME 2.1. — Construction d'une 3-chaîne. — Il existe $T_3 < +\infty$, $\eta > 0$ et $\alpha(\varepsilon, n)$ tels que quel que soit x de $[0, 1]^n$ à coordonnées toutes distinctes, $\mathbb{P}_{x_0}(\tau_{B_{3,\eta}} \leq T_3) \geq \alpha(\varepsilon, n) > 0$.

Démonstration. — a) Soit $x^0 = (x_1^0, \dots, x_n^0)$ un point quelconque de $[0, 1]^n$, à coordonnées toutes distinctes. Il existe nécessairement un couple de points écartés de plus de $1/n$.



On peut trouver $i_0, p \neq i_0 - 1$ et $i_0 + 1, q \neq i_0 - 1, i_0 + 1$ tels que

$$x_{i_0}^0 - x_p^0 \geq \frac{1}{n}$$

x_p^0 est le point le plus proche sur la gauche de $x_{i_0}^0$, avec $p \neq i_0 \pm 1$.

x_q^0 est le point le plus proche sur la droite de $x_{i_0}^0$, avec $q \neq i_0 \pm 1$.

On note $y = \frac{1}{2}(x_p^0 + x_{i_0}^0)$.

Soit T un entier > 0 . Pour α et β positifs tels que

$$\beta(1 - (1 - \varepsilon)^T) < \alpha < \beta < \frac{(1 - \varepsilon)^T}{2n},$$

et si $\omega^1, \dots, \omega^T$ appartiennent à $]y + \alpha, y + \beta[$, on a pour tout $t = 1, \dots, T$,

$$i) \quad X_{i_0}^t > (1 - \varepsilon)^T x_{i_0}^0 + (1 - (1 - \varepsilon)^T)y$$

et comme

$$X_{i_0}^t > (1 - \varepsilon)^T \left(\frac{x_{i_0}^0 - x_p^0}{2} \right) + y > \beta + y$$

et de même que

$$\frac{1}{2}(X_{i_0}^t + X_p^t) < (1 - \varepsilon)^T y + (1 - (1 - \varepsilon)^T)(y + \beta)$$

(car le déplacement du point X_{i_0} est maximum pour $\omega^t = y + \alpha$), on a

$$\frac{1}{2}(X_{i_0}^t + X_p^t) < y + (1 - (1 - \varepsilon)^T)\beta < y + \alpha$$

et par conséquent

$$ii) \quad i(X^{t-1}, \omega^t) = i_0 \quad \text{pour } t = 1, \dots, T.$$

Après T tirages dans $]y + \alpha, y + \beta[$ on a :

$$iii) \quad \sup(|X_{i_0}^T - X_{i_0-1}^T|, |X_{i_0}^T - X_{i_0+1}^T|) \leq (1 - \varepsilon)^T$$

$$iv) \quad X_{i_0}^T \in]y + \beta, y + \beta + (1 - \varepsilon)^T[.$$

$$D'où \quad \begin{cases} X_{i_0}^T, X_{i_0-1}^T, X_{i_0+1}^T \in]y + \beta - (1 - \varepsilon)^T, y + \beta + 2(1 - \varepsilon)^T[\\ \text{et } \forall j \neq i_0, i_0 - 1, i_0 + 1, \quad X_j^T = x_j^0. \end{cases}$$

La probabilité de cet événement est $\geq (\beta - \alpha)^T$.

Au temps T , on se trouve donc dans l'une des 6 configurations suivantes :

$$(1) \quad \begin{array}{c} | \quad | \quad | \\ \hline X_{i_0-1}^T \quad X_{i_0}^T \quad X_{i_0+1}^T \\ \hline \end{array}$$

$$(2) \quad \begin{array}{c} | \quad | \quad | \\ \hline X_{i_0+1}^T \quad X_{i_0}^T \quad X_{i_0-1}^T \\ \hline \end{array}$$

$$(3) \quad \begin{array}{c} | \quad | \quad | \\ \hline X_{i_0-1}^T \quad X_{i_0+1}^T \quad X_{i_0}^T \\ \hline \end{array}$$

$$(4) \quad \begin{array}{c} | \quad | \quad | \\ \hline X_{i_0+1}^T \quad X_{i_0-1}^T \quad X_{i_0}^T \\ \hline \end{array}$$

$$(5) \quad \begin{array}{c} | \quad | \quad | \\ \hline X_{i_0}^T \quad X_{i_0-1}^T \quad X_{i_0+1}^T \\ \hline \end{array}$$

$$(6) \quad \begin{array}{c} | \quad | \quad | \\ \hline X_{i_0}^T \quad X_{i_0+1}^T \quad X_{i_0-1}^T \\ \hline \end{array}$$

b) Avec une probabilité minorée, on peut ramener les situations (3) et (5) à la situation (1) (resp. (4) et (6) à (2)), au temps $T + 1$ (en choisissant T convenablement).

Prenons par exemple le cas de la situation (3).

Si l'on choisit $\omega^{T+1} \in]u + \alpha, u + \beta[$ avec $u = \frac{1}{2}(x_p^0 + y)$, l'indice « appelé » est $i_0 - 1$ et X_{i_0} est déplacé.

On peut alors appliquer le lemme 2.1 en imposant $4n(1 - \varepsilon)^T \leq \varepsilon$, et on est ramené à la situation (1).

Pour les calculs qui suivent, on prend en réalité T défini par

$$15n(1 - \varepsilon)^T \leq \varepsilon < 15n(1 - \varepsilon)^{T-1}.$$

Au temps T + 1, on se retrouve donc dans l'un des cas (1) ou (2). On suppose dans la suite qu'il s'agit du cas (1).

c) Module d'espacement de la 3-chaîne.

On a au temps T + 1

$$u = \frac{1}{2}(x_p^0 + y) < X_{i_0-1}^{T+1} < X_{i_0}^{T+1} < X_{i_0+1}^{T+1} < \frac{1}{2}(y + x_{i_0}^0) = v,$$

avec $v - u \geq \frac{1}{2n}$ et $X_{i_0}^{T+1} - X_{i_0-1}^{T+1} \geq \frac{\varepsilon}{5n}$.

Alors on vérifie qu'il existe $a > 0$ indépendant de x^0 , tel que si $\omega^{T+2} \in]u, u + a[$ et $\omega^{T+3} \in]v - a, v[$, on obtient au temps T + 3, une 3-chaîne $(i_0 - 1, i_0, i_0 + 1)$ de module d'espacement indépendant de x^0 . □

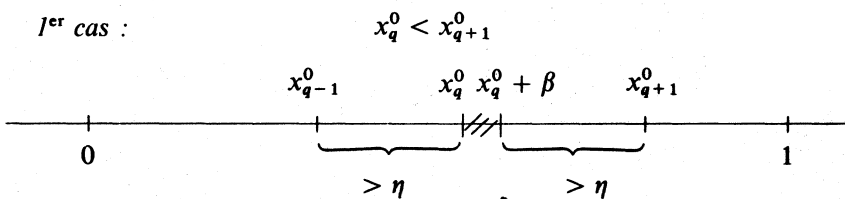
LEMME 2.2. — Pour $k \geq 3$, il existe une fonction h continue de \mathbb{R}^+ dans \mathbb{R}^+ , vérifiant $h(0) = 0$, un entier $T_{k+1} > 0$, et $\alpha_{k+1}(n, \varepsilon, \eta) > 0$ tels que

$$\forall x^0 \in B_{k,\eta} \quad \mathbb{P}_{x^0}(\tau_{B_{k+1},n(n)} \leq T_{k+1}) \geq \alpha_{k+1}(n, \varepsilon, \eta).$$

Ce lemme donne une minoration uniforme de la probabilité de passer d'une k -chaîne ordonnée à une $k + 1$ -chaîne, en conservant un module d'espacement pouvant être différent, mais que l'on sait calculer.

Démonstration. — Il n'y a rien à démontrer si $n \leq 3$. On se borne au cas d'une chaîne croissante, lorsque $n > 3$.

1^{er} cas :



On cherche à déterminer un β non nul, tel que si $\omega^t \in]x_q^0, x_q^0 + \beta[$ pour tout t moindre que T, on ait $i(x^{t-1}, \omega^t) = q$, de manière à « attirer

au plus vite x_{q+1} » et constituer une $(k + 1)$ chaîne de module $h(\eta)$ calculable, cela en majorant le temps T nécessaire.

Soit T tel que $(1 - \varepsilon)^T \leq \frac{\eta}{2} < (1 - \varepsilon)^{T-1}$ et $\beta = \inf \left\{ \frac{\eta}{2} (1 - \varepsilon)^T, \frac{\eta}{2} - \frac{\eta^2}{2} (1 - \varepsilon) \right\}$.

Alors on voit, par récurrence, que si pour $t \leq T$, $x_q^0 < \omega^t < x_q^0 + \beta$, on a pour tout $t = 1, \dots, T$,

- i) $x_q^0 < X_q^t < x_q^0 + \beta$
- ii) $i(X^{t-1}, \omega^t) = q$

et que X^t contient une $(k + 1, h(\eta))$ chaîne de début p et de fin $q + 1$, avec $h(\eta) \frac{\eta^2}{2} (1 - \varepsilon)$. D'où le résultat avec $T_{k+1} = T$ et $\alpha_{k+1}(n, \varepsilon, \eta) = \beta^T$.

On traite le cas symétrique $x_{p-1}^0 < x_p^0$ de la même façon. Reste donc le

2^e cas : $x_{q+1}^0 < x_p^0$ et $x_{p-1}^0 > x_q^0$.

a) Soit $a = \frac{\varepsilon^2 \eta}{2(1 + \varepsilon)}$, T défini par $(1 - \varepsilon)^T \leq a < (1 - \varepsilon)^{T-1}$, et $\beta \leq \frac{\eta}{2} (1 - \varepsilon)^T$.

Alors si pour $t = 1, \dots, T$, $\omega^t \in]x_q^0 - \beta, x_q^0[$, on a

i) $i(X^{t-1}, \omega^t) = q$ pour $t = 1, \dots, T$, et

ii) $X_q^T - X_{q+1}^T < a$.

b) En choisissant $\omega^{T+1} \in]x_{q-2}^0 - \frac{\eta}{2}, x_{q-2}^0[$, on a

i) $i(X^T, \omega^{T+1}) = q - 2$ et

ii) $X_{q-1}^{T+1} < X_q^{T+1} - \frac{\varepsilon \eta}{\varepsilon + 1} < x_q^0 - a$.

c) Enfin pour $\omega^{T+2} \in]X_{q-1}^{T+1} - \frac{(1 - \varepsilon)\eta}{2}, X_{q-1}^{T+1}[$ on a $i(X^{T+1}, \omega^{T+2}) = q - 1$

et on obtient alors une $(k + 1)$ chaîne de début p et de fin $q + 1$.

On peut vérifier que son module d'espacement est minoré par $\frac{(1 - \varepsilon)^2 \varepsilon^2 \eta^2}{2(1 + \varepsilon)}$. \square

De ces deux lemmes, découlent par récurrence immédiate le théorème 2.1 et son corollaire le théorème 1.1.

3. DÉMONSTRATION DU THÉORÈME 1.2. (EXISTENCE D'UNE DISTRIBUTION STATIONNAIRE)

Nous étudions l'existence d'une distribution stationnaire sur A (resp. sur A'). Pour cela, nous montrons que la chaîne est Doeblin [1].

Nous pouvons déjà remarquer que la transition $\mathbb{P}(\cdot, \cdot)$ de \mathbb{R}^n dans \mathbb{R}^n possède un support de dimension 1 et est donc singulière par rapport à la mesure de Lebesgue. La plus petite itérée de \mathbb{P} susceptible de posséder une partie absolument continue par rapport à la mesure de Lebesgue est donc \mathbb{P}^n , nous allons voir que c'est effectivement le cas.

La démonstration est faite pour l'ensemble absorbant A.

Nous démontrons tout d'abord le

THÉORÈME 3.1. — *Pour tout x de A, il existe un borélien C_x , deux réels strictement positifs a et δ tels que*

i) C_x soit de mesure de Lebesgue strictement positive ;

ii) Uniformément sur le voisinage $\prod_{i=1}^n \left[x_i - \frac{a}{2}, x_i + \frac{a}{2} \right]$ on ait :

$$\mathbb{P}^n(\xi, d\zeta) \geq \delta \mathbb{1}_{C_x}(\zeta) d\zeta.$$

Bien entendu C, a, δ dépendent aussi de n et du rapport d'homothétie $1 - \varepsilon$.

Démonstration. — Soit $x \in A$. Si $X^0 = x$, nous construisons une suite d'intervalles I_1, I_2, \dots, I_n telle que

$$(\forall \omega^j \in I_j) \quad i(X^{j-1}, \omega^j) = j \text{ si } n \equiv 0 \text{ ou } 1 \text{ modulo } 3.$$

$$(\forall \omega^j \in I_j) \quad i(X^{j-1}, \omega^j) = j + 1 \text{ et } i(X^{n-1}, \omega^n) = 1 \text{ si } n \equiv 2 \text{ modulo } 3.$$

1) Supposons une telle suite d'intervalles construite. On a $X^n = H(\dots H(H(x, \omega^1), \omega^2) \dots, \omega^n)$ (notations du § 1).

a) Si $n \equiv 0$ ou 1 modulo 3, en notation matricielle :

$$X^n = \text{diag} \begin{bmatrix} (1 - \varepsilon)^2 \\ (1 - \varepsilon)^3 \\ \vdots \\ (1 - \varepsilon)^3 \\ (1 - \varepsilon)^2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + D_n \begin{bmatrix} \omega^1 \\ \omega^2 \\ \vdots \\ \omega^{n-1} \\ \omega^n \end{bmatrix} = C_n x + D_n \omega$$

où

$$D_n = \begin{bmatrix} (1-\varepsilon)\varepsilon & \varepsilon & 0 & \dots & 0 \\ (1-\varepsilon)^2\varepsilon & (1-\varepsilon)\varepsilon & \varepsilon & & \\ 0 & (1-\varepsilon)^2\varepsilon & (1-\varepsilon)\varepsilon & \varepsilon & \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & (1-\varepsilon)^2\varepsilon & \varepsilon \\ 0 & \dots & 0 & (1-\varepsilon)\varepsilon & \varepsilon \\ 0 & \dots & 0 & (1-\varepsilon)\varepsilon & \varepsilon \end{bmatrix}$$

et on vérifie $\det D_n \neq 0$.

b) Si $n \equiv 2$ modulo 3 :

$$X^n = \text{diag} \begin{bmatrix} (1-\varepsilon)^2 \\ (1-\varepsilon)^3 \\ \vdots \\ (1-\varepsilon)^3 \\ (1-\varepsilon)^2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + D_n \begin{bmatrix} \omega^n \\ \omega^1 \\ \vdots \\ \omega^{n-1} \end{bmatrix} = C_n x + D_n \omega$$

où

$$D_n = \begin{bmatrix} \varepsilon & \varepsilon(1-\varepsilon) & 0 & \dots & 0 \\ \varepsilon & \varepsilon(1-\varepsilon)^2 & \varepsilon(1-\varepsilon) & 0 & \\ 0 & \varepsilon(1-\varepsilon)^2 & \varepsilon(1-\varepsilon) & \varepsilon & \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & \varepsilon(1-\varepsilon)^2 & \varepsilon(1-\varepsilon) \\ 0 & \dots & 0 & \varepsilon(1-\varepsilon) & \varepsilon \\ 0 & \dots & 0 & \varepsilon(1-\varepsilon) & \varepsilon \end{bmatrix}$$

et là encore $\det D_n \neq 0$.

Soit \mathcal{C}_n l'opérateur associé à D_n et \mathcal{C}_n l'opérateur associé à C_n .

$$(\forall (\omega^1, \dots, \omega^n) \in I_1 \times \dots \times I_n) \quad (\forall X^0 \in V_x) \quad |X_j^n - X_j^{n,x}| \leq (1-\varepsilon)^2 |X_j^0 - x_j|$$

où $X^{n,x}$ est la position à l'instant n partant précisément de x et V_x un voisinage de x .

Si donc $V_x \subset J_1 \times \dots \times J_n$ où les J_k sont des intervalles centrés sur x_k inclus strictement dans I_k , l'ensemble

$$C_x = \bigcap_{x^0 \in V_x} [\mathcal{D}_n(I_1 \times I_2 \dots \times I_n) + \mathcal{C}_n X^0]$$

est de mesure de Lebesgue non nulle et on a :

$$(\forall X^0 \in V_x) \quad \mathbb{P}^n(X^0, d\zeta) \geq \frac{1}{|\det D_n|} \mathbf{1}_{C_x}(\zeta) d\zeta.$$

2) Il reste donc à prouver l'existence des I_j . Il est facile de montrer que la restriction de la fonction H à l'ensemble $(A \times [0, 1]) \cap E$ (resp. $(A' \times [0, 1]) \cap E$) est continue.

On en déduit le lemme suivant (la norme choisie est $\|x\| = \sup |x_i|$).

LEMME 3.1. — Soit z_1, \dots, z_l , appartenant à $[0, 1]$. Soit $x \in A$. On pose $x^0 = x$, $x^j = H(x^{j-1}, z_j)$ pour $j = 1, \dots, l$. Pour tout $\gamma > 0$, il existe $\beta > 0$, tel que $\|X^0 - x^0\| < \beta$ et $\omega^k \in]z_k - \beta, z_k + \beta[$ pour $k = 1, \dots, l$, impliquent $\|X^k - x^k\| < \gamma$ pour $k = 1, \dots, l$ et on a donc

$$\|X^0 - x^0\| < \beta \Rightarrow \mathbb{P}_{X^0}(\|X^k - x^k\| < \gamma, k = 1, \dots, l) \geq (2\beta)^l.$$

Preuve. — Soit $\gamma > 0$. La continuité de H sur $(A \times [0, 1]) \cap E$ entraîne que l'on peut choisir $\beta > 0$ tel que $\|X^0 - x^0\| < \beta$ et $|\omega^1 - z| < \beta \Rightarrow \|X^1 - x^1\| < \gamma$, c'est-à-dire que

$$\|X^0 - x^0\| < \beta \Rightarrow \mathbb{P}_{X^0}(\|X^1 - x^1\| < \gamma) > 2\beta.$$

Par récurrence, en appliquant l fois la propriété de Markov, on obtient le résultat. \square

Ce lemme étant démontré, on construit les intervalles I_1, \dots, I_n en se bornant au cas $n \not\equiv 2$ modulo 3, la démonstration étant similaire pour $n \equiv 2$ modulo 3.

On sait que $x = (x_1, \dots, x_n) \in A$.

On définit par récurrence

$$x^0 = x, \quad x^j = H(x^{j-1}, x_j^{j-1}) \quad \text{et} \quad z_j = x_j^{j-1},$$

pour $j = 1, \dots, n$, de manière à ce que $i(x^{j-1}, z_j) = j$.

Alors d'après le lemme 3.1, on sait que pour tout $\gamma > 0$, il existe $\beta > 0$ tel que

$$\|X^0 - x^0\| < \gamma \Rightarrow \mathbb{P}_{X^0}(\|X^j - x^j\| < \gamma, j = 1, \dots, n) \geq (2\beta)^n.$$

Il est alors clair que si γ est suffisamment petit, les intervalles $I_j =]z_j - \beta, z_j + \beta[$ conviennent. D'où le théorème 3.1, avec $a = 2\gamma$ ne dépendant que de ε, n, x , et $\delta = \frac{1}{|\text{Det } D_n|}$. \square

Pour compléter ce premier résultat, nous allons prouver le

THÉORÈME 3.2. — $(\forall \varepsilon \in]0, 1[)$, il existe $x^\varepsilon \in A$, a_0 réel > 0 , tels que pour tout voisinage V_{x^ε} de x^ε de largeur $a < a_0$, il existe $T(n, \varepsilon, a)$ et $\gamma(n, \varepsilon, a)$ tels que :

$$(\forall x \in A) \quad \mathbb{P}_x(\tau_{V_{x^\varepsilon}} = T(n, \varepsilon, a)) \geq \gamma(n, \varepsilon, a) > 0.$$

Nous commençons par démontrer deux lemmes.

LEMME 3.2. — Pour tout $a > 0$, il existe un temps $T_0(n, \varepsilon, a)$ et une probabilité $\alpha(n, \varepsilon, a)$ tels que si $E_1 = \prod_{i=1}^n [1 - a, 1]$

$$(\forall x \in A) \quad \mathbb{P}_x \{ \tau_{E_1} \leq T_0(n, \varepsilon, a) \} \geq \alpha(n, \varepsilon, a) > 0.$$

Preuve. — Soit $z_1 = 1, z_2 = 1 - \frac{a}{n}, z_3 = 1 - \frac{2a}{n}, \dots, z_{n-1} = 1 - \frac{(n-2)a}{n}$.

Soit $x \in A$, soit T tel que : $(1 - \varepsilon)^T \leq \frac{a}{4n} < (1 - \varepsilon)^{T-1}$.

On prend :

$$\begin{aligned} \omega^1, \dots, \omega^{T_1(x)} & \text{ dans } [z_1, z_1 + \gamma] \\ \omega^{T_1(x)}, \dots, \omega^{T_1(x)+T_2(x)} & \text{ dans } [z_2, z_2 + \gamma] \\ \vdots & \\ \omega^{T_1(x)+\dots+T_{n-2}(x)}, \dots, \omega^{T_1(x)+T_2(x)+\dots+T_{n-1}(x)} & \text{ dans } [z_{n-1}, z_{n-1} + \gamma] \end{aligned}$$

où
$$\gamma = \frac{a}{2n} \left[\frac{(1 - \varepsilon)^T}{1 - (1 - \varepsilon)^T} \right]$$

et $T_1(x)$ est le premier instant t où partant de x , X_n^t et $X_{n-1}^t > 1 - \frac{a}{4n}$

.....

$T_1(x) + \dots + T_i(x)$ est le premier instant t où partant de x ,

$$X_{n-i}^t > 1 - \frac{i-1}{n} - \frac{a}{4n}.$$

Avec un tel choix des ω^t on a :

$$t \in [0, T_1(x)], i(X^{t-1}, \omega^t) = n \quad (\text{on a attiré } n \text{ et } n-1)$$

$$t \in [T_1(x) + \dots + T_{k-1}(x), T_1(x) + \dots + T_k(x)], i(X^{t-1}, \omega^t) = n - k$$

et $T_k(x) \leq T$ pour $k = 1, \dots, n - 1$.

Par conséquent : $P_x(\tau_{E_1} \leq (n-1)T) \geq \gamma^{(n-1)T}$ d'où $T_0(n, \varepsilon, a) = (n-1)T$. \square

LEMME 3.3 (Lemme d'attente). — Soit $x \in A$. Supposons $X_1^0 \in [x_1, x_1 + a[$, $a > 0$ et $t_0 \in \mathbb{N}^*$. Il existe alors β et γ , $\beta > \gamma > 0$ tels que $(\forall t \leq t_0)$, $\omega^t \in [x_1 - \beta, x_1 - \gamma]$.

On a

i)
$$i(X^t, \omega^{t+1}) = 1, \quad \forall t \leq t_0$$

ii)
$$X_1^t \in [x_1 - a, x_1 + a], \quad \forall t \leq t_0.$$

Preuve. — On a $X_1^{t+1} = (1 - \varepsilon)X_1^t + \varepsilon\omega^{t+1} \geq (1 - \varepsilon)X_1^t + \varepsilon(x_1 - \beta)$, par conséquent $X_1^{t+1} - (x_1 - \gamma) \geq (1 - \varepsilon)(X_1^t - (x_1 - \gamma)) - \varepsilon(\beta - \gamma)$, en posant $u_t = X_1^t - (x_1 - \gamma)$ on obtient

$$u_t \geq (1 - \varepsilon)^t [u_0 + \beta - \gamma] - (\beta - \gamma) \quad \text{et} \quad u_0 \geq \gamma.$$

Il suffit donc que $\frac{\beta}{\gamma} \geq 1 + (1 - \varepsilon)^{t_0}$ pour que $u_{t_0} \geq 0$ et pour assurer

$X_1^t \geq x_1 - a$. On choisit donc $\beta \leq a$ et $\gamma \leq (1 + (1 - \varepsilon)^{-t_0})^{-1}\beta$. \square

Ce lemme permet de remplacer l'inégalité du Lemme 3.2 par une égalité puisque l'on peut prendre patience jusqu'au temps $T(n, \varepsilon, a)$ en tirant le point n° 1, mais tout en restant dans un voisinage qu'on s'est fixé à l'avance.

Une fois ces deux lemmes acquis, la démonstration comprend deux étapes.

Démonstration (Théorème 3.2). — 1) La première étape consiste à placer X_1^t et X_2^t , X_3^t et X_4^t , X_5^t , ..., X_n^t respectivement dans des voisinages de points $y_1, y_4, y_5, \dots, y_n$ ($y_1 < y_4 < \dots < y_n$) fixés, de demi-largeur a , en un temps fixé fini, avec une probabilité minorée. On procède de la manière suivante :

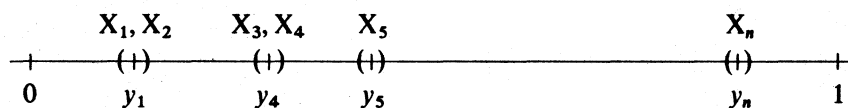
On applique le lemme 3.2 pour attirer tous les points « en 1 ».

On applique ce même lemme (ou son frère jumeau) pour attirer tous les points en y_n , puis on recommence en prenant pour point d'attraction y_{n-1} en évitant cependant d'attirer X_n^t hors du voisinage de y_n . Il suffit de s'arrêter lorsqu'on a fait rentrer X_{n-1}^t dans le voisinage de y_{n-1} et ainsi de suite, en laissant à chaque fois X_i^t au voisinage de y_i , jusqu'à ce que seuls se trouvent au voisinage de y_4 les points $X_1^t, X_2^t, X_3^t, X_4^t$. On recommence encore, mais cette fois en laissant sur place X_3^t, X_4^t , ceci en s'arrêtant dès que X_2^t (entraîné par X_1^t) est entré dans le voisinage de y_1 .

Le temps nécessaire pour effectuer ces opérations peut être rendu indé-

pendant du point initial x , en appliquant à chaque fois le lemme 3.3. Notons-le $T_1(n, \varepsilon, a)$.

Nous avons donc la situation suivante, avec une probabilité minorée indépendamment du point initial x :



2) La deuxième étape beaucoup plus courte, consiste à séparer X_1 de X_2 et X_3 de X_4 . Nous commençons par décrire une procédure déterministe : Choisissons z_1 et z_2 tels que :

$$y_1 < z_1 < \frac{(1 - \varepsilon)y_1 + \varepsilon z_1 + y_4}{2} < z_2 < y_4.$$

Si $X_1^0 = X_2^0 = y_1$ et $X_3^0 = X_4^0 = y_4$, alors si $(\omega^1, \omega^2) = (z^1, z^2)$, on a à l'instant 2 (seuls les quatre premiers points bougent) :

$$\begin{aligned}
 X_2^2 - X_1^2 &\geq \varepsilon(z_2 - \varepsilon z_1), \quad \text{idem pour } X_3^2, X_4^2, \\
 X_3^2 - X_2^2 &\geq (1 - \varepsilon)^2(y_4 - y_1)
 \end{aligned}$$

et pour a assez petit, on applique le lemme 3.1, pour trouver un intervalle où tirer ω^1 et ω^2 tout en préservant les positions obtenues à a près.

Ces deux étapes conduisent donc au théorème 3.2 avec :

$$\begin{aligned}
 T(n, \varepsilon, a) &= T_1(n, \varepsilon, a) + 2, \\
 \left\{ \begin{array}{l}
 x_1^\varepsilon &= (1 - \varepsilon)y_1 + \varepsilon z_1 \\
 x_2^\varepsilon &= (1 - \varepsilon)^2 y_1 + \varepsilon(1 - \varepsilon)z_1 + \varepsilon z_2 \\
 x_3^\varepsilon &= (1 - \varepsilon)^2 y_4 + \varepsilon(1 - \varepsilon)z_1 + \varepsilon z_2 \\
 x_4^\varepsilon &= (1 - \varepsilon)y_4 + \varepsilon z_2 \\
 x_j &= y_j \quad \text{pour } j \geq 5.
 \end{array} \right.
 \end{aligned}$$

Le point x^ε appartient à A et la probabilité de rentrer dans un voisinage de x^ε , de demi-largeur a est minorée indépendamment du point initial x . On a alors le

THÉORÈME 3.3. — *La chaîne (X^i) restreinte à A (resp. A') est Doeblin pour la mesure de Lebesgue.*

Preuve. — On a

$$\begin{aligned} \mathbb{P}^{\mathbb{T}(n,\varepsilon,a)+n}(\xi, d\zeta) &\geq \int_{V_{x^\varepsilon}} \mathbb{P}^{\mathbb{T}(n,\varepsilon,a)}(\xi, du) \mathbb{P}^n(u, d\zeta) \\ &\geq \gamma(n, \varepsilon, a) \frac{1}{|\det D_n|} 1_{C_{x^\varepsilon}}(\zeta) d\zeta. \quad \square \end{aligned}$$

D'où le théorème 1.2.

4. QUELQUES PROPRIÉTÉS COMPLÉMENTAIRES

Notons ${}^\varepsilon m = ({}^\varepsilon m_1, \dots, {}^\varepsilon m_n)$ l'espérance de la loi stationnaire ${}^\varepsilon \pi(\cdot)$.

a) Équations sur les moments de ${}^\varepsilon \pi(\cdot)$.

Soit $x \in A$, $x = (x_1, \dots, x_n)$.

En prenant les espérances pour la loi ${}^\varepsilon \pi$, on a

$$E(X_{i+2} + X_{i+1} - X_{i-1} - X_{i-2})(X_{i+2} + X_{i+1} + X_{i-1} + X_{i-2} - 4X_i) = 0$$

pour $i \neq 1, 2, n-1, n$.

$$E(X_{j+2} + X_{j+1})(X_{j+2} + X_{j+1} - 4X_j) = 0 \quad \text{pour } j = 1, 2$$

$$E(2 - X_{j-1} - X_{j-2})(2 + X_{j-2} + X_{j-1} - 4X_j) = 0 \quad \text{pour } j = n-1, n.$$

Ces équations portent sur les moments d'ordre 2 de la loi ${}^\varepsilon \pi(\cdot)$ et ne permettent pas de déterminer l'espérance ${}^\varepsilon m$.

Remarque. — Si X^0 suit la loi stationnaire ${}^\varepsilon \pi$, on a

$${}^\varepsilon m = EX^1 = E(EX^1/X^0) = \int d{}^\varepsilon \pi(x) \int y d\mathbb{P}_x(y) = \int d{}^\varepsilon \pi(x) h(x).$$

Si $h(x)$ était linéaire en x , $h(x) = Hx$, on obtiendrait l'espérance ${}^\varepsilon m$ par la résolution d'un système linéaire. Ce n'est pas le cas ici.

b) Convergence de l'algorithme lorsque ε dépend du temps.

On impose $X^0 \in A$ et ε dépendant du temps avec les conditions classiques des algorithmes de Robbins-Monro :

$$\varepsilon_t \geq 0; \quad \sum_0^\infty \varepsilon_t = +\infty; \quad \sum_0^\infty \varepsilon_t^2 < +\infty.$$

(Voir par exemple Métivier [5]).

On écrit

$$X^{t+1} = X^t + \varepsilon_t V(X^t, \omega^{t+1})$$

La fonction $H(x)$ s'écrit donc

$$H(x) = C(x)M(x - x^*) = B(x)(x - x^*).$$

D'après le théorème de Gerschgorin (cf. [6]), toutes les valeurs propres de la matrice $B(x)$ ont leur partie réelle strictement négatives puisque

$\sum_{j=1}^n |b_{ij}(x)| + b_{ii}(x) \leq 0$ quel que soit i et que ces valeurs propres sont différentes de 0 (Kohonen fait cette remarque pour la matrice M dans [3]).

THÉORÈME 4.1. — Supposons $\varepsilon_i \geq 0$, $\sum_0^{\infty} \varepsilon_i = +\infty$, $\sum_0^{\infty} \varepsilon_i^2 < +\infty$. Alors

si $X^0 \in A$, le processus X^t converge p. s. et dans L^2 vers x^* unique solution de $Mx = a$. (Un résultat analogue s'obtient si $X^0 \in A'$.)

Démonstration. — Le système différentiel déterministe

$$\frac{dx}{dt} = H(x)$$

admet la solution $x = x^*$ comme solution uniformément exponentiellement stable au sens large, car toutes les valeurs propres de $B(x)$ ont leur partie réelle strictement négative.

On remarque ensuite que les dérivées premières et secondes de $H(x)$ sont bornées sur A et on conclut à la convergence p. s. et dans L^2 du processus $X(t)$ vers x^* , lorsque $t \rightarrow +\infty$ (cf. Nevelson et Khaminski [7], § IV.4.3). \square

CONCLUSION

Bien que l'algorithme étudié soit simple à définir (et à simuler) quelle que soit la dimension, en démontrant la convergence en dimension 1, a nécessité beaucoup de calculs très spécifiques, et longs à expliciter (§ 2 et 3).

Il nous semble qu'il resterait trois points à vérifier :

— que lorsque ε tend vers 0 comme $\frac{1}{\log t}$, le processus réordonne les points (x_1, \dots, x_n) . Les minoration de probabilités et les majorations de temps d'attente du § 2 permettent d'avancer ce résultat ;

— que les mesures stationnaires (sur A et A') sont absolument continues par rapport à la mesure de Lebesgue ;

— que le point x^* (cf. § 4. b) est l'espérance de la mesure stationnaire quel que soit ε .

REMERCIEMENTS

Nous remercions chaleureusement J. Bretagnolle pour l'aide qu'il nous a apportée.

RÉFÉRENCES

- [1] J. L. DOOB, *Stochastic Processes*, Wiley, 1956.
- [2] T. KOHONEN, Self-organized formation of topologically correct feature maps, *Biol. Cybern.*, t. 43, 1982, p. 59 à 69.
- [3] T. KOHONEN, Analysis of a simple self-organizing process, *Biol. Cybern.*, t. 44, 1982, p. 135-140.
- [4] T. KOHONEN, *Self-organization and associative memory*, Springer-Verlag, 1984.
- [5] M. MÉTIVIER, *Semi-martingales*, de Gruyter, 1982.
- [6] A. CODDINGTON et N. LEVINSON, *Theory of Ordinary Differential Equations*, Mc Graw-Hill, 1955.
- [7] M. B. NEVEL'SON et R. Z. HAS'MINSKII, Stochastic approximation and recursive estimation. *A. M. S. Translations*, t. 47, 1973.
- [8] G. W. STEWART, *Introduction to matrix computations*. Academic Press, 1973.

(Manuscrit reçu le 15 janvier 1985)

A Stochastic Model of Retinotopy: A Self Organizing Process

M. Cottrell¹ and J. C. Fort²

¹ Université Paris-Sud, UA 743 C.N.R.S., Statistique Appliquée, Bâtiment 425 Mathématique, F-91405 Orsay, France

² Université Paris V, Laboratoire de Statistiques Médicales, 45, Rue des Saints Pères, F-75006 Paris, France

Abstract. Following Kohonen and using the Hebb principle, we define a self organizing stochastic process, which is a simple modelization of the retinotopy, i.e. the establishment of well-ordered connexions between the retina and the cortex.

We give some mathematical results about convergence of this process. These results are illustrated by computer simulations.

Introduction

Many papers have been published about the modelization of the retinotopy; i.e. the self-building of a bijective and bicontinuous mapping from the retina to some part of the brain cortex. Some are oriented towards the analysis of theoretical models proposed on biological grounds (Willshaw and Malsburg 1979; Overton and Arbib 1982). Others are further from biological reality but try to highlight some simple mechanisms of self-organization (Kohonen 1982, Erdi and Barna 1984). This paper belongs to the latter category.

Here we construct a stochastic algorithm by combining two points of view: we use the Hebb principle (or one of its possible expressions) together with an algorithm similar to that proposed by Kohonen (1982).

In Sect. I, we present the mathematical formalization of the model.

In Sect. II, we study this model in a simplified case: the one-dimensional case.

In Sect. III, we indicate which results are preserved from the one-dimensional to the two-dimensional case.

The last section is devoted to the description of the behavior of the model when the adaptation parameter decreases with time.

I The Model

Notations. The retina R is modeled by a two-dimensional network: $\{0, 1, \dots, n+1\}^2$ (resp. in Sect. II an integer interval $\{0, 1, \dots, N+1\}$). Its cells are denoted by (i, j) (resp. i).

The relevant part of the cortex C is represented by the square $[0, 1]^2$ (resp. the interval $[0, 1]$). Its cells are denoted by x .

Each cell (i, j) is mainly related to a cortex cell x_{ij} . However, a density of synaptic weights is present around the cell x_{ij} . We denote this density by $f_{ij}(x)$. That means that there are many connexions from (i, j) to the cells in C , but we essentially consider the center of gravity of these connexions with respect to the synaptic weights. We suppose that $f_{ij}(x) = f(\|x - x_{ij}\|)$ with f decreasing, positive and $f(0) = 1$.

If we stimulate the retina, by illuminating the cells of a part σ of R the cells $(x_{ij}, (i, j) \in \sigma)$ are excited and the total postsynaptic activity in a point of C is:

$$\varphi^\sigma(x) = \sum_{(i, j) \in \sigma} f_{ij}(x).$$

We enounce the Hebb principle in the form: "The variation of f_{ij} is proportional to the product of the post- and pre-synaptic activities", which is resumed in:

for every $(i, j) \notin \sigma$ f_{ij} does not change
for every $(i, j) \in \sigma$ f_{ij} is changed to \bar{f}_{ij} , with

$$\bar{f}_{ij}(x) = \frac{f_{ij}(x) + \varepsilon \varphi^\sigma(x) f_{ij}(x_{ij})}{\int \{f_{ij}(x) + \varepsilon \varphi^\sigma(x) f_{ij}(x_{ij})\} dx}, \quad (\varepsilon \text{ small, } > 0)$$

for which the center of gravity is

$$\begin{aligned} \bar{x}_{ij} &= \frac{x_{ij} + \varepsilon \int x \varphi^\sigma(x) dx}{1 + \varepsilon \int \varphi^\sigma(x) dx} \\ &= \frac{1}{1 + \varepsilon |\sigma|} x_{ij} + \frac{\varepsilon}{1 + \varepsilon |\sigma|} \sum_{(k, l) \in \sigma} x_{kl}, \end{aligned}$$

where $|\sigma|$ is the cardinal of σ .

Thus \bar{x}_{ij} is a weighted average of x_{ij} and the average of $\{x_{k,l}, (k,l) \in \sigma\}$. We naturally choose \bar{x}_{ij} as the new cell in C mainly connected with the retina cell (i,j) , after the effect of the stimulation σ .

We note that the global post-synaptic activity of the image in C of the stimulation σ has been increased:

$$\sum_{(i,j),(k,l) \in \sigma} \bar{f}_{ij}(\bar{x}_{kl}) \geq \sum_{(i,j),(k,l) \in \sigma} f_{ij}(x_{kl}).$$

However, if the same stimulation σ is iterated successively many times, the points $\{(x_{ij}), (i,j) \in \sigma\}$ become closer and closer. To avoid this effect, we must force the points x_{ij} to occupy all the free place in $[0,1]^2$. A suitable assumption having this effect is to suppose that the cells of the boundary of R are connected to the corresponding cells of the boundary of C : that means that the boundary relations between R and C are well defined before the process starts, and will not change with time.

We now set this hypothesis in a mathematical algorithm:

Let $X^t = (X^t_{ij})_{(i,j) \in \{0, \dots, n+1\}^2}$ [resp. $(X^t_i)_{i \in \{0, \dots, n+1\}}$ in the simplified case] denote the positions of the main connexions of the cells (i,j) . It is a mapping from R to C .

The initial mapping X^0 , at time $t=0$, may be random, or deterministic.

The stimulations σ are randomly chosen: a stimulation σ is defined by its center (i,j) and the closest neighbours of (i,j) . We study the simplest case: $\sigma(i,j) = \{(i-1,j), (i,j-1), (i,j), (i+1,j), (i,j+1)\}$ [resp. $\sigma(i) = \{i-1, i, i+1\}$].

One can take larger stimulations without modifying the main mathematical results.

Thus stimulations are defined by their centers and we assume that for all $t \geq 1$ they are independent random variables with probability distribution μ given by $(p_{ij})_{(i,j) \in \{1, \dots, n\}^2}$, $p_{ij} > 0$ (resp. $p_i, p_i > 0$). In the numerical applications we have chosen $p_{ij} = \frac{1}{n^2}$ (resp.

$p_i = \frac{1}{n}$). Note that (i,j) can't be chosen on the boundary of R .

If at time t the application from R to C is X^t , at time $t+1$ one picks at random a stimulation with center (i_0, j_0) , denoted by $\sigma(i_0, j_0)$ giving

$$X^{t+1}_{ij} = X^t_{ij} \quad \text{if } (i,j) \notin \sigma(i_0, j_0)$$

$$X^{t+1}_{ij} = (1 - |\sigma(i_0, j_0)| \cdot \lambda) X^t_{ij} + \lambda \left(\sum_{(i',j') \in \sigma(i_0, j_0)} X^t_{i'j'} \right)$$

resp. $i \neq (i_0, j_0)$

$$X^{t+1}_i = X^t_i \quad \text{if } i \notin \sigma(i_0)$$

$$X^{t+1}_i = (1 - |\sigma(i_0)| \cdot \lambda) X^t_i + \lambda \left(\sum_{i \in \sigma(i_0)} X^t_i \right).$$

if $i \in \sigma(i_0)$

Here $|\sigma(i_0, j_0)| = 5$ [resp. $|\sigma(i_0)| = 3$] and we must have $\lambda < \frac{1}{5}$ (resp. $\frac{1}{3}$); λ plays the role of $\frac{\epsilon}{1 + \epsilon|\sigma|}$.

The boundary conditions are given by: for every t ,

$$X^{t+1}_{0,j} = \left(\frac{j}{n+1}, 0 \right) \quad X^{t+1}_{i,0} = \left(0, \frac{i}{n+1} \right) \quad i, j \in \{0, \dots, n+1\}$$

$$X^{t+1}_{n+1,j} = \left(\frac{j}{n+1}, 1 \right) \quad X^{t+1}_{i,n+1} = \left(1, \frac{i}{n+1} \right)$$

(resp. $X^t_{0,j} = 0, X^t_{i,n+1} = 1$).

Disregarding the fixed boundary values the above mechanism is a Markov chain taking its values in $([0,1]^2)^n$ (resp. $[0,1]^n$).

Our goal in Sects. II and III is to prove the convergence of this Markov chain to a stationary distribution which establishes the retinotopy. We begin with the one-dimensional case, the simplest. There the retinotopy is equivalent to the convergence of X^t toward an increasing mapping from $\{0, 1, \dots, n+1\}$ to $[0, 1]$.

II The One-Dimensional Case

The algorithm previously defined in Sect. I, is a Markov chain on $[0, 1]^n$.

1 Matrix Notation

We denote by I_p the identity matrix. To choose an integer $i_0 \in \{1, \dots, n\}$ according to the distribution μ is equivalent to choosing a pair (A_{i_0}, B_{i_0}) , where A_{i_0} is a $n \times n$ -matrix in $M_n[0, 1]$, i.e. with elements in $[0, 1]$ and B_{i_0} a vector in $[0, 1]^n$, with

$$A_1 = \left(\begin{array}{cc|c} 1-2\lambda & \lambda & 0 \\ \lambda & 1-2\lambda & 0 \\ \hline 0 & 0 & I_{n-2} \end{array} \right), \quad B_1 = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

$$A_i = \left(\begin{array}{c|cc|c} I_{i-2} & 0 & 0 & \\ \hline 0 & 1-2\lambda & \lambda & \lambda \\ & \lambda & 1-2\lambda & \lambda \\ \hline 0 & \lambda & \lambda & 1-2\lambda \\ \hline 0 & 0 & 0 & I_{n-i-1} \end{array} \right), \quad B_i = \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{pmatrix},$$

for $1 < i < n$ and

$$A_n = \left(\begin{array}{c|cc} I_{n-2} & 0 & \\ \hline 0 & 1-2\lambda & \lambda \\ & \lambda & 1-2\lambda \end{array} \right), \quad B_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \lambda \\ \lambda \end{pmatrix}.$$

Then we can define a sequence of independent random variables (Z^t, W^t) , with range $\{(A_1, B_1), \dots, (A_n, B_n)\}$, and distribution $\mu = (p_1, \dots, p_n)$ and the process X^t can be described by

$$\begin{cases} X^{t+1} = Z^{t+1}X^t + W^{t+1} \\ X^0. \end{cases}$$

In the following, the matrices and the corresponding linear mappings, in canonical base (e_1, \dots, e_n) in \mathbf{R}^n , are identified.

2 Order Stability

Proposition 1. Let \mathcal{A} be the set of non-decreasing functions from R to C . Then \mathcal{A} is an absorbing set for the process X^t .

Proof. $\mathcal{A} = \{x \in [0, 1]^n \mid 0 \leq x_1 \leq \dots \leq x_n \leq 1\}$.

For $i = 1, \dots, n$ and $x \in \mathcal{A}$, it is easily verified that $A_i x \in \mathcal{A}$. Furthermore, $B_i \in \mathcal{A}$ for every i . \square

3 Weak Convergence of the Process X^t

For $X^0 = x$, we may write

$$X^t = Z^t Z^{t-1} \dots Z^1 x + \sum_{j=1}^{t-1} Z^t \dots Z^{j+1} W^j + W^t,$$

which has the same distribution as

$$Y^t = Z^1 Z^2 \dots Z^t x + \sum_{j=1}^{t-1} Z^1 \dots Z^j W^{j+1} + W^1,$$

since the variables (Z^t, W^t) are iid.

We study the behavior of products $Z^1 Z^2 \dots Z^t$ when $t \rightarrow +\infty$.

We first show:

Proposition 2. Let \mathbf{R}^n be endowed with the norm $\|x\| = \sum_{i=1}^n |x_i|$, and $M_n(\mathbf{R})$ with the associated norm for linear functions $\left(\|A\| = \sup_x \frac{\|Ax\|}{\|x\|} \right)$. Then there exists $c \in [0, 1[$ such that $\|A_1 \dots A_n\| \leq c$.

Proof. We have $\|A_i\| = 1$ for $i = 2, \dots, n-1$, so we need some calculations to prove that $\|A_1 \dots A_n\| < 1$.

For $i \neq 1, n$ and $j > i+1$ or $j < i-1$, $A_j e_i = e_i$ and

$$A_i e_i = \lambda e_{i-1} + (1-2\lambda)e_i + \lambda e_{i+1}$$

$$A_{i-1} e_i = \lambda e_{i-2} + \lambda e_{i-1} + (1-2\lambda)e_i$$

$$A_{i+1} e_i = (1-2\lambda)e_i + \lambda e_{i+1} + \lambda e_{i+2}.$$

For $1 < i < n-1$,

$$A_2 \dots A_{i+1} e_i = \alpha_i e_1 + f_1 \quad \text{with } \alpha_i > 0$$

and f_1 a positive linear combination of (e_2, \dots, e_{i+2}) . It is easily shown that $\alpha_i \geq \lambda^{i-1} (1-2\lambda)$ and $\|f_1\| \leq 1 - \alpha_i$.

Thus

$$\|A_1 \dots A_n e_i\| = \|A_1 \dots A_{i+1} e_i\| \leq 1 - \lambda^i (1-2\lambda). \quad (*)$$

Moreover, $(*)$ is true also for $i = 1, 2, n-1$ and n . Thus for $x \in [0, 1]^n$, we have

$$\|A_1 \dots A_n x\| \leq c \|x\| \quad (c = 1 - \lambda^n (1-2\lambda) < 1). \quad \square$$

We deduce the

Theorem 1. The process X^t converges weakly to a random variable Y , independent of X^0 ; Y is supported by \mathcal{A} . Consequently the reaching time $\tau(n, \lambda)$ of \mathcal{A} is finite.

That means that the algorithm reorders the points x_1, \dots, x_n in a finite time.

Proof. We have:

$$P((Z^1, \dots, Z^n) = (A_1, \dots, A_n)) = p_1 \dots p_n = a > 0.$$

By the law of large numbers we get, with N_k defined in the second term,

$$\frac{1}{k} N_k = \frac{1}{k} \sum_{j=1}^k \mathbf{1}_{((Z^{(j-1)n+1}, \dots, Z^{jn}) = (A_1, \dots, A_n))}$$

$$\xrightarrow[k \rightarrow \infty]{a.s.} a$$

which implies that for almost all ω , and all $\varepsilon > 0$ sufficiently small, there exists $k_0(\omega)$ such that:

$$k \geq k_0(\omega) \Rightarrow \frac{1}{k} N_k(\omega) \geq a - \varepsilon.$$

Applying proposition 2, it follows that

$$\|Z^1 \dots Z^{nk}(\omega)\| \leq c^{\lfloor k(a-\varepsilon) \rfloor}$$

(where $\lfloor \cdot \rfloor$ is the integer part).

Thus putting $k_t = \left\lfloor \frac{t}{a-\varepsilon} \right\rfloor$, we have for almost all ω :

$$\exists T_0(\omega), T_0(\omega) = nk_0(\omega) \text{ such that:}$$

$$t \geq T_0(\omega) \Rightarrow \|Z^1 \dots Z^t\| \leq c^{\lfloor k_t(a-\varepsilon) \rfloor}$$

from which we deduce (following Kesten 1973):

$$\forall x \in [0, 1]^n$$

$$i) Z^1 \dots Z^t x \rightarrow 0 \text{ a.s.}$$

$$ii) \text{ Since } W^t \text{ is bounded, the serie}$$

$$W^t + Z^t W^{t-1} + \dots + Z^t \dots Z^{j+1} W_j \text{ converges a.s.}$$

Consequently Y^t converges a.s. toward a random variable Y independent of X^0 and X^t converges weakly toward Y .

Since \mathcal{A} is an absorbing class for X^t and by taking X^0 in \mathcal{A} , we conclude that Y is supported by \mathcal{A} . \square

4 X^t Does not Converge in Probability

As the distribution of Y is not concentrated on the diagonal of \mathbb{R}^n (which is quite obvious) then:

Proposition 3. *The process X^t does not converge in probability.*

Proof. There exist a point x and a neighborhood $V(x)$ of x such that:

$$\sup |x_{i+1} - x_i| = \alpha > 0 \quad \text{and} \quad P(Y \in V(x)) \geq \delta > 0$$

(we can choose $V(x)$ included in $\left\{ \|y - x\| < \frac{\alpha}{4} \right\}$).

Then: for all X^0 , there exists a time T such that

$$(\forall t \geq T) \quad P(X^t \in V(x)) \geq \frac{\delta}{2}$$

because X^t converges weakly to Y . We also have

$$P(X^t \in V(x) \cap \{Z^t = A_{i_0}, W^t = B_{i_0}\}) \geq \frac{\delta}{2u},$$

if we choose i_0 such that $x_{i_0+1} - x_{i_0} = \alpha$. It is then easily verified that

$$P(\|X^{t+1} - X^t\| \geq \varepsilon_0) \geq \frac{\delta}{2u} \quad \text{where} \quad \varepsilon_0 = \frac{\alpha\lambda}{2},$$

and we have proved that X^t cannot converge in probability. \square

5 Mean and Variance of the Limit Distribution

We denote by π the distribution of Y and m its mean value. If X^0 has distribution π then X^1 has too. Thus $m = (EZ)m + EW$, Z and W having the same distribution as (Z^1, W^1) .

We remark that this equation does not depend on λ .

For example: if $p_1 = \dots = p_n = \frac{1}{n}$, m is solution of (S):

$$(S) \begin{pmatrix} -4 & 2 & 1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 2 & -6 & 2 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 2 & -6 & 2 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 1 & 2 & -6 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} m_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ m_n \end{pmatrix} = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ -1 \\ -1 \end{pmatrix}$$

which gives:

- for $n=3$ $m = (\frac{2}{3}, \frac{1}{3}, \frac{1}{3})$
- $n=4$ $m = (\frac{77}{217}, \frac{91}{217}, \frac{126}{217}, \frac{140}{217})$
- $n=5$ $m = 0.31, 0.37, 0.5, 0.63, 0.69$
- $n=6$ $m = 0.28, 0.34, 0.45, 0.55, 0.66, 0.72$
- $n=7$ $m = 0.25, 0.30, 0.41, 0.5, 0.59, 0.70, 0.75$
- $n=8$ $m = 0.23, 0.28, 0.37, 0.46, 0.54, 0.63, 0.72, 0.77$
- $n=9$ $m = 0.21, 0.26, 0.34, 0.42, 0.5, 0.58, 0.66, 0.74, 0.79$
- $n=10$ $m = 0.20, 0.24, 0.32, 0.40, 0.46, 0.54, 0.60, 0.68, 0.76, 0.80$
- $n=20$ $m = 0.12, 0.14, 0.19, 0.23, 0.27, 0.31, 0.35, 0.39, 0.44, 0.48, 0.52, 0.56, 0.61, 0.65, 0.69, 0.73, 0.77, 0.81, 0.86, 0.88.$

In the same way we obtain an equation for the variance of the distribution π . Putting $A_{ij} = E(Y_i Y_j)$, the A_{ij} are solutions of the following linear system:

$$A_{ij} = \sum_{k,l} A_{kl} E(z_{ik} z_{lj}) + \sum_k m_k E(z_{ik} w_j) + \sum_l m_l E(w_i z_{lj}) + E(w_i w_j),$$

where z_{ij} is the (i,j) -element of Z , w_i is the i -element of W .

6 Simulations

We have made some computer simulations in order to study the reaching time of the set \mathcal{A} , i.e. the time before X^t becomes a non-decreasing function.

As could be guessed, $\tau_{n,\lambda}$ is almost exponentially distributed and we have for example:

Table 1. $\lambda=0.1$

n	10	20	30	40	50	60	70	80
Mean reaching time	63,7	336	866	1939	3284	5646	8210	13 508
Standard deviation	66,6	358	934	2067	3470	6004	8662	14 468

Table 2. $\lambda=0.3$

n	10	20	30	40	50	60	70	80
Mean reaching time	25.3	101	294	616	1195	1791	2730	2911
Standard deviation	27.7	109	316	650	1273	1890	4223	4432

III The Two-Dimensional Case

It is not easy to generalize the results of Sect. II. The process still converges in distribution, but the support of the limit distribution is not easy to describe any more, since there is no natural order in \mathbb{R}^2 . Up to now, we have not been able to show that this distribution is supported by a “well-ordered” set, but we know that its mean value is represented by a well-ordered network and every simulation we have done leads to a “well-ordered” network, in which the neighborhood relations in R are preserved in C . In Sect. IV we prove that if the adaptation parameter λ goes to zero as $t \rightarrow \infty$, X^t converges toward the mean value and retinotopy is achieved.

1 Asymptotic Behavior

Again we can write:

$$X^{t+1} = Z^{t+1}X^t + W^{t+1},$$

where the pair (Z^{t+1}, W^{t+1}) has distribution μ in $\{(A_{11}, B_{11}), \dots, (A_{nn}, B_{nn})\}$ $A_{ij} \in (M_n[0, 1])^2$, $B_{ij} \in ([0, 1]^n)^2$. We omit the details of this description, these being straight forward generalizations of the representation in Part II. The same argument as in Sect. II leads to:

Theorem 2. *The process X^t converges weakly as $t \rightarrow \infty$ toward Y , whose distribution is independent of X^0 . It does not converge in probability.*

Here we can't describe the support of Y . However, we can calculate its mean value.

2 Mean Value (For simplicity we put $p_{ij} = \frac{1}{n^2}$).

Let m be the mean value of Y . Thus m is the solution of the linear system:

$$m = (EZ)m + EW.$$

The coefficients of the system can be computed explicitly.

For $i \neq 1, n$ and $j \neq 1, n$, we have

$$m_{ij} = m_{ij} \left(\frac{5}{n^2} (1 - 4\lambda) + \left(1 - \frac{5}{n^2} \right) \right) + \frac{\lambda}{n^2} \Sigma,$$

where Σ is obtained by summing the m_{ki} (belonging to the neighbor set represented below), weighted with the coefficients 1 or 2 according to the following scheme:

$$\begin{array}{cccc}
 & & & j \\
 i \downarrow & & & \\
 & & .1 & \\
 & .2 & .2 & .2 \\
 1 \cdot & .2 & (i,j) & .2 \cdot 1 \\
 & .2 & .2 & .2 \\
 & & .1 &
 \end{array}$$

For other points we proceed in a similar way.

Thus we have a method to construct this linear system and to compute the value of (m_{ij}) . Again m does not depend on λ . For instance we obtain:

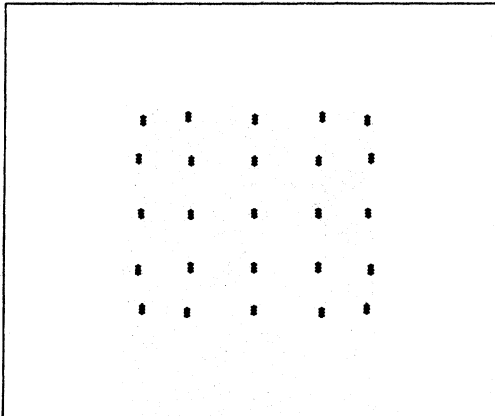


Fig. 1. For $n=5$

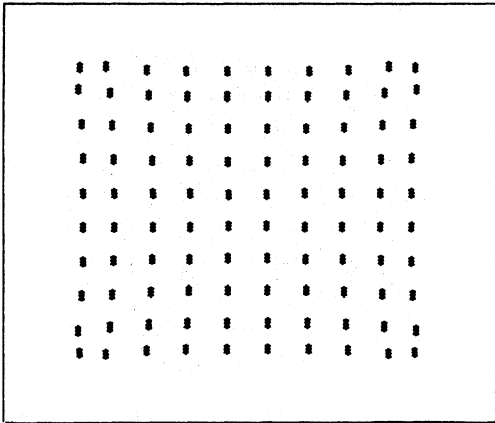


Fig. 2. For $n=10$

IV Adaptation Effect Decreasing with Time

We first prove a technical Lemma (in the case where

$$p_{ij} = \frac{1}{n^2} \left(\text{resp. } p_i = \frac{1}{n} \right).$$

Lemma. Let $V = \frac{1}{\lambda}(EZ - I)$. The matrix V has all its eigenvalues real and negative.

Proof. In the one-dimensional case V verifies:

- i) all the diagonal terms are negative.
- ii) the other terms are non negative.
- iii) all row sums are non positive and, in particular for $i = 1, 2, n-1$ and n , negative.

In the two-dimensional case, V is a $n^2 \times n^2$ matrix and we have

a) for $i \neq 1, n$ and $j \neq 1, n$, the line (i, j) of the system (S) is

$$-20m_{ij} + 2(\Sigma_1) + \Sigma_2$$

and Σ_1 contains $8m_{kr}$ -terms and Σ_2 $4m_{kr}$ -terms.

b) for $i = 1$ or n , or $j = 1$ or n , some of the terms in Σ_1 and Σ_2 are missing; thus V verifies i), ii), and iii).

Moreover, in the two cases V is symmetric. Consequently (see Gantmacher 1959) V admits only negative eigenvalues (this justifies "a posteriori" that the system S is an inversible system). \square

Now we can prove that when the parameter λ decreases with time, the process X^t converges to the mean value of the limit distributions of Sects. II and III. As this mean value is a "regular" network, the limit of X^t achieves the retinotopy: it preserves the topological order of the net (i, j) .

Theorem. Suppose that $\lambda = \lambda(t)$ is a function of time, such that $\sum_{t \in \mathbb{N}} \lambda(t) = +\infty$, $\sum_{t \in \mathbb{N}} \lambda(t)^2 < +\infty$. Then X^t converges in L^2 and a.s. to m , the mean value of the limit distribution of the process.

Proof. We have

$$X^{t+1} = Z^{t+1}X^t + W^{t+1}$$

and

$$E[X^{t+1}|X^t] = X^t + \lambda(t)VX^t + EW$$

(keeping the previous notations). Thus

$$E(X^{t+1} - m|X^t) = (X^t - m) + \lambda(t)V(X^t - m).$$

Applying a classical result for the Robbins-Monro algorithm, we get the convergence of X^t to m in L^2 and a.s. since 0 is the only solution of $Vy = 0$ and is asymptotically stable as a consequence of the Lemma. \square

V Conclusion

The model we have presented here is a self-organizing system based upon two principles: one is the Hebb principle which ensures the reinforcement of a connexion when it is used. The other is the assumption that the boundary is predetermined.

In Sect. IV we saw that if the process has a decreasing power of adaptation ($\lambda(t) \rightarrow 0$), it results in a topologically correct map from $\{0, 1, \dots, n+1\}^2$ on $[0, 1]^2$.

Even when the power of adaptation does not vary with time, in the one-dimensional case, such a map is attained in a finite time (more precisely, a distribution over such maps). In the two-dimensional case we could

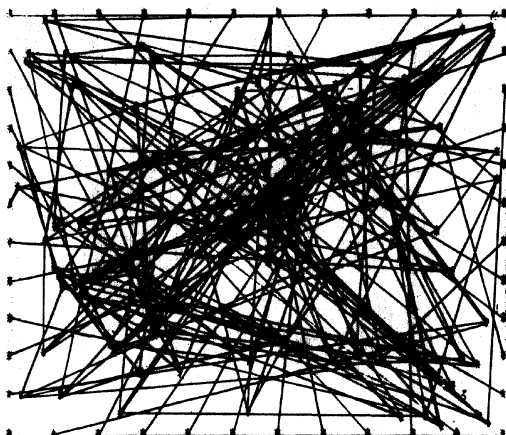


Fig. 3. Convergence to the mean value when $\lambda(t) = 1/(1+t/1000)$, with $n=10$. The graph is obtained by joining each point to its neighbours. Initial state

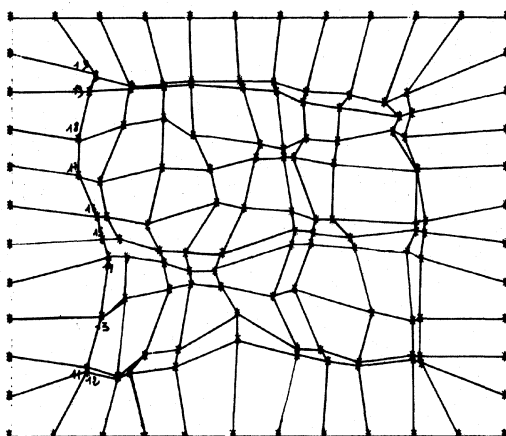


Fig. 4. After 3000 steps

not prove the same result essentially because we have not defined appropriately the notion of a topologically correct map. However, we have observed from simulations that, after a finite time, a "well-ordered" situation is attained which also seems stable. For future work, there remains to introduce a suitable notion of "well-ordering" and prove that the limit distribution is supported by topologically correct maps.

Acknowledgements. We warmly thank Prof. R. Azencott for his helpful comments and stimulating discussions.

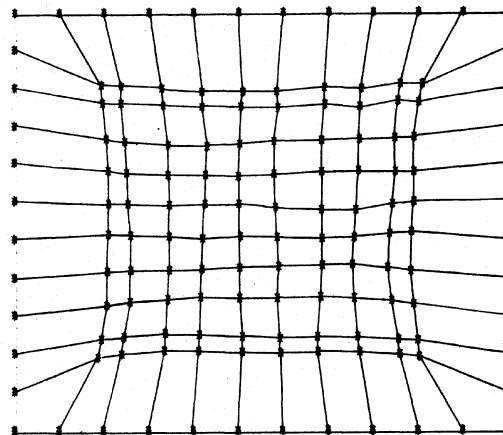


Fig. 5. After 20000 steps

References

- Erdi P, Barna G (1984) Self organizing mechanism for the formation of ordered neural mappings. *Biol Cybern* 51:93-101
- Gantmacher F.R. (1959) *The theory of matrices*, vol. 1 and 2. Chelsea Publishing Company, New York
- Kesten H (1973) Random difference equations and renewal theory for products of random matrices. *Acta Math* 131:207-248
- Kohonen T (1982) Self-organized formation of topologically correct feature maps. *Biol Cybern* 43:59-69
- Kohonen T (1984) *Self-organization and associative memory*. Springer, Berlin Heidelberg New York Tokyo
- Overton K, Arbib M (1982) The extended Branch-Arrow model of the formation of retinotectal connexions. *Biol Cybern* 45:157-175
- Willshaw DJ, Malsburg C von der (1979) A marker induction for the establishment of ordered neural mapping: its application to the retino-tectal problem. *Philos Trans R Soc London* 287:B1021

Received: November 8, 1985

Madame Marie Cottrell
 Université Paris-Sud
 UA 743 C.N.R.S.
 Statistique Appliquée
 Bâtiment 425 Mathématique
 F-91405 Orsay
 France

TROISIEME PARTIE

STABILITE ET ATTRACTIVITE DANS DES RESEAUX DE
MEMOIRE ASSOCIATIVE.

Stability and Attractivity in Associative Memory Networks

M. Cottrell

UA 743 CNRS Statistique Appliquée, Laboratoire de Mathématique Bât. 425, Université Paris XI, F-91405 Orsay Cedex, France

Abstract. We focus on stable and attractive states in a network having two-state neuron-like elements. We calculate the connection matrix which guarantees the stability and the strongest attractivity of p memorized patterns. We present an analytical evaluation of the patterns' attractivity. These results are illustrated by some computer simulations.

1 Introduction

The ability to recall memorized patterns is a very important human feature. Many models of neural networks include it, and the capacity of memory is usually spatially distributed throughout the network. It is exactly contained in the "efficiencies" of synaptic junctions.

The study of Distributed Associative Memory Networks was initiated in the fifties. We refer e.g. to Rosenblatt (1958); Caianiello (1961); Kohonen (1970, 1972, 1976); Nakano (1972); Kohonen et al. (1974, 1977) etc.

In these models, the state of each neuron is represented by output spike frequencies. The memorization of patterns relies upon changes in the synaptic efficiencies according to the presentation of patterns. The further presentation of a perturbed pattern (or a part of it) leads to its recollection. This is the property of selective recall.

This theory is largely developed in Kohonen (1984).

This kind of models can also be used for constructing some associative memory which need not be a realistic model of neural network. They can be conceived for memorization and retrieval of patterns in another context, e.g. error corrections in transmissions. In this case, the like-neuron automata may have continuous – or discrete-valued states. They must just work as required. The simplest networks with two-state threshold "neurons" were studied by many

authors: Little (1974), Hopfield (1982), Peretto (1984), Amit et al. (1985a, b), Weichbuch and Fogelman-Soulie (1985), Personnaz et al. (1985), etc.

The use of conceptual tools of statistical mechanics, especially spin glasses models, has allowed a good advance in understanding their behavior, and led to asymptotic results, when the number of units grows to infinity (Amit et al. 1985a, b).

In all these papers, the networks are used for recognizing, i.e. retrieving a given set of configurations referred to as patterns. However *these patterns, whether deterministically or stochastically chosen, are not always attractors, and not even stable states*, in the network defined by the classical connections suggested by Hebb (1949) and advocated by Cooper et al. (1978).

On the other hand, these connections are symmetric, and this is an unpleasant restriction, even if it allows to define one Hamiltonian whose local minima contain the patterns (Hopfield 1982; Peretto and Niez 1985).

All contributions have studied either a deterministic algorithm, with temperature $T=0$ (Hopfield 1982; Personnaz et al. 1985; Weichbuch and Fogelman-Soulie 1985) or a stochastic one with temperature $T>0$ (Amit et al. 1985a, b; Peretto 1984).

In this paper we shall study the deterministic algorithm. We consider a network consisting of N two-state units. We define a *configuration* or *network state* as an element $S=(S_1, S_2, \dots, S_N)$ of the hypercube $\{-1, +1\}^N$, with $S_i = +1$ (or -1) if the i -th unit is *active* (or *inactive*). The configuration at time t is denoted by S^t .

The collective behavior of such a network is entirely specified by the strengths of the connections C_{ij} between the source j and the receiver i , and by the threshold values. The $N \times N$ matrix $C=(C_{ij})$ acts as a decoding machine and will be called *connection matrix*.

Each unit receives inputs from all the others weighted by the strengths of connections. Like in

neural networks, the sum of the weighted inputs represents the "membrane potential of the neuron": it becomes (or remains) active if this potential is higher than the threshold, and it becomes (or remains) inactive if this potential is smaller than the threshold.

There are p learned, previously known configurations referred to as *the patterns*, denoted by S^1, \dots, S^p . We wish, starting from a configuration S , which contains some errors, to retrieve one of the S^m , $m=1, \dots, p$, which is to be the nearest one.

Thus it is necessary to determine how to choose the (C_{ij}) matrix in order to get the patterns as attractive as possible. Moreover we have to use an iteration mode, either the serial one, or the parallel one.

Dynamic Description

At time 0, the initial configuration in S^0 , issued from one of the patterns, but containing some errors (transmission errors, miscellaneous disturbances).

We call $S^t = (S_j^t)$ the configuration at time t .

At time t , the unit i receives the signal $\sum_j C_{ij} \left(\frac{S_j^t + 1}{2} \right)$, the variable $\frac{S_j^t + 1}{2}$ is equal to $+1$ if the unit j is active, 0 if not. If unit i receives a signal greater (resp. smaller) than the threshold, then it becomes active (resp. inactive). We shall choose, as usual, the threshold $\theta_i = \frac{1}{2} \sum_j C_{ij}$, for this choice implies that:

$$\sum_j C_{ij} \left(\frac{S_j^t + 1}{2} \right) > \theta_i \Leftrightarrow \sum_j C_{ij} S_j^t > 0.$$

Then the two types of algorithm are:

a) *Sequential Iteration Algorithm*. At time t , pick at random a unit $i \in \{1, \dots, N\}$, with uniform distribution, or in any case, in such a way that all the units can be selected. Then

$$\text{if } \sum_j C_{ij} S_j^t > 0, \quad S_i^{t+1} = +1$$

$$\text{if } \sum_j C_{ij} S_j^t < 0, \quad S_i^{t+1} = -1$$

$$\text{if } \sum_j C_{ij} S_j^t = 0, \quad S_i^{t+1} = S_i^t.$$

At each step, only one unit is checked.

b) *Parallel Iteration Algorithm*. At time t , calculate all the sums $\sum_j C_{ij} S_j^t$ and set $S_i^{t+1} = +1, -1$, or S_i^t as $\sum_j C_{ij} S_j^t > 0, < 0$, or $= 0$.

The system evolves by lining up the state S_t with the local field defined as $\frac{1}{2} \sum_j C_{ij}$.

For the two kinds of dynamics, two notions are interesting:

The stability. A configuration S^0 is stable iff $S^t = S^0$ for every t .

The attractivity. A configuration S^0 is a k -attractor (for $1 \leq k \leq N$) iff starting from a configuration S which presents k errors with respect to S^0 , the dynamic leads to S^0 .

Obviously the stability and the attractivity of a configuration S are defined for each kind of iteration mode and for each given matrix.

In this paper we attempt to solve the following problem: *Given patterns S^1, \dots, S^p , build a matrix C (i.e. an algorithm) in such a way that the S^1, \dots, S^p are stable and k -attractors with k as great as possible.*

The paper is organized as follows:

In Sect. 2, we give exact definitions of stability, attractivity, domain of attraction of a configuration.

Section 3 is devoted to the study of orthogonal patterns.

In Sect. 4, following Personnaz et al. (1985), we give the general formulation of the connection matrix C which provides stability to patterns S^1, \dots, S^p . Then in Sect. 5, we get the mathematical expression of patterns attractivity for a given matrix C .

Section 6 describes a construction of the matrix C , that maximizes all the patterns attractivity.

In Sect. 7, we come back to the case where all patterns are pairwise orthogonal, and stress the interest of that situation.

In Sect. 8, we solve the settled problem, by considering the situation where all the patterns have the same degree of attractivity.

In Sect. 9, we introduce the usual notion of configuration energy, which is only defined when the matrix C is symmetric, and we give the relations between energy minima and attractors.

Section 10 is devoted to a discussion.

In Appendix 1, some properties of the Hamming distance are recalled, and in Appendices 2 and 3, numerical examples illustrate our results and show that it is impossible to get a better estimate of the attractivity than the given one.

2 Stability-Attractivity

Let C be a connection matrix, and $S = (S_j)$ a configuration. We consider each of the two kinds of algorithms (sequential or parallel ones).

We denote by $\delta(S, S')$ the Hamming distance of two configurations S and S' , i.e. the number of components where S and S' differ, and by $d(S, S')$ the Euclidian distance of S and S' , viewed as elements of \mathbb{R}^N .

One has $d(S, S') = 2\sqrt{\delta(S, S')}$ (see Appendix 1 for properties of the Hamming distance).

Definition 2.1. A configuration S is *stable* (with respect to C) iff starting from S , the network state remains S , i.e. iff

$$\forall i \in \{1, \dots, N\}, \quad S_i \left(\sum_j C_{ij} S_j \right) \geq 0. \quad (2.1)$$

The set of stable configurations is denoted by E_0 .

Definition 2.2. Let k be an integer. A configuration S is a *k-attractor* (with respect to C) iff starting from S' with $\delta(S, S') = k$, the network state evolves in one step towards S'' with $\delta(S, S'') \leq k-1$ or leaves S' invariant (only in the sequential algorithm). This is equivalent to the condition:

$$\forall i, \forall j_1, \dots, j_k \text{ (with } j_1, \dots, j_k \text{ all different)}$$

$$S_i \left(\sum_j C_{ij} S_j - 2 \sum_{l=1}^k C_{ij_l} S_{j_l} \right) > 0. \quad (2.2)$$

The set of k -attractors is denoted by E_k .

Proof of the Above Equivalence. Let $S = (S_i)$ be a configuration which is a k -attractor and S' such that $\delta(S, S') = k$. Thus $S' = (S'_i)$ with

$$S'_i = -S_i \text{ for } i \in \{j_1, \dots, j_k\}$$

$$S'_i = S_i \text{ otherwise.}$$

We start from S' . For i integer picked at random (sequential algorithm) or for every i (parallel algorithm) we compute

$$\alpha_i = \sum_j C_{ij} S'_j = \sum_j C_{ij} S_j - 2 \sum_{l=1}^k C_{ij_l} S_{j_l}.$$

If $\alpha_i > 0$, (resp. $\alpha_i < 0$) we set the component i to be $+1$ (resp. -1).

The condition $\alpha_i S_i > 0$ means that we line up the spin i with the corresponding value of S , and therefore at the next time for the sequential algorithm, the new configuration will be S'' with $\delta(S, S'') = k-1$ if $i \in \{j_1, \dots, j_k\}$ or S' if not. For the parallel algorithm the new configuration will be S . \square

Remarks 2.3. 1) S and $-S$ satisfy the same inequalities.

2) S is stable (resp. k -attractor) for the sequential algorithm iff it is for the parallel algorithm. Indeed, for both algorithms, the inequalities to check for every integer i picked at random are the same. This will appear as a consequence of the following proposition which ensures that the Definitions 2.1 and 2.2 are coherent.

Proposition 2.4. For $k < \frac{N}{2}$, one has

$$E_k \subset E_{k-1} \subset \dots \subset E_1 \subset E_0.$$

Proof. Adding up the inequalities defining E_k , over all integers j_1, \dots, j_k , we obtain that $E_k \subset E_0$. Then adding

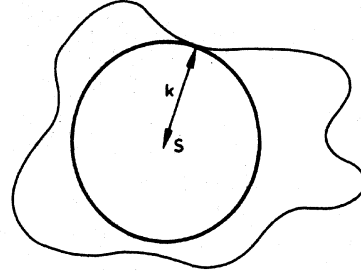


Fig. 1. Hamming radius

up over any subset of $(k-1)$ integers taken among $\{j_1, \dots, j_k\}$, we get $E_k \subset E_{k-1}$, under the condition that $N > 2k$. This last condition is obvious if we note that $\delta(S, -S) = N$, and that $S \in E_k$ iff $-S \in E_k$: indeed if $2k \geq N$, there would exist a state S' with $\delta(S', S) = k = \delta(S', -S)$, which is impossible since the spheres $B_\delta(S, k)$, $B_\delta(-S, k)$ are disjoint when $S \in E_k$ [sphere $B_\delta(S, k)$ is the set of configurations whose distance of S is less than k].

Moreover, note that the inequalities defining E_k are incompatible for $k \geq \frac{N}{2}$. \square

Let us now to define the domain of attraction of a configuration.

Definition 2.5. The *domain of attraction (DA)* of a configuration S is the (maybe empty) set of configurations S' such that, starting from S' , the algorithm leads to S .

However, it is convenient to consider only circular domains, i.e. spheres for the Hamming distance.

Definition 2.6. The *Hamming radius of the DA* of a configuration S is the Hamming radius of the greatest sphere included in it.

Consequently, if $S \in E_k$, the radius of its domain of attraction is $\geq k$.

Note that a Hamming sphere with center S and radius k , $B_\delta(S, k)$, is the intersection of the Euclidian sphere with center S and radius $2\sqrt{k}$, $B_\delta(S, 2\sqrt{k})$, and the hypercube $\{-1, +1\}^N$.

Remark 2.7. The maximal size of the DA of the patterns is necessarily bounded by the mutual Hamming distances of the patterns (and their opposite), since $S \in E_k$ and $S' \in E_k$ require $k+k < \delta(S, S')$ and $k+k < \delta(S, -S') = N - \delta(S, S')$.

3 Case of Orthogonal Patterns

Two configurations S and S' are orthogonal iff their Euclidian product $\langle S, S' \rangle = \sum_i S_i S'_i = 0$. (Thus orthogonality is defined as usual in the Euclidian space \mathbb{R}^N .)

The case of pairwise orthogonal patterns is well known, and easy to study (see Hopfield 1982; Peretto 1984; Personnaz et al. 1985) especially because of the following properties:

First, all the mutual Hamming distances between the patterns (and their opposite) are equal to $\frac{N}{2}$, so the patterns are well separated.

Furthermore, this case corresponds to the spin glasses one, in which case values S_i^m are picked at random, independently, with $\frac{1}{2} = \mathbf{P}(S_i^m = +1) = \mathbf{P}(S_i^m = -1)$ for all $i = 1, \dots, N$ and $m = 1, \dots, p$. The mean value of the Euclidian product $= \sum_i S_i^m, S_i^{m'}$ is then 0, (for $m \neq m'$) and the patterns are on average orthogonal, because the mean values of the number of components equal to +1 or to -1 are equal.

The usual connections (Hebbian connections) are then

$$C_{ij} = \frac{1}{N} \sum_{m=1}^p S_i^m S_j^m. \quad (3.1)$$

They are suggested by the principle of Hebb. In neural networks, following Hebb (1949), the synaptic efficacies are modified according to the neural activity, the strength of a synapse being proportional to the correlated activities of the neurons it connects. Here each coincidence of values for the patterns S^m in units i and j increases C_{ij} , and conversely.

With the above definition of matrix (C_{ij}) , in the orthogonal case, numerical simulations show strong attractivity of patterns, thus good efficiency of information retrieval. In fact, we prove that:

Proposition 3.2. *If the patterns S^1, \dots, S^p are pairwise orthogonal, these configurations are k -attractors at least up to $k = \left\lfloor \frac{N}{2p} \right\rfloor$ (integer part of $\frac{N}{2p}$), for the Hebbian connection matrix.*

Proof. According to (2.2) and (3.1), for $m = 1, \dots, p$, S^m is a k -attractor for C iff

$$\frac{1}{N} S_i^m \left(\sum_{j=1}^N \left(\sum_{l=1}^p S_l^j S_l^m \right) S_j^m - 2 \sum_{\theta=1}^k \left(\sum_{l=1}^p S_l^{\theta} S_l^m \right) S_j^{\theta} \right) > 0$$

for every $i = 1, \dots, N$, and distinct j_1, \dots, j_k in $\{1, \dots, N\}$. Or iff

$$\frac{1}{N} \left[\sum_{l=1}^p S_l^m S_l^m \left(\sum_{j=1}^N S_j^m S_j^m \right) - 2 \sum_{\theta=1}^k \sum_{l=1}^p S_l^{\theta} S_l^m S_j^{\theta} S_j^m \right] > 0.$$

The patterns S^1, \dots, S^p being pairwise orthogonal, $\sum_{j=1}^N S_j^l S_j^m = 0$ if $l \neq m$, and $= N$ if $l = m$. So the condition

becomes

$$\frac{1}{N} \left[N - 2 \sum_{\theta=1}^k \sum_{l=1}^p S_l^{\theta} S_l^m S_j^{\theta} S_j^m \right] > 0.$$

Now

$$\left| 2 \sum_{\theta=1}^k \sum_{l=1}^p S_l^{\theta} S_l^m S_j^{\theta} S_j^m \right| \leq 2 \sum_{\theta=1}^k \sum_{l=1}^p |S_l^{\theta}| |S_l^m| |S_j^{\theta}| |S_j^m| \leq 2kp,$$

since $|S_l^{\theta}| = |S_l^m| = |S_j^{\theta}| = |S_j^m| = 1$. The condition is satisfied if $N - 2kp > 0$.

Since $2 \left\lfloor \frac{N}{2p} \right\rfloor < \frac{N}{2}$ - Hamming distance of two orthogonal patterns - as soon as $p > 2$, we see that each pattern is k -attractor at least for $k = \left\lfloor \frac{N}{2p} \right\rfloor$. \square

4 Searching the Matrix C

We denote by A' the transposed of the matrix A , and we identify a vector of \mathbf{R}^N with the column matrix of its components in the canonical basis. We denote by $\langle \cdot, \cdot \rangle$ the usual Euclidian product, and $\| \cdot \|$ the associated norm. For a linear subspace E , we denote by E^\perp the subspace of all vectors which are orthogonal to every vector in E .

We may express the stability of the patterns S^1, \dots, S^p (see Definition 2.1) by

$$\sum_j C_{ij} S_j^m S_j^m = \langle C_i, D_i^m \rangle > 0 \quad (4.1)$$

for every $i = 1, \dots, N$ and every $m = 1, \dots, p$, where C_i is the i -th row of C , viewed as a vector of \mathbf{R}^N , and $D_i^m = S_i^m S^m$. The vector D_i^m is $\pm S^m$, so that its i -th component is equal to +1.

Thus, we want to determine N vectors C_1, \dots, C_N of \mathbf{R}^N , such that, for all i , the p vectors D_i^m are on the same side of hyperplane H_i , defined as the orthogonal space of C_i .

Such an hyperplane is not unique: more important is the volume left around each D_i^m , more performing will be the choice of C_i (and of H_i).

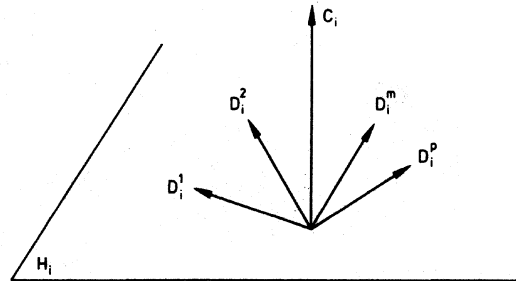


Fig. 2. Choice of hyperplane H_i

We see immediately that for $C_i = \frac{1}{p} \sum_{m=1}^p D_i^m$ (natural choice which corresponds to $C_{ij} = \frac{1}{p} \sum_{m=1}^p S_i^m S_j^m$, i.e. to Hebbian connections (3.1), apart from constant $\frac{N}{p}$), the S^m may not be stable, as can be observed by numerical simulations. (See also Personnaz et al. 1985.)

However this choice is often convenient, especially when the S^m are chosen at random, for then they are almost orthogonal (see Sect. 3).

In Personnaz et al. (1985), one can find the general expression for the matrix C ensuring stability of patterns S^1, \dots, S^p , under the condition that they are linearly independent.

Their formula is

$$C = (A_1 S^1, \dots, A_p S^p) (\Sigma' \Sigma)^{-1} \Sigma' + \tilde{C}, \quad (4.2)$$

where Σ is the $(N \times p)$ matrix with columns S^1, \dots, S^p , A_1, \dots, A_p are arbitrary positive diagonal N -matrices and \tilde{C} is a $(N \times N)$ -matrix such that $\tilde{C}\Sigma = 0$. Indeed, the system to be solved is $\langle C_i, S_i^m S^m \rangle > 0$ ($i=1, \dots, N$; $m=1, \dots, p$) or equivalently:

$$\langle C_i, S^m \rangle = \alpha_i^m S_i^m \text{ for arbitrariness } \alpha_i^m > 0. \quad (4.3)$$

A more condensed form is: $CS^m = A_m S^m$, for $m=1, \dots, p$, where A_m is an arbitrary positive diagonal N -matrix.

The general expression of C_i is

$$C_i = ((\alpha_i^1 S_i^1, \dots, \alpha_i^p S_i^p) (\Sigma' \Sigma)^{-1} \Sigma') + \tilde{C}_i, \quad (4.4)$$

where $\alpha_i^1, \dots, \alpha_i^p$ are arbitrary positive scalars and \tilde{C}_i is orthogonal to S^1, \dots, S^p . We notice that C_i is the sum of a linear combination of S^1, \dots, S^p , and of an orthogonal vector, and that it is defined up to a positive multiplicative constant.

We must determine how to choose the $N \times p$ constants α_i^m , and the vectors \tilde{C}_i to optimize the attractivity of S^1, \dots, S^p . So, we shall study the size of the domain of attraction of the S^m , as a function of arbitrary coefficients of matrix C .

5 Size of the Domains of Attractivity (DA) of S^1, \dots, S^p

If we denote by k_m the radius of the DA of S^m (Definition 2.6), we have $k_m + k_{m'} < \inf(\delta(S^m, S^{m'}), N - \delta(S^m, S^{m'}))$ for any pair m, m' , with $m \neq m'$ (see Remark 2.7).

Furthermore k_m is limited by the "free room" around S^m , i.e. the subset of \mathbb{R}^N which contains S^m and none of the other patterns.

Let us denote by T_{j_1, \dots, j_k} the sign modifier in position j_1, \dots, j_k of a configuration, i.e. the mapping defined by

$$T_{j_1, \dots, j_k}(S) = S^*$$

with

$$S_i^* = S_i \text{ for } i \notin \{j_1, \dots, j_k\}$$

$$S_i^* = -S_i \text{ for } i \in \{j_1, \dots, j_k\}.$$

Thus $S^m \in E_k$ (S^m is a k -attractor) iff

$$\langle C_i, T_{j_1, \dots, j_k}(D_i^m) \rangle > 0 \quad (5.1)$$

for every $i \in \{1, \dots, N\}$ and for every subset $\{j_1, \dots, j_k\}$ of distinct integers of $\{1 \dots N\}$. [Notations of (4.1) and Definition 2.2.]

We may interpret these inequalities in a geometric way: we denote by a_i^m , (resp. b) the endpoints of the vectors D_i^m [resp. $T_{j_1, \dots, j_k}(D_i^m)$].

The condition (5.1) means that S^m is a k -attractor iff for all i , the Hamming sphere with center a_i^m and radius k is entirely on the same side of $H_i = C_i^\perp$ (orthogonal space of vector C_i) see Fig. 3. It means that

$$\begin{aligned} d(a_i^m, b) &= 2\sqrt{\delta(D_i^m, T_{j_1, \dots, j_k}(D_i^m))} \text{ (by Appendix 1)} \\ &= 2\sqrt{k} \\ \langle a_i^m, H_i \rangle &= \frac{\langle C_i, D_i^m \rangle}{\|C_i\|}. \end{aligned}$$

More precisely,

$$\begin{aligned} \langle C_i, T_{j_1, \dots, j_k}(D_i^m) \rangle &= \langle C_i, D_i^m \rangle \\ &= |\langle C_i, Ob \rangle - \langle C_i, Oa_i^m \rangle| \\ &= |\langle C_i, Ob - Oa_i^m \rangle| \\ &= 2 \sum_{i=1}^k |C_{ij}| \leq 2k \max_j |C_{ij}|, \end{aligned}$$

(because only k components differ). So (5.1) holds true as soon as

$$k < \frac{\langle C_i, D_i^m \rangle}{2 \max_j |C_{ij}|} \text{ for every } i.$$

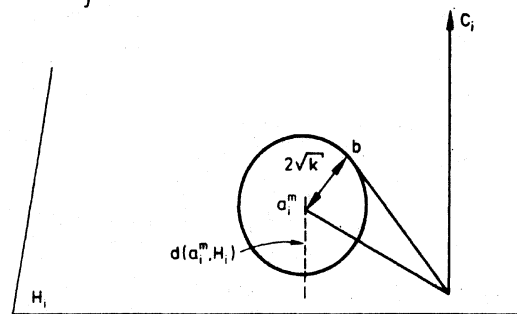


Fig. 3. Hamming sphere $B_d(a_i^m, k)$

Hence

Proposition 5.2. For a connection matrix $C=(C_{ij})$, whose rows are C_1, \dots, C_N , each pattern S^m is k -attractor at least up to any k such that

$$k < \frac{1}{2} \inf_i \frac{\langle C_b, D_i^m \rangle}{\max_j |C_{ij}|} = \gamma_m, \quad (5.2)$$

where $D_i^m = S_i^m S^m$.

So if k and k' are two integers satisfying (5.2), w.r.t. S^m (for k) and $S^{m'}$ (for k'), $S^m \in E_k$, $S^{m'} \in E_{k'}$, and $k+k' \leq \inf(\delta(S^m, S^{m'}), N - \delta(S^m, S^{m'}))$.

Now we proceed to simplify the expression of γ_m [in (5.2)] when C is given by (4.2), (ensuring patterns stability).

The choice of the C_i 's, in (4.4) has to ensure that $d(a_i^m, H_i)$ or γ_m are as big as possible.

First, we see that in (4.4) $\tilde{C}_i = 0$ is the best choice. Indeed if $C_i = \hat{C}_i + \tilde{C}_i$ with \hat{C}_i in the subspace \mathcal{V} spanned by $\{S^1, \dots, S^p\}$ and \tilde{C}_i in \mathcal{V}^\perp , we have

$$\langle C_b, D_i^m \rangle = \langle \hat{C}_b, D_i^m \rangle$$

and

$$\|C_i\|^2 = \|\hat{C}_i\|^2 + \|\tilde{C}_i\|^2, \text{ so } d(a_i^m, H_i) = \frac{\langle C_b, D_i^m \rangle}{\|C_i\|}$$

will be greater for $\tilde{C}_i = 0$.

Using (4.3),

$$\langle C_b, D_i^m \rangle = S_i^m \langle C_b, S^m \rangle = S_i^m (\alpha_i^m S_i^m) = \alpha_i^m. \quad (5.3)$$

As to the vector C_b , we write, from (4.4),

$$C_b = \Sigma(\Sigma' \Sigma)^{-1} \begin{pmatrix} \alpha_1^1 & S_1^1 \\ \vdots & \vdots \\ \alpha_p^p & S_p^p \end{pmatrix} \quad (5.4)$$

and

$$\begin{aligned} \|C_b\|^2 &= C_b' C_b = (\alpha_1^1 S_1^1, \dots, \alpha_p^p S_p^p) (\Sigma' \Sigma)^{-1} \begin{pmatrix} \alpha_1^1 & S_1^1 \\ \vdots & \vdots \\ \alpha_p^p & S_p^p \end{pmatrix} \\ &= (\alpha_1^1, \dots, \alpha_p^p) \text{diag}(S_1^1 \dots S_p^p) (\Sigma' \Sigma)^{-1} \\ &\quad \times \text{diag}(S_1^1 \dots S_p^p) \begin{pmatrix} \alpha_1^1 \\ \vdots \\ \alpha_p^p \end{pmatrix} \end{aligned}$$

[where $\text{diag}(S_1^1 \dots S_p^p)$ is the matrix whose diagonal elements are S_1^1, \dots, S_p^p and others are 0]

$$= \alpha_i^m W_i \alpha_i \text{ (obvious notations).}$$

We remark that if we denote by $D_i = (D_i^1, \dots, D_i^p) = (S_i^1 S^1, \dots, S_i^p S^p)$, i.e. the matrix of the patterns nor-

malized to +1 in the i -th component, we have

$$W_i = (D_i' D_i)^{-1}.$$

So

$$d^2(a_i^m, H_i) = \frac{(\alpha_i^m)^2}{\alpha_i^m W_i \alpha_i} \quad (5.5)$$

and

$$\gamma_m = \frac{1}{2} \inf_i \frac{\alpha_i^m}{\max_j |C_{ij}|}. \quad (5.6)$$

The inequality (5.2) does not give exact values of the sizes k_1, \dots, k_p of the domains of attraction (DA) of S^1, \dots, S^p , because it is only a sufficient condition.

However, the pattern S^m is attractive at least up to γ_m for $m=1, \dots, p$.

Using geometric notations let us consider patterns S^1, S^2 . Let us assume that $\delta(S^1, S^2)$ is small with respect to N . Since for every i such that $S_i^1 = S_i^2$, the distance $d(a_i^m, H_i)$ is "great", whereas the sizes of the DA of S^1 and S^2 are small, [since less than $\delta(S^1, S^2)$] this seems contradiction. But we must notice that on the contrary, for i such that $S_i^1 = -S_i^2$, we have $D_i^1 = -S^1$ or $D_i^2 = -S^2$ and $d(a_i^m, H_i)$ small, which leads to a small value of γ_m for $m=1, 2$.

Now, Proposition 5.2 can be completed by:

Proposition 5.7. For a connection matrix $C=(C_{ij})$, whose rows C_1, \dots, C_N are given by (4.4) with $\tilde{C}_i = 0$, each pattern S^m is attractor at least up to Hamming distance

$$\gamma_m = \frac{1}{2} \inf_i \frac{\alpha_i^m}{\max_j |C_{ij}|}. \quad (5.7)$$

6 Optimal Matrix

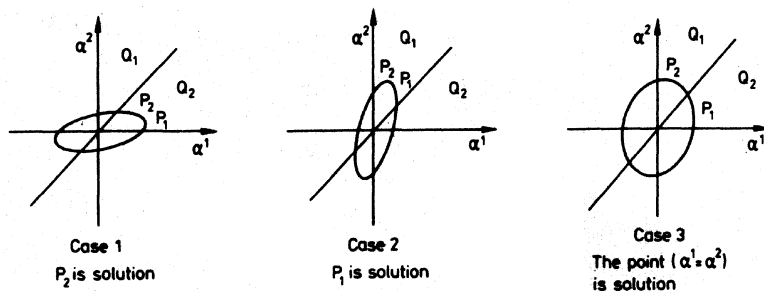
Let us start from an initial configuration S^0 , obtained by distorting one of the patterns, e.g. S^{m_0} . There will not be any identification error if S^0 belongs to the domain of attraction (DA) of S^{m_0} .

Hence the next definition:

Definition 6.1. A matrix C for which the patterns S^1, \dots, S^p are attractors, is *optimal* if it maximizes the minimum radius (Definition 2.6) of the DAs of the patterns.

Since we have not the exact value of these radii, we try to determine a matrix C , called semi-optimal, and which maximizes the minimum distance $d(a_i^m, H_i)$ (Fig. 3).

Thus we look for positive constants $\alpha_i^m, m=1, \dots, p, i=1, \dots, N$, which maximize for each $i, \inf_m d^2(a_i^m, H_i)$

Fig. 4. Case $p=2$

$= \inf_m \frac{(\alpha_i^m)^2}{\alpha_i^m W_i \alpha_i}$ [by (5.5)]. Let us sketch a construction of the α_i^m for fixed i .

a) We look for a vector $\alpha_i = (\alpha_i^1, \dots, \alpha_i^p)$ in $(\mathbb{R}^+)^p$, with $\alpha_i^m W_i \alpha_i = 1$.

b) We cut out $(\mathbb{R}^+)^p$ into quadrants Q_j defined by $\alpha_i^j = \min(\alpha_i^m)$, for $j=1, \dots, p$.

In each quadrant Q_j , we want to maximize $(\alpha_i^j)^2$, i.e. α_i^j . For instance, if $p=2$, we consider the ellipse whose equation is $\alpha_i^m W_i \alpha_i = 1$,

c) We compute the p points $P_j = (\alpha_i^1, \dots, \alpha_i^p)$, $j=1, \dots, p$, solutions of $\alpha_i^m W_i \alpha_i = 0$ and $\frac{\partial \alpha_i^j}{\partial \alpha_i^m} = -\frac{(W_i \alpha_i)_m}{(W_i \alpha_i)_j} = 0$ for $m \neq j$, and keep the points P_j belonging to the corresponding quadrant Q_j . We have

$$\max_{\alpha_i} \left(\min_m \frac{(\alpha_i^m)^2}{\alpha_i^m W_i \alpha_i} \right) = \max_j \{ (\alpha_i^j)^2 / P_j = (\alpha_i^1, \dots, \alpha_i^p) \in Q_j \}$$

under the condition that the above set is not empty.

d) If for all j , $P_j \notin Q_j$ (case 3), we restrict ourselves in the quadrants of $(\mathbb{R}^+)^{p-1}$, $(\mathbb{R}^+)^{p-2}$, etc. ... defined by inequalities such as:

$$\alpha_i^j = \alpha_i^j = \min(\alpha_i^m)$$

$$\alpha_i^j = \alpha_i^j = \alpha_i^j = \min(\alpha_i^m) \dots \text{and so on.}$$

If none of the P_j s successively found belongs to the convenient domain, we get the solution $\alpha_i^1 = \alpha_i^2 = \dots = \alpha_i^p$ (see Sect. 8).

We can sum up the results as follows:

Proposition 6.2. By iterating this approach for every $i=1, \dots, N$, we construct a matrix semi-optimal C , which leaves as much volume as possible, around each point α_i^m , for every i , with

$$C_i = \Sigma (\Sigma' \Sigma)^{-1} (\alpha_i^1 S_i^1 \dots \alpha_i^p S_i^p)' \quad (6.2)$$

for $\Sigma = (S^1, \dots, S^p)$ and $(\alpha_i^1, \dots, \alpha_i^p)$ positive yielding the maximum of $\min_m \frac{(\alpha_i^m)^2}{\alpha_i^m W_i \alpha_i}$, with $W_i = (D_i' D_i)^{-1}$ and $D_i = (S_i^1 S^1, \dots, S_i^p S^p)$.

7 Case of Orthogonal Patterns

In this section we assume that the patterns S^1, \dots, S^p are pairwise orthogonal.

In this case, $\Sigma' \Sigma = N \text{Id}_p$ and for every i , $D_i' D_i = N \text{Id}_p$ since the vectors D_i^m are also pairwise orthogonal.

So the ellipses of equations $\alpha_i^m W_i \alpha_i = 1$ are spheres and the research of the semi-optimal α_i^m , leads to the solution $\alpha_i^1 = \dots = \alpha_i^p$, (see the two-dimensional example, case 3, in Sect. 6).

Since the α_i^m are defined (for each i) up to a positive multiplicative constant, we may choose $\alpha_i^m = 1$ for every i and every m .

So from (5.4),

$$C_i = \Sigma \left(\frac{1}{N} \text{Id}_p \right) (S_i^1 \dots S_i^p)' \quad \text{and} \quad C_{ij} = \frac{1}{N} \sum_{m=1}^p S_i^m S_j^m. \quad (7.1)$$

In other words, in the orthogonal case, the semi-optimal solution corresponds to the equality of the α_i^m , and to the classical Hebbian connections. [In that case C is the projection matrix introduced by Kohonen (1970).]

Thus, we may compute γ_m and $\inf_i d^2(\alpha_i^m, H_i)$, using (5.5) and (5.6)

$$d^2(\alpha_i^m, H_i) = \frac{(\alpha_i^m)^2}{\alpha_i^m W_i \alpha_i} = \frac{N}{p}, \quad (7.2)$$

and

$$\gamma_m = \frac{1}{2} \inf_i \frac{\alpha_i^m}{\max_j |C_{ij}|} = \frac{N}{2} \inf_i \frac{1}{\max_j \left| \sum_{m=1}^p S_i^m S_j^m \right|}.$$

But $\max_j \left| \sum_{m=1}^p S_i^m S_j^m \right| = p$ is obtained for $j=i$, since all terms of the sum are then equal to $+1$.

Hence we get

$$\gamma_m = \frac{N}{2p} \quad (\text{independent of } m). \quad (7.3)$$

So we find again the result of Sect. 3, which we now enounce as follows:

Proposition 7.4. *In the case of pairwise orthogonal patterns, the Hebbian connection matrix (3.1) is at least semi-optimal and ensures the attractivity of the patterns at least up to the Hamming distance $\frac{N}{2p}$.*

In the not orthogonal case, we remark that if $\delta(S^m, S^{m'}) > \frac{N}{2}$ for some (m, m') , then $\delta(S^m, -S^{m'}) < \frac{N}{2}$; therefore the volume "winned" by the component i will be "lost" by component i' such that $D_i^m = D_i^{m'}$ and $D_i^{m'} = -D_i^m$.

The sizes of the domains of attraction depend on the minima over i ; we see that the favourable case is the orthogonal one.

Hence the practical interest of
 - picking at random [with $\mathbf{P}(S_i^m = +1) = \mathbf{P}(S_i^m = -1) = \frac{1}{2}$], as in spin glasses, ensuring, on average, orthogonality.
 - a deterministic encoding of the objects to be recognized, by means of pairwise orthogonal patterns.

8 Domains of Attraction with Equal Sizes

The research of the optimal matrix C sketched in Sect. 6 is tedious. We shall simplify it by taking all constants $\alpha_i^m = +1$ (Personnaz et al.'s method), which is equivalent to equal attractivity of each pattern.

We need now a geometric interpretation and a lower bound of the common size of the domains of attraction (DA).

Since $\alpha_i^m = \langle C_i, D_i^m \rangle$, by (5.3), in the case $\alpha_i^m = 1$ for every $m = 1, \dots, p$, we have to determine a hyperplane H_i , orthogonal to C_i at equal distance of all the points a_i^m for $m = 1, \dots, p$ (see Figs. 2 and 3). The vector C_i is orthogonal to the affine space containing all points a_i^1, \dots, a_i^p .

We write, from (5.4),

$$C_i = \Sigma(\Sigma' \Sigma)^{-1} \begin{pmatrix} S_i^1 \\ \vdots \\ S_i^p \end{pmatrix} \quad \text{i.e.} \quad C = \Sigma(\Sigma' \Sigma)^{-1} \Sigma', \quad (8.1)$$

which is the orthogonal projection matrix of \mathbf{R}^N on the vector space \mathcal{V} , spanned by S^1, \dots, S^p (introduced by Kohonen 1970). The matrix C is symmetric, and its columns (equal to its rows) are images of the vectors e_1, \dots, e_N (canonical basis of \mathbf{R}^N) by this projection.

So we have $C_i = \text{proj}_{\mathcal{V}}(e_i)$, $D_i^m \in \mathcal{V}$, and $\langle C_i, D_i^m \rangle = \langle e_i, D_i^m \rangle = +1$ by the definition of the vectors D_i^m , whose i -th component is equal to $+1$.

Hence, noting that $\max_{i,j} |C_{ij}| = \max_i (C_{ii}) = \max_i \|C_i\|^2$ since C is a projection matrix ($C^2 = C$), we get

Proposition 8.2. *For the matrix $C = \Sigma(\Sigma' \Sigma)^{-1} \Sigma'$ with $\Sigma = (S^1, \dots, S^p)$, each pattern S^m is attractor at least up to Hamming distance*

$$\gamma = \frac{1}{2} \frac{1}{\max_i \|C_i\|^2} = \frac{1}{2} \frac{1}{\max_i C_{ii}}. \quad (8.2)$$

See numerical examples of evaluation of γ in Appendix 2.

Remark 8.3. The matrix C obtained when $\alpha_i^m = 1$, for every i, m , is symmetric. However for what concerns the algorithm, it is equivalent to the matrix obtained by multiplying each row by a positive arbitrary constant: this matrix is no more symmetric.

Neither it is in the optimal case of Sect. 6.

9 Energy

The patterns S^1, \dots, S^p span the subspace \mathcal{V} , and the matrix

$$C = \Sigma(\Sigma' \Sigma)^{-1} \Sigma' \quad [\text{see (8.1)}]$$

is the symmetric projection matrix on \mathcal{V} , whose elements are in interval $[-1, +1]$, with $C = C^2 = C' C$, hence

$$C_{ij} = \langle C_i, C_j \rangle = \langle \text{proj}_{\mathcal{V}}(e_i), \text{proj}_{\mathcal{V}}(e_j) \rangle.$$

In that case, we define an energy function

$$E(S) = -\frac{1}{2} H(S) \quad (9.1)$$

with

$$H(S) = \sum_{i,j} C_{ij} S_i S_j. \quad (9.2)$$

We have the following theorem (with C symmetric).

Theorem 9.3. i) $H(S)$ increases when the system evolves (in sequential or parallel algorithm).

ii) The k -attractor states are local maxima of H . More precisely, if S is k -attractor, and $S^* = T_{j_1, \dots, j_k}(S)$ (notations of Sect. 5), then

$$H(S) > H(S^*).$$

iii) The learned patterns are absolute maxima of H and any absolute maximum is a stable state.

Demonstration. i) is clear. We prove ii) noting that $H(S) - H(S^*) = 4 \sum_{i \in I} S_i \left(\sum_{j \notin I} C_{ij} S_j \right)$ for $I = \{j_1, \dots, j_k\}$.

Since S is k -attractor, it is stable and for all i ,

$S_i \left(\sum_j C_{ij} S_j \right) > 0$ (2.1) and $S_i \left(\sum_{j \neq i} C_{ij} S_j - \sum_{j \neq i} C_{ij} S_j \right) > 0$ (2.2), hence $H(S) > H(S^*)$. For iii), we write $H(S) = S'CS = (CS)'CS = \|CS\|^2 = \|\text{proj}_{\mathcal{V}}(S)\|^2$ for all configuration $S \in \{-1, +1\}^N$.

Since $\|S\|^2 = N$, $H(S) \leq N$ for all S with equality iff $S \in \mathcal{V} = \text{Vect}(S^1, \dots, S^p)$. So patterns S^1, \dots, S^p are absolute maxima, and also the opposite $-S^1, \dots, -S^p$.

Any linear combination of the learned patterns, element of $\{-1, +1\}^N$, (if exists) will be absolute maxima and spurious stable state: $H(S) = N$ iff $S \in \mathcal{V}$, i.e. iff $CS = S$ which is the stability. \square

Note that if the patterns S^1, \dots, S^p are orthogonal,

$$H(S) = \sum_{m=1}^p \|\text{proj}_{S^m}(S)\|^2 = \frac{1}{N} \sum_{m=1}^p \langle S, S^m \rangle^2.$$

See in Appendix 3 examples of spurious stable and attractor states, which are local maxima of H . At temperature $T=0$ (studied here) the algorithm may reach some of these states, but when $T>0$ using the annealing method, only the absolute maxima of H will be reached, i.e. the spurious stable states belonging to \mathcal{V} .

10 Provisional Conclusions

1) To choose the sizes of the domains of attraction (DA) in such a way they are equal for all patterns S^1, \dots, S^p enables us to give no preference to any pattern.

Assume that we start from an initial configuration S^0 distorted from S^{m_0} . We modelize this disturbance: an error occurs in each component $i=1, \dots, N$, independently with a small probability q . The number of errors $\delta(S^0, S^{m_0})$ is a Binomial distribution $\mathcal{B}(N, q)$.

Starting from S^0 , the algorithm acts and the probability that it gives a good answer (S^{m_0}), is bounded from below by the probability that S^0 belongs to the DA of S^{m_0} , with radius k_{m_0} , i.e. by $\mathbb{P}(\delta(S^0, S^{m_0}) \leq k_{m_0})$.

Of course, we can approximate this probability, by substituting the Normal Distribution $\mathcal{N}(Nq, Nq)$ to the Binomial distribution. (N is great and q small.)

So selecting an algorithm, i.e. a matrix C such that all the k_m are equal, does yield the same lower bounds for all the probabilities of correct identification, whatever is the configuration S^m to identify.

2) Now let us assume that the choice of the configuration S^m to identify, is made with a probability $(p_m) \left(\sum_m p_m = 1 \right)$. Then the probability of wrong identification is less than $\varepsilon = \sum_m p_m \mathbb{P}(\delta(S^0, S^m) > k_m/S^0)$ arises from S^m and if an identification error for S^m costs g_m ,

the mean error cost is less than $G = \sum_m g_m p_m \mathbb{P}(\delta(S^0, S^m) > k_m/S^0)$ arises from S^m .

In that case, the choice of the semi-optimal matrix C , described in Sect. 6, yields a reasonable lower bound ε and G .

It remains that these calculations are approximations (apart from the fact we cannot calculate the exact size k_m of the DA of the patterns S^m , w.r.t. a given matrix C).

3) If $\delta(S^0, S^m) > k_m$, there is an error if S^0 falls in the DA of another configuration, but we do not know what happens if S^0 does not belong to a DA of the S^m . So S^0 may be attracted by one of the spurious configurations made attractive by the matrix C , for instance a configuration $-S^m$, or other linear combination of the S^m . (See Appendix for numerical examples.)

4) The complete study of the deterministic algorithm (Temperature $T=0$) enables us to see that only the spurious stable state belonging to the subspace \mathcal{V} spanned by the patterns, remain when the temperature $T>0$ using the annealing method. This result confirms the results of Amit et al. (1985a, b), and shows it is true for all N , and not only when $N \rightarrow +\infty$.

5) The algorithm ensures a perfect retrieval of patterns if the initial state S^0 satisfies $\delta(S^0, S^m) < \gamma$ [defined in (8.2)] for some S^m . But, of course, the algorithm ensures a very good retrieval if S^0 is more distant from the patterns. Let be \mathcal{B} the set of initial states which lead to some pattern. The simulations show that \mathcal{B} occupies a great portion of the hypercube [see numerical evaluations in Peretto and Niez (1985), which agree with examples of Appendix 3]. Of course, the size of \mathcal{B} decreases when p increases, for a fixed N , like the attractivity γ .

Appendix 1: Hamming Distance

For $S, S' \in \{-1, +1\}^N$, we denote by $\delta(S, S')$ the number of distinct components of S and S' : it is the Hamming distance of S and S' .

The following properties are easy to check

1) δ is a distance in $\{-1, +1\}^N$,

2) $\delta(S, -S) = N$; if $\sum_{i=1}^N S_i S'_i = 0$, i.e. if S and S' are orthogonal

(only with N pair), $\delta(S, S') = \frac{N}{2}$.

3) $\delta(S, -S') = N - \delta(S, S')$.

If d is the Euclidian distance in \mathbb{R}^N , $\langle \cdot \rangle$ the Euclidian product, $\|\cdot\|$ the Euclidian norm, \cos and \sin the usual trigonometric functions, we have

4) $\|S\| = \sqrt{N}$, $\delta(S, S') = \frac{1}{2}(N - \langle S, S' \rangle) = N \sin^2 \left(\frac{\delta(S, S')}{2} \right)$ and

$d(S, S') = 2\sqrt{\delta(S, S')}$.

Appendix 2: Evaluation of Gamma

For different values of N and p , we compare γ in various cases

- a) orthogonal case $\gamma = N/2p$
- b) random choice of patterns
- c) random choice of (S^2, \dots, S^p) and $S^1 = (1, \dots, 1)$
- d) a non orthogonal case.

$N=20$	a	b	c	d
$p=3$	3.3	1.9	2.4	1.5
$p=4$	2.5	1.4	1.9	1.1
$p=5$	2	1.2	1.5	0.7

$N=30$	a	b	c	d
$p=3$	5	3.7	3.8	2
$p=4$	3.7	2.6	2.7	1
$p=5$	3	1.8	1.9	0.7

$N=40$	a	b	c	d
$p=3$	6.7	5.3	4.9	3.1
$p=4$	5	4	3.9	1.2
$p=5$	4	3.1	3.1	0.8
$p=6$	3.3	2.3	2.4	0.7

$N=50$	a	b	c	d
$p=4$	6.3	5.8	5.1	1.4
$p=5$	5	3.8	3.9	0.9
$p=6$	4.2	3.0	3.1	0.8

Note that if we want to retrieve exactly a pattern transmitted with 10% errors at most, we must choose $p \ll N$, approximately $p \sim 0.10N$ (for random case). This agrees with numerical results of Hopfield (1982) or Amit et al. (1985a, b) for instance. (Take account they use a connection matrix which does not ensure the stability of all patterns.)

Appendix 3

We indicate numerical results of various simulations.

We determine all the stable states, their order of attractivity, their Hamming distance from patterns, and their energy.

In each case, 1000 trials are performed with initial state at random. The system evolves very quickly (at most 4 parallel iteration steps) to one of the stable states. We indicate the number of trials ending into each stable states. See that the value of γ above calculated is not subestimated, and gives a good estimation of the BA' size.

Example 1

$N=20, p=3, \gamma=2.4$ (case c of Appendix 2)

$$S^1 = 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1$$

$$S^2 = 1 \ 1 \ 1 \ -1 \ 1 \ -1 \ -1 \ -1 \ -1 \ 1 \ 1 \ -1 \ 1 \ -1 \ 1 \ -1 \ -1 \ 1 \ -1$$

$$S^3 = -1 \ -1 \ 1 \ 1 \ -1 \ 1 \ -1 \ 1 \ 1 \ 1 \ -1 \ -1 \ -1 \ 1 \ -1 \ -1 \ -1 \ -1 \ 1 \ 1$$

Example 2

$N=20, p=3, \gamma=1.9$

S^1, S^2, S^3 are random (case b)

Example 3

$N=20, p=4, \gamma=1.4$

S^1, S^2, S^3, S^4 are random (case b)

Example 4

$N=20, p=4, \gamma=1.1$

S^1, S^2, S^3, S^4 are not orthogonal ($\delta(S^1, S^3) = \delta(S^1, S^4) = 3$)

Example 1	Nb of trials ending into one of the patterns (or opposite) = 67%													
State	S^1	$-S^1$	S^2	$-S^2$	S^3	$-S^3$	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}	S_{14}
Att. order	2	2	2	2	2	2	1	1	0	0	0	0	0	0
$\delta(S^m)$ {	0	20	11	9	11	9	15	5	17	3	14	6	6	14
	11	9	0	20	12	8	14	6	6	14	15	5	17	3
	11	9	12	8	0	20	14	6	6	14	3	17	5	15
$H(S)$	20	20	20	20	20	20	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4
% n.b. trials ending	11.7	10.7	10.2	11.6	11.1	12.2	6.7	5.8	5.4	4.0	1.9	3.1	3.3	2.3

Example 2	Nb of trials ending into one of the patterns (or opposite) = 54%													
State	S^1	$-S^1$	S^2	$-S^2$	S^3	$-S^3$	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}	S_{14}
Att. order	1	1	1	1	1	1	2	2	0	0	0	0	0	0
$\delta(S^m)$	0	20	12	8	12	8	14	6	18	6	6	2	14	14
	12	8	0	20	12	8	14	6	6	18	6	14	2	14
	12	8	12	8	0	20	14	6	6	6	18	14	14	2
$H(S)$	20	20	20	20	20	20	16	16	16	16	16	16	16	16
% n.b. trials ending	9.6	8.8	9.9	8.1	8.3	9.6	9.9	8.2	4.7	6.0	4.3	5.0	2.8	4.8

Example 3. In that case S^1, S^2, S^3, S^4 and their opposite are 1-attractors (compare with $\gamma = 1.4$). And there are 40 stable states whose energy are 14.1, 15.1, 15.2, 15.6 or 16. The Hamming distance between a pattern and a spurious stable state can be 2, and that shows that the evaluation of γ is exact.

Among 1000 trials, the system ends into one of the pattern in 52% of the cases

Example 4. There are 8 1-attractors, S^1, S^2, S^3, S^4 , their opposites and 4 configurations more which have the same energy ($H(S) = 20$). There are 4 stable states, and 4 1-attractors,

State	S^1	$-S^1$	S^2	$-S^2$	S^3	$-S^3$	S^4	$-S^4$	S_9	S_{10}	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}	S_{16}
Att. order	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
$\delta(S^m)$	0	20	11	9	3	17	3	17	6	14	6	14	12	9	8	11
	11	9	0	20	14	6	10	10	13	3	17	7	5	2	15	18
	3	17	14	6	0	20	6	14	3	11	9	17	9	12	11	8
	3	17	10	10	6	14	0	20	3	13	7	17	15	12	5	8
$H(S)$	20	20	20	20	20	20	20	20	20	20	20	20	17.8	17.8	17.8	17.8

The system ends into one of the pattern is 52% of the cases

References

- Amit DJ, Gutfreund H, Sompolinsky H (1985a) Phys Rev A 32:1007
 Amit DJ, Gutfreund H, Sompolinsky H (1985b) Phys Rev Lett 55:1530
 Caianiello ER (1961) J Theor Biol 1:204
 Cooper LN, Liberman F, Oja E (1978) Biol Cybern 33:9
 Hebb DO (1949) The organization of behavior. Wiley, New York
 Hopfield J (1982) Proc Natl Acad Sci USA 79:2554
 Kohonen T (1970) Helsinki University of Technical Report TKK-F-A130
 Kohonen T (1972) IEEE Trans C-21:353-359
 Kohonen T (1984) Self-organization and associative memory. Springer, Berlin Heidelberg New York
 Kohonen T, Oja E (1976) Biol Cybern 21:85-95
 Kohonen T, Lehtiö P, Rovamo J (1974) Annales Academiae Scientiarum Fennicae A: V Medica 167
 Kohonen T, Lehtiö P, Rovamo J, Hyvärinen J, Bry K, Vainio L (1977) Neuroscience 2:1065-1076
 Little WA (1974) Math Biosci 19:101

- Nakano K (1972) IEEE Trans SMC-2:380-388
 Peretto P (1984) Biol Cybern 50:51
 Peretto P, Niez JJ (1985) Collective properties of neural networks. In: Proceedings of the winter school on "Disordered Systems and Biological Organization"
 Personnaz L, Guyon I, Dreyfus G (1985) J Phys Lett 46:359
 Personnaz L, Guyon I, Dreyfus G, Toulouse G (1986) J Stat Phys 43:411
 Rosenblatt F (1958) Psych Rev 65:386
 Weichbuch G, Fogelman-Soulie F (1985) J Phys. Lett 46:624

Received: February 26, 1987

Accepted in revised form: July 11, 1987

Dr. Marie Cottrell
 UA 743 CNRS Statistique Appliquée
 Laboratoire de Mathématique Bât. 425
 Université Paris XI
 F-91405 Orsay Cedex
 France

QUATRIEME PARTIE

ANALYSE MATHEMATIQUE D'UN MODELE DU CORTEX
CEREBELLEUX.

ANALYSE MATHÉMATIQUE D'UN MODÈLE DU CORTEX CÉRÉBELLEUX :

I. EFFET DE L'ACTION INHIBITRICE DES COLLATÉRALES RÉCURRENTES

*Marie COTTRELL
U.A. C.N.R.S. 743 "Statistique Appliquée"
Laboratoire de Mathématique, Bât. 425
Université Paris 11
F - 91405 ORSAY Cedex*

Nous analysons le comportement d'un réseau de neurones reliés uniquement par des liaisons inhibitrices, en l'absence de stimuli extérieurs.

L'état du réseau est un processus de Markov dont nous étudions l'ergodicité ou caractérisons la divergence, suivant les paramètres du système.

Ce réseau modélise l'état du cortex cérébelleux chez le rat de quelques jours.

MOTS CLES :

**Processus de Markov
Réseaux de neurones
Cortex cérébelleux
Liaisons inhibitrices**

I. MODELISATION DU CORTEX CEREBELLEUX.

INTRODUCTION.

Cet article a eu pour point de départ une collaboration avec le Docteur H. Axelrad et C. Bernard (Laboratoire de Neurophysiologie de la Faculté de Médecine Pitié-Salpêtrière) et B. Giraud (Laboratoire de Physique théorique au CEA).

Il s'agit d'une contribution à la compréhension du mode de fonctionnement du cortex du cervelet et en particulier du rôle de l'inhibition par les collatérales récurrentes, en menant ensemble

- un travail de modélisation et de simulation.
- l'analyse mathématique du modèle
- un travail expérimental *in vivo*.

Voir AXELRAD et al. (1985), GIRAUD et al. (1985), AXELRAD et al. (1987).

1. PLACE DE NOTRE TRAVAIL.

De nombreux modèles de réseaux de neurones formels, d'automates cellulaires, ont été proposés ces dernières années. La bibliographie en est très importante c'est un sujet qui connaît actuellement un immense développement. Voir par exemple dans KOHONEN (1984) ou RUMELHART et Mc CLELLAND (1986) une présentation de beaucoup des aspects des recherches en cours et une abondante bibliographie. On peut consulter aussi la liste des intervenants au congrès qui s'est tenu à San Diego (USA) en juin 1987, sous les auspices de IEEE.

Ces neurones sont organisés en réseaux fonctionnant de façon dynamique, avec des règles (déterministes ou stochastiques) d'excitation, d'inhibition, de coopération ou compétition, bref d'interaction. Ces réseaux ont des fonctions d'apprentissage, de mémoire, de reconnaissance de formes, font du traitement d'information en parallèle, et ont d'ailleurs permis un extraordinaire développement et la construction de calculateurs parallèles.

Mais ils ont peu de chose à voir avec les systèmes neurobiologiques réels, même si l'ensemble de ces travaux font progresser l'étude des mécanismes du cerveau et du système nerveux central.

Ces réseaux ont des comportements collectifs intéressants, mais il est difficile de les comparer à des réseaux de neurones réels. Beaucoup de biologistes font de sérieuses réserves sur l'adéquation de ces modèles à la réalité neurobiologique.

Par ailleurs, on connaît assez précisément les mécanismes physico-chimiques de l'activité neuronale et on sait enregistrer à l'aide de micro-électrodes les décharges électriques (*spikes*) successives d'un ou plusieurs neurones. Jusqu'à ces dernières années, on ne pouvait enregistrer qu'un ou deux neurones à la fois, mais maintenant les méthodes optiques permettent l'enregistrement simultané de toute une population de neurones (par exemple BLASDEL et SALAMA (1986)).

Cependant si de nombreux modèles ont été proposés pour mimer l'activité d'un seul neurone (voir par exemple SAMPATH et SRINIVASAN (1977)), la modélisation "réaliste" de l'activité simultanée d'un réseau de neurones reste une tâche délicate.

D'où l'idée de notre travail : étudier un système neurobiologique réel : le cortex cérébelleux, dont la structure morphologique et le fonctionnement physiologique soit assez bien décrit et compris pour qu'on puisse en faire un modèle simplifié mais réaliste, pour lequel on pourra comparer les résultats théoriques et expérimentaux.

2. LE CORTÈX CÉRÉBELLÉUX.

Le cortex cérébelleux, dont le rôle est de coordonner les mouvements, est une partie du système nerveux, dont les caractéristiques sont bien particulières.

Il a fait l'objet d'un très grand nombre d'études neurophysiologiques et théoriques, (ECCLES et al. (1967), MARR (1969), ALBUS (1971), PALAY et CHAN-PALAY (1974), ITO (1984), CHAUVET (1986)) pour ne citer que quelques auteurs.

Sa structure est cristalline, invariante par déplacement le long de sa surface.

Il est composé d'un nombre limité -5- de types cellulaires :

- * les plus grandes cellules sont les *cellules de Purkinjie* (notées CP). Ce sont les seules qui sont en communication avec l'extérieur du cervelet, avec deux entrées (fibres grimpantes et moussues) et une sortie.

- * entre les cellules du Purkinjie, se trouvent les *interneurones* (trois types : cellules en panier, cellules en étoiles, cellules de Golgi) qui jouent un rôle *inhibiteur*.

- * les CP sont reliées par des *axones collatéraux récurrents*, qui jouent également un rôle *inhibiteur*.

- * par l'intermédiaire des *cellules grains*, les afférences moussues (entrées) amenant l'information extérieure se ramifient en *fibres parallèles* qui en contactant plusieurs cellules de Purkinjie jouent un rôle *activateur*.

Les cellules de Purkinje sont disposées régulièrement suivant un réseau bidimensionnel le long de la surface du cortex cérébelleux.

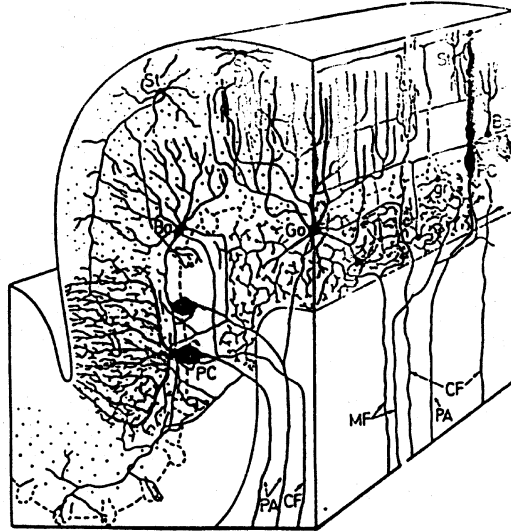


Figure 1.a

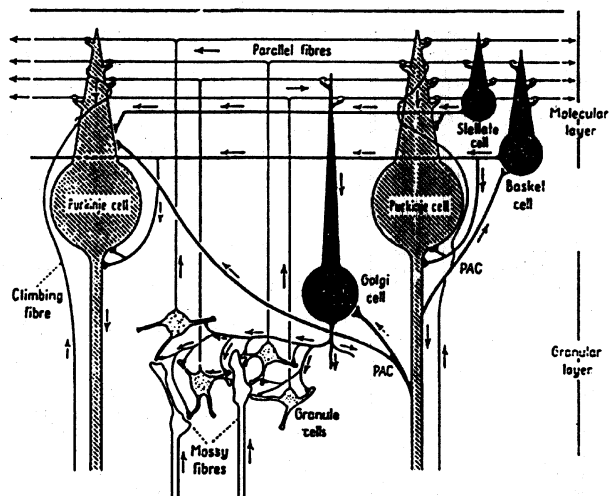


Figure 1.b

(D'après ECCLES)

Pour étudier le rôle propre du cortex cérébelleux, on schématise les différents constituants de la couche moléculaire, et on décompose l'analyse du système global en analysant des sous-systèmes de complexité croissante comme suit :

Système A



Figure 2.a

Système B



Figure 2.b

Système C



Figure 2.c

Le système complet contient de plus des interneurones (inhibiteurs), et les fibres grimpantes (excitatrices), dont on peut reconnaître facilement les décharges en salves.

Dans cette étude, on se borne aux systèmes A et B, pour dégager le rôle organisateur, structurant au niveau du réseau tout entier, de l'inhibition.

Il se trouve en outre que chez le jeune rat entre 6 et 9 jours, les fibres parallèles et les interneurones sont absents. Les collatérales récurrentes sont présentes en grand nombre. Elles régressent d'ailleurs au fur et à mesure du développement des fibres et des interneurones (CREPEL et al. (1980)). Les fibres grimpantes sont présentes dès le 3^e jour, et chacune contacte alors plusieurs cellules de Purkinjie, dans le même plan sagittal (axe transversal).

Comme la fréquence de décharge des fibres grimpantes est plus faible que la fréquence des décharges spontanées des cellules de Purkinjie, et que la forme de leurs salves est caractéristique, on peut étudier le modèle B *in vivo*, en enregistrant les activités des cellules de Purkinjie chez de jeunes rats. De plus un enregistrement du tissu *in vitro* permet d'éliminer les décharges des fibres grimpantes.

Par ailleurs, comme on connaît les antagonistes (bicuculline et bicrotoxine) du neurotransmetteur des cellules de Purkinjie (GABA), on peut découpler chimiquement les cellules et donc comparer le modèle B au modèle A *in vivo*.

3. VARIABLE D'ETUDE. MODELE PROPOSE PAR H. AXELRAD.

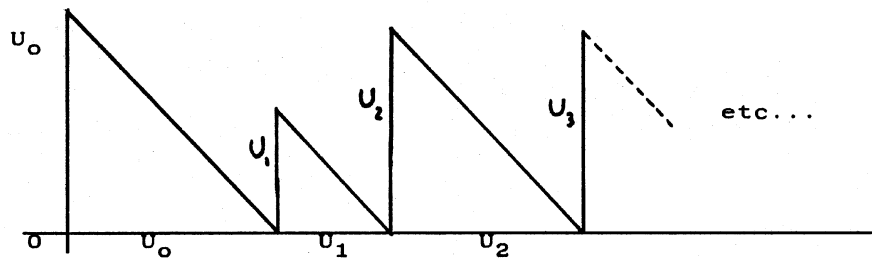
Les décharges (*spikes*) émises par les cellules de Purkinjie étant d'amplitude constante et de durée fixe (1 ms), ce qui caractérise l'activité d'une cellule, c'est la fréquence des *spikes* et donc les intervalles inter-*spikes*.

En l'absence de liaisons intra-cérébelleuses et avec l'extérieur, les intervalles inter-*spikes* forment un processus de renouvellement de loi notée \mathcal{F} , c'est-à-dire une suite de variables aléatoires indépendantes et de même loi \mathcal{F} (COX et LEWIS (1966), LANDOLT et CORREIA (1978)).

On prend comme convention que l'état du neurone i , à l'instant t est le temps à attendre jusqu'à la prochaine décharge. Ce temps décroît linéairement en fonction du temps et la décharge se produit quand il atteint la valeur 0 (convention commode).

A chaque instant de décharge, une variable aléatoire U de loi \mathcal{F} est tirée, indépendamment du passé. La loi \mathcal{F} correspond à une v.a. positive, bornée ou non.

Une trajectoire de l'état du neurone i est donc de la forme :



Evolution de l'état d'un neurone isolé.

Figure 3

Mais dans le modèle B, les neurones, numérotés (i, j) , rangés dans un réseau bidimensionnel, sont reliés par les collatérales inhibitrices. Chaque neurone (i, j) est lié aux neurones d'un voisinage $\mathcal{V}(i, j)$.

L'effet de l'inhibition collatérale est de retarder ou même supprimer les décharges prévues des neurones voisins.

Le modèle proposé par H. Axelrad et al. a des règles assez fines pour mimer précisément la réalité.

On en trouvera la description précise dans AXELRAD et al. (1987).

La suite de cet article est l'analyse mathématique d'un modèle de fonctionnement du système B, (Fig. 2.b : cellules de Purkinjie avec collatérales inhibitrices) simplifié pour rendre possibles les démonstrations, tout en gardant les mêmes principes essentiels.

II. MODELE MATHEMATIQUE DU RESEAU. MODELE (M). CELLULES AVEC COLLATERALES INHIBITRICES.

1. DEFINITION. DESCRIPTION.

Résumons les notations utilisées.

Les neurones sont disposés dans un carré plan, et numérotés (i, j) , $1 \leq i, j \leq n$.

L'état du système à l'instant t est décrit par le vecteur $x^t = (x_{ij}^t)$, où x_{ij}^t est une valeur positive état du neurone (i, j) .

A chaque instant de décharge, le neurone concerné est réinitialisé par le tirage d'un état suivant une loi de distribution \mathcal{F} . On envisage différentes lois, qu'on suppose à valeurs positives.

Dans le système de neurones non couplés (système A, Fig. 2.a), chaque neurone fonctionne indépendamment des autres, et en chaque site on observe un processus de renouvellement de loi \mathcal{F} (Fig. 3).

Dans le système B, Fig. 2.b, les neurones interagissent et l'inhibition est modélisée de la manière suivante : chaque fois qu'un neurone émet un spike, il inhibe ses voisins : leurs états sont alors incrémentés d'une quantité positive, ce qui retarde d'autant leur prochaine décharge. On considère différents systèmes de voisins d'un neurone (i, j) : par exemple

$$\begin{array}{l}
 V_4(i, j) = \begin{array}{ccccc} & & * & & \\ & * & \ominus & * & \\ & & * & & \end{array} \\
 \\
 V_8(i, j) = \begin{array}{ccccc} & * & * & * & \\ & * & \ominus & * & \\ & * & * & * & \end{array} \\
 \\
 V_{14}(i, j) = \begin{array}{cccccc} & * & * & * & * & * \\ & * & * & \ominus & * & * \\ & * & * & * & * & * \end{array} \\
 \\
 V_6(i, j) = \begin{array}{ccccc} & & * & & \\ & * & * & \ominus & * & * \\ & & * & & & \end{array} \\
 \\
 V_{12}(i, j) = \begin{array}{ccccc} & & & * & & \\ & & * & * & * & \\ & * & * & \ominus & * & * \\ & & * & * & * & \\ & & & * & & \end{array} \\
 \\
 V_{24}(i, j) = \begin{array}{cccccc} & * & * & * & * & * \\ & * & * & * & * & * \\ & * & * & \ominus & * & * \\ & * & * & * & * & * \\ & * & * & * & * & * \end{array}
 \end{array}$$

(les voisinages V_8 et V_{14} correspondent au fait que les cellules contactées par les collatérales récurrentes issues d'une cellule de Purkinjé forment approximativement une ellipse allongée dans le plan sagittal).

Le processus $x^t = (x_{ij}^t)$ est donc défini de la manière suivante :

i) *Algorithme à temps et espace discret.* (II.1.1)

Au temps 0, chaque état est initialisé "au hasard" suivant la loi \mathcal{F} , et indépendamment des autres. La loi \mathcal{F} est une loi discrète sur N .

Soit $X^t = (X_{ij}^t)$ l'état du réseau à l'instant t ,

- si $X_{ij}^t = 0$, le neurone (i,j) décharge et X_{ij}^{t+1} est une réalisation d'une variable aléatoire U de loi \mathcal{F} , indépendante de X^t .
- si (i,j) est voisin d'un neurone (i_0, j_0) qui émet un spike au temps t , $X_{ij}^{t+1} = X_{ij}^t + \theta - 1$ où $\theta \in \mathbb{N}$ est le "délai" modélisant l'inhibition causée par la décharge du neurone (i_0, j_0) .
- sinon $X_{ij}^{t+1} = X_{ij}^t - 1$.

Le processus X^t défini ainsi est une chaîne de Markov, à temps discret, à valeurs dans \mathbb{N}^{n^2} , à probabilités de transition stationnaires. L'unité de temps correspond à la durée d'un spike.

En fonctionnement réel, un certain nombre de neurones sont simultanément excités, d'autant que la durée d'un spike est typiquement de 1ms. C'est ce que l'on observe dans les simulations de la chaîne X^t . Notons que si un neurone est voisin de plusieurs neurones qui déchargent, il n'est inhibé qu'une fois, on ne cumule pas les inhibitions. Mais l'étude mathématique est un peu compliquée par la possibilité de décharges simultanées de plusieurs neurones.

On utilise donc une version continue du processus X^t , où le temps est continu, l'instant de décharge identifié avec l'instant de début du spike, et la probabilité que plusieurs neurones initient leurs décharges simultanément est nulle.

D'où la définition du modèle (M).

ii) Algorithme à temps et espace continus.

(II.1.2)

Au temps 0, les (X_{ij}^0) sont initialisées de manière quelconque (mais positives et sans ex aequo). La loi \mathcal{F} est alors une loi qui admet une densité f strictement positive par rapport à la mesure de Lebesgue sur \mathbb{R}^+ (ou éventuellement sur un intervalle de la forme $[0, L]$ si l'on choisit une loi \mathcal{F} à support borné).

Si $X^t = (X_{ij}^t)$ est l'état du réseau à l'instant t

- si aucune des composantes de X^t n'est nulle, (aucun neurone ne décharge), $X_{ij}^{t+dt} = X_{ij}^t - dt$, pour tout (i,j) .
- si t est instant de décharge d'un des neurones, soit $X_{i_0 j_0}^t = 0$, alors
 - $X_{i_0 j_0}^{t+dt}$ est une réalisation d'une v.a. U de loi \mathcal{F} , indépendante de $X^t, -dt$.
 - $X_{ij}^{t+dt} = X_{ij}^t + \theta - dt$, pour $(i,j) \in V(i_0, j_0)$, avec θ fixe, réel positif.

$$\bullet x_{ij}^{t+dt} = x_{ij}^t - dt, \text{ pour } (i,j) \in V(i_0, j_0).$$

Le processus X^t défini ci-dessus, dans sa version continue à droite est un processus de Markov, à temps continu, à valeurs dans $(\mathbb{R}^+)^{n^2}$, à noyau de transition stationnaire donné par

$$\begin{cases} P(x_{11}, x_{12}, \dots, x_{nn}, dy_{11}, dy_{12}, \dots, dy_{nn}, dt) \\ = \delta(y_{ij} = x_{ij} - dt), \text{ si } \prod_{(i,j)} x_{ij} \neq 0 ; \\ = f(y_{i_0 j_0}) dy_{i_0 j_0} \otimes \left[\sum_{(i,j) \in V(i_0, j_0)} \delta(y_{ij} = x_{ij} + \theta - dt) \right] \\ \otimes \left[\sum_{(i,j) \in V(i_0, j_0) \cup \{(i_0, j_0)\}} \delta(y_{ij} = x_{ij} - dt) \right], \text{ si } x_{i_0 j_0} = 0 ; \end{cases} \quad (\text{II.1.3})$$

Pour éviter les ex aequo, et donc définir l'indice (i_0, j_0) de manière unique p.s., on prend comme système d'états

$$\varphi = (\mathbb{R}^+)^{n^2} \cap \{x/x_{ij} - x_{i'j'} \in \theta\mathbb{Z}, \text{ pour } (i,j) \neq (i',j')\}$$

ce qui autorise au plus une coordonnée nulle.

Une trajectoire de (X^t) est typiquement constituée de segments parallèles à la "diagonale" (droite d'équation $y_{11} = \dots = y_{nn}$ dans $(\mathbb{R}^+)^{n^2}$), séparés par des sauts chaque fois qu'elle rencontre une face (d'équation $y_{ij} = 0$). Par exemple si le réseau est simplement constitué de deux neurones) on obtient une trajectoire de la forme indiquée figure 4

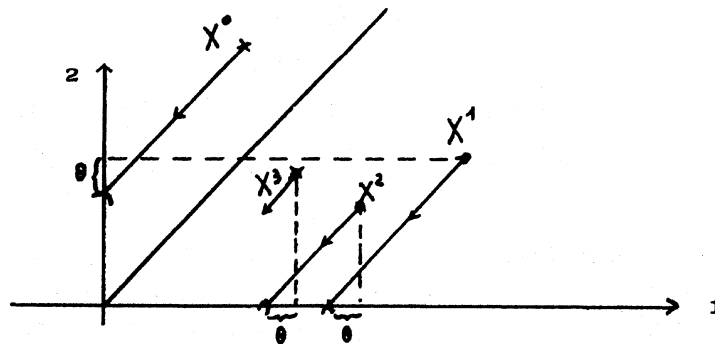


Figure 4

Une trajectoire d'une coordonnée X_{ij}^t est du type :

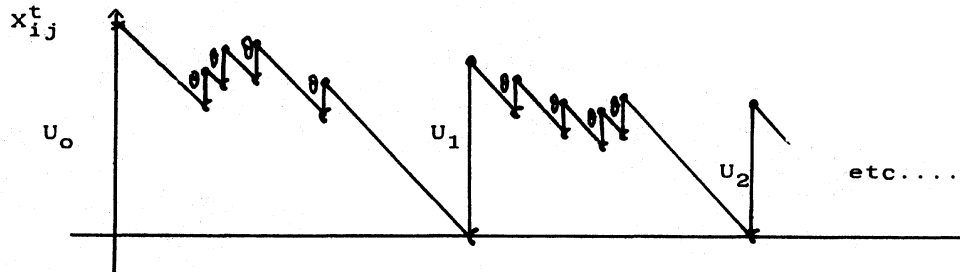


Figure 5

- Remarquons que les coordonnées X_{ij}^t ne sont pas individuellement markoviennes.

Dans la suite, on considère toujours (sauf mention expresse) le modèle continu (M), où θ est donc un réel positif représentant le délai dû à l'inhibition, et \mathcal{F} est une loi continue de support \mathbb{R}^+ admettant une densité $f > 0$ par rapport à la mesure de Lebesgue.

On suppose que \mathcal{F} est d'espérance et de variance finies.

2. SIMULATIONS.

Les simulations (voir appendices) du modèle discret (i) défini au II.1 montrent expérimentalement

- l'établissement d'un régime stationnaire correspondant à l'ergodicité du processus (X^t) , lorsque θ est plus petit qu'un certain seuil θ_0 , dépendant du système de voisinage.

- au contraire, si $\theta > \theta_0$, l'installation d'un régime non stationnaire, approximativement en "quinconce", ou en "bandes" suivant la forme des voisinages \mathcal{V} . On obtient des moirures, alternances de bandes de neurones actifs, déchargeant suivant la loi \mathcal{F} , avec une fréquence moyenne à peu près constante (inverse de $E\mathcal{F}$), et de bandes de neurones complètement inhibés, la largeur de ces bandes dépendant de la taille des voisinages.

Dans les paragraphes suivants, on donne les démonstrations de ces résultats : au paragraphe III, on montre que le processus (X^t) est irréductible et aperiodique et que lorsque $\theta < \theta_0$, les temps de retour à 0 en chaque site sont des v.a. d'espérance finie.

Dans le paragraphe IV, on démontre que lorsque $\theta < \theta_0$, le processus (X^t) est récurrent positif et ergodique, on étudie son régime stationnaire et on calcule alors l'intervalle inter-spike moyen de chaque neurone.

On étudie le cas $\theta > \theta_0$ au paragraphe V pour un système de deux neurones et on montre alors que p.s. l'une des deux composantes tend vers $+\infty$, et que l'autre converge en loi vers sa loi propre (sans inhibition, c'est-à-dire le processus de renouvellement de loi \mathcal{F}).

Enfin dans le § VI, on donne des conditions de stabilité de certaines configurations divergentes lorsque $\theta > \theta_0$.

Dans les Appendices 1, 2, 3, 4 etc., sont présentés de nombreux exemples à 2 ou n^2 neurones dans les cas convergent et divergent.

Dans ce qui suit on oublie la structure géométrique du modèle. Les sites sont nommés i , et sont en nombre fini N .

Rappelons que les variables U_i sont supposées indépendantes de même loi \mathcal{F} , à densité strictement positive sur \mathbb{R}^+ , et que $E(U)$ notée $E(\mathcal{F})$ et $E(U^2)$ notée $E(\mathcal{F}^2)$ sont finis.

III. CHAINE EXTRAITE. IRREDUCTIBILITE. APERIODICITE. TEMPS DE RETOUR.

1. CHAINE EXTRAITE.

Au processus de Markov défini par le modèle (M) au (II.1.2), on associe la suite croissante des temps d'arrêt T_1, T_2, T_3, \dots instants des décharges successives, instants des sauts du processus (X^t) .

Plus précisément, avec $T_0 = 0$, on pose

$$T_{\ell+1} = \inf\{t > T_\ell, \exists i_0, X_{i_0}^{t-} = 0\}. \quad (\text{III.1.1})$$

Remarque. Les T_ℓ sont tous finis p.s. et même $\overline{\lim} \frac{T_\ell}{\ell} < E(\mathcal{F})$.
(Voir ci-dessous juste avant le lemme III.3.1.).

Pour $x \in \mathcal{F}$, on note $\mathcal{J}(x) = \min_i(x_i)$ et $i_0(x)$ l'indice (unique) qui réalise ce minimum.

Alors après le temps T_ℓ , le numéro du neurone qui sera le premier à décharger est $i_0(X^{T_\ell})$, noté i^ℓ , et $T_{\ell+1} = T_\ell + \mathcal{J}(X^{T_\ell}) = T_\ell + X_{i_0}^{T_\ell}$.

En posant $Z(x) = x - \mathcal{J}(x) \mathbf{1}$, et en définissant comme équivalents deux vecteurs qui ont le même $Z(x)$, on voit que pour $T_\ell \leq t < T_{\ell+1}$, $X^t = X^{T_\ell} - (t - T_\ell) \mathbf{1}$ et que X^t reste équivalent à X^{T_ℓ} : on a $Z(X^t) = Z(X^{T_\ell}) = Z(X^{T_{\ell+1}^-}) = X^{T_{\ell+1}^-}$, noté Z^ℓ dans la suite.

On associe au temps $T_{\ell+1}$, l'état du système juste après la décharge et on a donc

$$X_i^{T_{\ell+1}} = Z_i^\ell + U_{\ell+1} \mathbf{1}_{i=i^\ell} + \theta \mathbf{1}_{i \in V(i^\ell)},$$

où $V(i)$ est l'ensemble des voisins de i , sans y inclure i , inhibés par i .

Les processus (X^{T_ℓ}) et (Z^ℓ) sont tous deux des chaînes de Markov, à temps discret. La chaîne (Z^ℓ) a pour ensemble d'états $\mathcal{F}_0 = \{x \in \mathcal{F} / \exists i, x_i = 0\}$. Sa transition est notée Q .

Quant à (X^{T_ℓ}) , l'ensemble de ses états est

$\mathcal{F}_p = \{x \in \mathcal{F} / \exists i, \forall j \in V(i), x_j > \theta\}$, puisque les voisins d'un certain i , (celui qui vient de sauter) sont simultanément incrémentés de θ à chaque décharge.

Bien sûr, la chaîne (X^{T_ℓ}) contient toute l'information du processus (X^t) , puisque la partie de trajectoire issue de (X^{T_ℓ}) jusqu'à l'instant $T_{\ell+1}$ du prochain saut est déterministe.

La transition de cette chaîne de Markov (X^{T_ℓ}) , homogène dans le temps est P donnée par

$$\begin{aligned} & P(x_1, x_2, \dots, x_N, dy_1, dy_2, \dots, dy_N) \\ &= f(y_{i_0}) dy_{i_0} \otimes \left[\prod_{i \in \mathcal{V}(i_0)} \delta(y_i = x_i - x_{i_0} + \theta) \right] \\ & \quad \otimes \left[\prod_{i \in \mathcal{V}(i_0) \cup \{(i_0)\}} \delta(y_i = x_i - x_{i_0}) \right] \end{aligned} \quad (\text{III.1.2})$$

lorsque $i_0 = i_0(x)$.

Ses itérées sont $p^2, p^3, \dots, p^m, \dots$, la loi de (X^{T_ℓ}) correspondant à la loi initiale μ est notée \mathbb{P}_μ (ou \mathbb{P}) lorsque cela ne prête pas à confusion. On note aussi \mathbb{P}_μ^k la loi de X^{T_k} sous μ (notations usuelles μ^k). On note aussi \mathbb{P}_μ la loi de (X^t) .

2. IRREDUCTIBILITE, APERIODICITE.

La loi \mathcal{F} admettant une densité $f > 0$ sur \mathbb{R}^+ , on peut clairement construire un ensemble de probabilité strictement positive de trajectoires de X^t menant un x quelconque de \mathcal{F} à toute boule ouverte $B_\varepsilon(y)$ de centre y et de rayon $\varepsilon > 0$. Cette construction se fait par exemple comme illustrée ci-dessous *figure 6*, pour 2 neurones

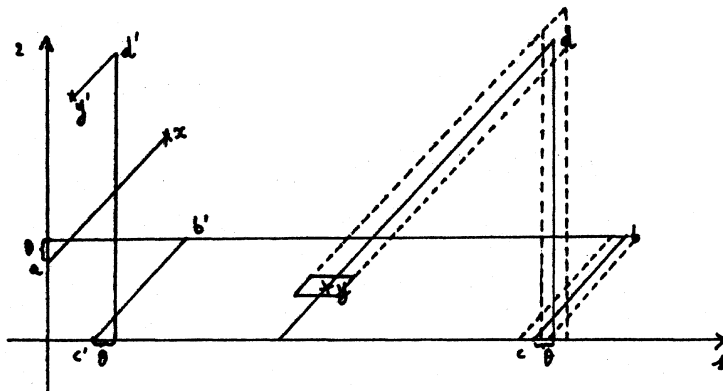


Figure 6

Précisons le cas général. Posons $Z^0 = X^0 - \mathcal{J}(X^0)$. Le premier saut aura lieu au temps $T_1 = \mathcal{J}(X^0)$ et au site $i_1 = i_0(X^0)$ tel que $Z_{i_0}^0 = 0$. On a $X^{T_1} = Z^0 + U_1 \mathbf{1}_{i_1} + \theta \mathbf{1}_{U(i_1)}$ (notation évidente).

On peut choisir le tirage de U_1 de sorte que l'indice du 2^e saut soit $i_2 \neq i_1$: en posant $\mathcal{M}(x) = \max_i(x_i) - \mathcal{J}(x)$, on voit que pour tout $j \neq i_1$, $X_j^{T_1} < \mathcal{M}(Z^0) + \theta$ et qu'il suffit que $U_1 > \mathcal{M}(Z^0) + \theta$.

Alors au temps T_2 , on aura

$$X_{i_1}^{T_2} > U_1 - \mathcal{M}(Z^0) - \theta, \quad X_{i_2}^{T_2} = U_2,$$

$$X_j^{T_2} < \mathcal{M}(Z^0) + 2\theta \quad \text{et ainsi de suite.}$$

Si donc à chaque tirage, l'innovation U_j vérifie

$$U_j > N \mathcal{M}(Z^0) + (1 + 2 + \dots + N) \theta = N \mathcal{M}(Z^0) + \frac{N(N+1)}{2} \theta$$

on est assuré de définir une permutation σ de l'ensemble des entiers $\{1, \dots, N\}$, ne dépendant que de Z^0 et définissant l'ordre des sites qui sautent. Alors au temps T_N , chaque coordonnée $X_i^{T_N}$ s'écrit $U_{\sigma(i)} + \varphi_i(Z^0)$, où $\varphi_i(Z^0)$ est déterministe, ne dépend pas des tirages des U_i et est entièrement déterminée par Z^0 .

On note U_σ le vecteur $U_{\sigma(i)}$, $\varphi(Z^0)$ le vecteur $(\varphi_i(Z^0))$, $\|x\|$ la norme $\sup_i |x_i|$ et $U > a$ pour $(U_i > a)$ pour tout i .

On peut résumer ce qui précède par le

LEMME III.2.1. Evolution déterministe.

Soit $M(Z^0) = N \mathcal{M}(Z^0) + \frac{N(N+1)}{2} \theta$. Il existe une permutation déterministe $\sigma(Z^0)$ de $[1, N]$, et un vecteur déterministe $\varphi(Z^0)$ vérifiant $\|\varphi(Z^0)\| \leq M(Z^0)$ tels que, conditionnellement aux tirages $U = (U_1, U_2, \dots, U_N) > M(Z^0)$, on ait $X^{T_N} = U_\sigma + \varphi(Z^0)$ avec $T_N \leq M(Z^0)$, (on a alors $T_{N+1} - T_N \leq M(Z^0) + \|U\|$).

Comme $\mathcal{F}_0 = \bigoplus_j \mathcal{F} \cap \{x_j = 0\}$, on prend sur $\mathcal{F} \cap \{x_j = 0\}$ la mesure de Lebesgue sur les coordonnées différentes de j et δ_0 sur la j -ième coordonnée, et on définit ainsi (par somme directe) la mesure Λ sur \mathcal{F}_0 .

LEMME III.2.2. Minoration.

Si K est un compact de $\overline{\mathcal{P}_0}$, il existe une constante $\gamma_K > 0$ telle que $Q^N(z, dz') > \gamma_K \mathbf{1}_K(z) \mathbf{1}_K(z') \Lambda(dz')$.

Démonstration. Il suffit de démontrer ceci pour $z'_1 = 0$.

Soit $z' = (0, z'_2, \dots, z'_N)$. D'après le lemme précédent, en N coups, on est en $(U_1 + \varphi_1(z), \dots, U_N + \varphi_N(z))$ et on a le choix des U_i dès que $U_i > M(Z^0)$. On peut donc minorer $Q^N(z, dz')$ par

$$\mathbf{1}_{z'_1=0} \times \Lambda(dz') \times \int_A^B f(u) \prod_{j=1}^N f(u + z'_j - \varphi_j(z) + \varphi_1(z)) du,$$

en choisissant A et B pour réaliser les conditions du lemme III.2.1.

Comme $z \in K$ compact, on peut majorer $M(z)$ par une constante finie $M(K)$. Remarquons qu'il faut $U_{\sigma(j)} = z'_{\sigma(j)} + U_1 - \varphi_{\sigma(j)}(z) + \varphi_1(z)$ pour "descendre" en z' , et que de plus, comme

$|\varphi_{\sigma(j)}(z)|, |\varphi_1(z)| < M(z) \leq M(K)$, il suffit de prendre $B > A > 3M(K)$ pour avoir $U_{\sigma(j)} > M(K)$ pour tout j (conditions du lemme) et pour que tous les arguments de l'expression intégrale soient positifs.

Quitte à remplacer f par $f \wedge 1$, on peut supposer (dans la minoration de l'expression intégrale), que f est intégrable, strictement positive et bornée. L'expression minorante, étant continue alors des $(z'_j - \varphi_j(z) + \varphi_1(z))$ est donc minorée par $C_K > 0$. \square

De plus, en faisant un "coup pour rien", on peut aussi arriver en n'importe quel z' à partir de z , en $N+1$ transitions au lieu de N . Illustrons ceci, dans le cas de deux neurones :

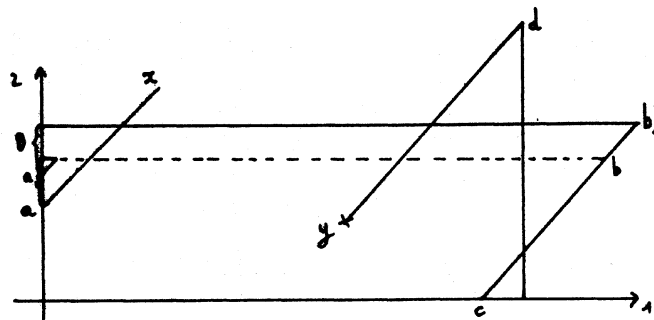


Figure 7

Le lemme III.2.2. entraîne l'irréductibilité de la chaîne Z^{ℓ} (avec Λ comme mesure de référence), et la dernière remarque entraîne l'apériodicité.

En fait, pour tout $c \in]0,1[$, tout x initial

$G^c(x, dy) = \sum c^n Q^n(x, dy)$ est équivalent à Λ , et on obtient donc

COROLLAIRE III.2.3. La chaîne (Z^{ℓ}) est irréductible et apériodique.

Pour que la chaîne soit récurrente (pour tout x , tout E tel que $\Lambda(E) > 0$, partant de x on repasse indéfiniment dans E) il suffit de montrer qu'il existe un x et un compact K ($\Lambda(K) > 0$) tel que

i) Soit, partant de x , on repasse infiniment dans K

ii) Soit $G(x, K) = \infty$.

(voir par exemple [REVUZ]).

On peut en déduire les mêmes propriétés pour la chaîne (X^{ℓ}) .

Soit x et x' appartenant à un compact K de \mathcal{F}_p . Comme la loi conditionnelle après le premier saut ne dépend que de $Z(x)$, en N coups, la chaîne (Z^{ℓ}) peut passer de $z = Z(x)$ à z' pour z' susceptible de conduire à x' de \mathcal{F}_p : si $x' = (x'_1, \dots, x'_N) \in \mathcal{F}_p$, il existe i tel que $x'_j > \theta$ pour $j \in V(i)$ et $z' = (z'_1, \dots, z'_N)$ peut conduire à x' dès lors que $z'_i = 0$, $z'_j = x'_j - \theta$ pour $j \in V(i)$ et $z'_{j'} = x'_{j'}$ pour $j' \neq i$ et $j' \notin V(i)$, en tirant $U = x'_i$.

Alors en $(N+1)$ étapes, on peut aller de x à x' , avec probabilité minorée sur K .

Bien sûr, comme pour (Z^{ℓ}) , on peut aussi faire "un coup pour rien" et aller de x à x' en $N+2$ étapes. D'où

COROLLAIRE III.2.4. La chaîne (X^{ℓ}) est irréductible, apériodique sur \mathcal{F}_p (muni de la mesure de Lebesgue λ). S'il existe un compact de potentiel infini, elle est récurrente dans tous les ensembles λ -positifs de \mathcal{F}_p . De même (X^t) est irréductible, apériodique sur \mathcal{F} , et la récurrence de (X^{ℓ}) entraîne celle de (X^t) .

En effet, comme la chaîne (X^{ℓ}) contient toute l'information du processus (X^t) , (X^t) a les mêmes propriétés que (X^{ℓ}) sur son support \mathcal{F} .

3. TEMPS DE RETOUR A ZERO.

Intuitivement, la récurrence du processus correspond au fait que les neurones déchargent tous alternativement, c'est-à-dire que la loi de la v.a. $i^{\ell} = i_0(X_i^{\ell})$ du site où est réalisé $\min_i(X_i^{\ell})$ charge indéfiniment tous les entiers de 1 à N. La récurrence positive est le fait que la "bascule" a lieu en un temps intégrable. Au contraire, lorsqu'un neurone i ne décharge plus, il ne "descend" jamais assez pour être à son tour minimum, il est surchargé et $X_i^t \rightarrow +\infty$.

Soit i un site donné, d'état $x = X_i^0$ en $t = 0$, et notons τ la v.a. temps de retour de ce site à 0.

Pour $t < \tau$, $t < t + X_i^t = x + \theta \sum_{j \in \mathcal{V}(i)} \sum_m \mathbf{1}_{(T_{j,m} < t)}$ où $T_{j,m}$ est le m -ième temps de retour du site j à l'état 0.

Notons $N_j(t)$ le nombre de retours à 0 du site j avant t :

$$N_j(t) = \sum_m \mathbf{1}_{T_{j,m} < t}.$$

Dans le modèle initial, on s'est donné une unique suite U_k iid (\mathcal{F}). On aurait pu tout aussi bien affecter à chaque site j une suite $U_m^{(j)}$ iid (\mathcal{F}), ces suites étant indépendantes, réinitialisant l'état j à chacun de ses passages en 0, de sorte que, pour $m > 1$,

$$T_{j,(m+1)} - T_{j,m} > U_m^{(j)}. \quad (*)$$

Soit $S_m^{(j)} = U_1^{(j)} + \dots + U_m^{(j)}$ et ν_j le compteur associé : $\{\nu_j(t) > m+1\} = \{S_m^{(j)} < t\}$. (*) donne $N_j(t) < \nu_j(t)$.

On note que ν_j ne dépend pas de la loi initiale, et que

$$\frac{\nu_j(t)}{t} \xrightarrow{\text{p.s.}} 1/E(\mathcal{F}).$$

Remarque. Soit $N(t)$ le "compteur général" $(N(t) > \ell+1) = \{T_{\ell} < t\}$. Comme en T_{ℓ} , $\mathcal{J}(X_i^{\ell}) < U_{\ell}$, il est clair que $N(t) > \nu(t)$, compteur associé à la marche des U , soit $\liminf \frac{N(t)}{t} > 1/E(\mathcal{F})$ ou encore $\limsup \frac{T_{\ell}}{\ell} < E(\mathcal{F})$.

Comme $(t < \tau)$ implique $\left\{ 1 < \frac{x}{t} + \theta \sum_{j \in \mathcal{V}(i)} \frac{\nu_j(t)}{t} \right\}$
 $(\tau = \infty)$ est impossible si $\theta |\mathcal{V}(i)| / E(\mathcal{F}) < 1$, soit

LEMME III.3.1. Si le site i vérifie $\theta |V(i)| < E(\mathcal{F})$ alors, pour toute loi initiale, le temps de retour à 0 du site i est fini p.s.

Démonstration. Notons v le nombre $|V(i)|$ de sites voisins de i . La même hypothèse $\theta v < E(\mathcal{F})$ va entraîner l'intégrabilité de τ (avec majoration exponentielle). Notons (abusivement) P_x la probabilité sous une loi initiale donnant au site i la valeur initiale x , et ν la loi commune des ν_j .

Comme $P_x[\tau > t+x] \leq P[\nu \sum_{j \in V(i)} \nu_j(t+x) > t]$, on obtient comme majorant $\nu P[\nu(t+x) > t/v\theta]$. Si $[z]$ désigne "partie entière de z ", il vient $P_x[\tau > t+x] \leq \nu P[S_{[t/v\theta]-1} \leq t+x]$. Mais la variable (centrée) $U - E(\mathcal{F})$ est bornée à gauche et satisfait donc une inégalité de Cramer du type : pour tout $\alpha > 0$, il existe $\gamma > 0$ tel que

$$P[S_m \leq m[E(\mathcal{F}) - \alpha]] \leq \exp - m\gamma .$$

On en déduit que :

Soit $E(\mathcal{F}) = v\theta(1+\rho)$ avec $\rho > 0$. On pose $\alpha = \rho v\theta/2$, et on note γ le nombre associé. Alors, pour $m > 2(2+x/v\theta)/\rho$, si $(m+1)v\theta \leq t < (m+2)v\theta$, $P_x[\tau > t+x] \leq \nu \exp(-m\gamma)$. On peut donc énoncer

LEMME III.3.2. Lorsque $\theta < \theta_0 = E(\mathcal{F})/|V(i)|$ il existe des constantes A et B ne dépendant que de la loi \mathcal{F} et de θ telles que pour toute loi initiale μ donnant x comme état initial, $E_\mu(\tau) \leq A+Bx$ et $E_\mu(\tau^2) \leq A+Bx^2$.

La majoration $N_i(t) \leq 1 + \nu_i(t)$ montre que

LEMME III.3.3. Il existe t_0 et $\gamma > 0$ ne dépendant que de la loi \mathcal{F} tels que pour $t > t_0$,

$$P(N_i(t) > \left[\frac{2t}{E(\mathcal{F})} \right] - 1) \leq e^{-\gamma t}.$$

IV. REGIME STATIONNAIRE

$$\theta < \theta_0 .$$

1. RECURRENCE POSITIVE POUR

$$\theta < \theta_0 = E(\mathcal{F})/|U| .$$

Nous prenons comme loi initiale $\nu = \mathcal{F}^{\otimes N}$. On a

LEMME IV.1.1. Si $\theta < \theta_0$, la suite des probabilités

$$\mu_n = \frac{1}{n} \sum_{k=1}^n \mathbb{P}_\nu^k \quad \text{est tendue.}$$

On en déduit immédiatement que toute probabilité adhérente à cette suite est une probabilité invariante de la chaîne, donc unique, équivalente à la mesure λ , et que la suite des itérées \mathbb{P}^k tend vers elle. La chaîne est bien récurrente positive.

Preuve de IV.1.1. Fixons un site i , et soit $T_{i,k}$ la suite des temps de retour à 0 du site i , (pour le processus X^t) avec la convention $T_{i,0} = 0$. Si $T_{i,k} \leq s < T_{i,k+1}$, $X_i^s \leq T_{i,k+1} - s$ donc

$$\int_{T_{i,k}}^{T_{i,k+1}} X_i^s ds \leq \frac{1}{2} (T_{i,k+1} - T_{i,k})^2 . \quad \text{Mais, y compris pour } k = 0,$$

$X_i^{T_{i,k}}$ est de loi \mathcal{F} , de sorte qu'en utilisant le lemme III.3.2, où on pose $C = A + B E(\mathcal{F}^2)$ il vient

$$E_\nu \int_0^{T_{i,k}} X_i^s ds \leq C k/2 . \quad (\text{IV.1.2})$$

Notons \mathcal{E}_s la variable $\sum_i X_i^s$. Si T_ℓ désigne la suite des instants de retour à 0 (tous sites confondus) comme $T_{i,\ell} \geq T_\ell$,

$$E_\nu \int_0^{T_\ell} \mathcal{E}_s ds \leq C N \ell/2 . \quad (\text{IV.1.3})$$

Notons X_i^k la quantité $X_i^{T_k}$, $\mathcal{E}_k = \sum_i X_i^k$, et \mathcal{J}_k le minimum des X_i^k . L'évolution de la chaîne entre passages en zéro donne :

$$T_{k+1} = T_k + \mathcal{J}_k , \quad X_i^{T_{k+1}+s} = X_i^k - s \quad \text{pour } 0 \leq s < \mathcal{J}_k ,$$

donc $\int_{T_k}^{T_{k+1}} \mathcal{E}_s ds \geq (\mathcal{E}_k - N \mathcal{J}_k/2) \mathcal{J}_k \geq \mathcal{E}_k \mathcal{J}_k/2$.

Notons que $\mathcal{E}_{k+1} = \mathcal{E}_k - N \mathcal{J}_k + |V|\theta + U_{k+1}$, où (U_j) désigne la suite des innovations, indépendantes et de loi \mathcal{F} , de sorte que, pour $\alpha = \theta/(N+2)$, on a

$$\text{si } \mathcal{J}_k < \alpha, \quad \mathcal{E}_{k+1} > \mathcal{E}_k. \quad (\text{IV.1.4})$$

Supposons (1er cas) que $\text{Max}(\mathcal{J}_k, \mathcal{J}_{k+1}, \dots, \mathcal{J}_{k+N+2}) > \alpha$. Les formules précédentes montrent que $\int_{T_k}^{T_{k+N+3}} \mathcal{E}_s ds > \alpha \mathcal{E}_k/2$.

Mais parmi les instants $T_{k+1} \dots T_{k+N+2}$, l'un des sites a sauté au moins deux fois, de sorte que si on pose $W_{k+1} = \text{Min}\{U_{k+1}, \dots, U_{k+N+2}\}$, $T_{k+N+2} - T_k > W_{k+1}$.

Dans le second cas, ($\text{Max}(\mathcal{J}_k, \dots, \mathcal{J}_{k+N+2}) < \alpha$) on sait que $\mathcal{E}_{k+j} > \mathcal{E}_k$ ($1 < j < N+2$) donc $\int_{T_k}^{T_{k+N+3}} \mathcal{E}_s ds > \mathcal{J}_k W_{k+1}/2$.

Posons $\tilde{\alpha}_{k+1} = \alpha \wedge W_{k+1}$. En combinant les deux minoration, il vient

$$\int_{T_k}^{T_{k+N+3}} \mathcal{E}_s ds > \tilde{\alpha}_{k+1} \mathcal{E}_k/2. \quad (\text{IV.1.5})$$

Notons que par construction, $\tilde{\alpha}_{k+1}$ est indépendant de l'évolution du processus jusqu'à T_k , et d'espérance $2\beta > 0$. Il vient donc :

Il existe $\beta > 0$ telle que, pour toute loi initiale

$$E \int_{T_k}^{T_{k+N+3}} \mathcal{E}_s ds > \beta E \mathcal{E}_k \quad (k > 0). \quad (\text{IV.1.6})$$

Aussi bien, pour toute loi initiale

$$(N+3)E \int_{T_1}^{T_{1+(j+1)(N+3)}} \mathcal{E}_s ds > \beta \sum_{k=1}^{1+j(N+3)} E \mathcal{E}_k.$$

Combinant ceci avec IV.1.3, reprenant la définition de μ_n et notant \mathcal{E} la variable $\sum X_1$, il vient

$$\int \mathcal{E} d\mu_{1+j(N+3)} < \left(\frac{C}{2\beta}\right) \left(\frac{(j+1)(N+3)+1}{1+j(N+3)}\right) N(N+3), \quad (\text{IV.1.7})$$

qui prouve, C et N étant finis et $\beta > 0$ que

$$\int \mathcal{E} d\mu_n = o(1).$$

Prenant pour K_T le compact $(\mathcal{E} \leq T)$, il vient $\mu_n(K_T) = o(1/T)$ qui prouve la tension des mesures. \square

La chaîne étant récurrente dans tous les λ positifs de son espace d'états, on a déjà noté que le processus X^t l'est aussi dans tous les λ positifs de son système d'états. Bien entendu, il possède lui aussi une probabilité invariante, équivalente à λ : comme dans le début de ce paragraphe, il suffit de montrer que, pour la même mesure initiale, les

$$\mu_t = \frac{1}{t} \int_0^t \nu P_s ds \text{ sont tendues quand } t \rightarrow \infty. \quad (\text{IV.1.8})$$

Soit A_t l'événement : {pour tout site i , $1 + N_i(t) \leq k_t$ } où k_t est l'entier $[2t/E(\mathcal{F})]$. Soit $Z_t = 1_{A_t} \int_0^t \xi_s ds$.

Sur A_t , $T_{i,k_t} > t$ et d'après (IV.1.2),

$$E_\nu(Z_t) \leq C N k_t / 2 \leq C N t / E[\mathcal{F}]. \quad (\text{IV.1.9})$$

Soit $C K_T$ le complémentaire du compact ($\mathcal{E} \leq T$).

L'inégalité de Markov montre que

$$\mu_t [C K_T] \leq P_\nu(A_t) + C N / (T E[\mathcal{F}]).$$

Mais (III.3.3) montre que, quand t tend vers l'infini, $P_\nu(A_t)$ tend vers 0 et (IV.1.8) est prouvé.

Nous pouvons donc résumer

PROPOSITION IV.1.10. Si $\theta < E(\mathcal{F}) / |V|$ et $E(\mathcal{F}^2) < \infty$, les processus (X^t) et (X^{T_k}) sont ergodiques, irréductibles, apériodiques, récurrents positifs.

On remarquera qu'on n'a pas utilisé dans la minoration le fait qu'un i influence ou soit influencé par ses voisins.

En fait, la seule hypothèse est que si $\nu = \sup |V(i)|$,

$$\nu \theta < E[\mathcal{F}]$$

car cette hypothèse assure que tous les $T_{i,k}$

sont finis p.s. !

2. LOI INVARIANTE.

La loi invariante m de la chaîne X^T vérifie

$$m(dy) = \int_{x \in \mathcal{P}} m(dx) P(x, dy),$$

et par (III.1.1) (en posant $A_i = \{x \in \mathcal{P} / x_i = \min(x_j)\}$)

$$\begin{aligned} m(dy) &= \sum_{i_0} \int_{x \in A_{i_0}} m(dx) f(y_{i_0}) dy_{i_0} \otimes \left[\sum_{i \in V(i_0)} \delta(y_i = x_i - x_{i_0} + \theta) \right] \\ &\quad \otimes \left[\sum_{i \notin V(i_0) \cup \{i_0\}} \delta(y_i = x_i - x_{i_0}) \right] \\ &= \sum_{i_0} f(y_{i_0}) dy_{i_0} \int_{x_{i_0}} m(dx_1, dx_2, \dots, dx_N) \end{aligned}$$

$$\text{avec} \quad \begin{aligned} x_i &= x_{i_0} + y_i - \theta & \text{si } i \in V(i_0) \\ x_i &= x_{i_0} + y_i & \text{si } i \notin V(i_0) \cup \{i_0\}. \end{aligned}$$

Cette équation, discrétisée, permet de déterminer la mesure stationnaire dans des cas particuliers.

On peut en déduire les équations vérifiées par les lois marginales

$$m(dy_{i'}) = \int_{y_i} m(dy_1, \dots, dy_N) \quad \text{pour } i' \neq i.$$

On a bien sûr

$$y_{i'} > \theta \quad \text{si } i' \in V(i)$$

et

$$y_{i'} > 0 \quad \text{si } i' \notin V(i) \cup \{i\}$$

$$m(dy_i) = f(y_i) dy_i m(A_i) + \sum_{i'} m(A_{i'} \cap \partial B_i) \times \delta_{(y_i > \theta)}$$

où ∂B_i est défini par les inégalités

$$x_{i'} + y_i - \theta < x_{i''} < x_{i'} + y_i - \theta + dy_i$$

pour $i'' \in V(i)$.

La symétrie du comportement de tous les neurones implique que

$$m(A(i)) = \frac{1}{N}.$$

On a donc :

$$E(X_i) = \int_{y_i} m(dy_i) y_i = \frac{1}{N} E(\mathcal{F}) + \int_{y_i > \theta} \sum_{i'} m(A_{i'} \cap \partial B_i) \quad (\text{IV.2.1})$$

Quant au processus (X^t) , sa loi invariante, notée aussi m , a une densité g par rapport à la mesure de Lebesgue sur \mathcal{P} et on peut écrire une équation de bilan, vérifiée par la fonction g :

$$\sum_i \frac{\partial g}{\partial x_i} + \sum_{i_0} g(x'_1, x'_2, \dots, x'_N) f(x_{i_0}) = 0 \quad (\text{IV.2.2})$$

avec $x'_{i_0} = 0$, $x'_i = x_i - \theta$ pour $i \in V(i_0)$

et $x'_i = x_i$ pour $i \notin V(i_0) \cup \{i_0\}$.

Remarque.

On observe des auto-corrélations peu significatives de ces intervalles, ce qui s'explique par le fait qu'à chaque décharge, la variable U est indépendante de l'état du réseau. Quant aux inter-corrélations des activités de deux neurones, mesurées par exemple entre les nombres de décharges des deux neurones pendant des fenêtres successives de temps, elles sont faibles quand θ est petit, et augmentent avec θ . Voir les Appendices 1 (2 neurones, dernière colonne) et 3 (n^2 neurones).

3. ETUDE DE L'INTERVALLE INTER-SPIKE D'UN NEURONE DANS UN RESEAU DE DEUX NEURONES QUAND

$$\theta < \theta_0.$$

On considère le processus (X^t) dans le cas de deux neurones, notés 1 et 2, et on s'intéresse aux intervalles inter-spikes du neurone 2. Dans ce cas $\theta_0 = E(\mathcal{F})$ puisque $|V| = 1$.

Supposons qu'à l'instant t_0^- , l'état du système est $(x_0, 0)$.

On conditionne par $X^{t_0^-} = (x_0, 0)$.

$$\text{Etat :} \quad \begin{array}{c|c} t_0^- & t_0^+ \\ \hline (x_0, 0) & (x_0 + \theta, u_0) \end{array}$$

i) Si $u_0 < x_0 + \theta$, la prochaine décharge de 2 a lieu à $t_0 + u_0$

$$\begin{array}{c|c} t_1^- = t_0^- + u_0 & t_1^+ \\ \hline (x_0 + \theta - u_0, 0) & (x_0 - u_0 + 2\theta, u_1) \end{array}$$

Sinon, la prochaine décharge est du neurone 1 et a lieu à $t_0 + x_0 + \theta$ en $(0, u_0 - x_0 - \theta)$ et $(u_1, u_0 - x_0)$ après le saut.

ii) Si $u_1 > u_0 - x_0$, la prochaine décharge de 2 a lieu à $t_0 + x_0 + \theta + u_0 - x_0 = t_0 + \theta + u_0 = t_2$ avec

$$\begin{array}{c|c} t_2^- & t_2^+ \\ \hline (u_1 - u_0 + x_0, 0) & (u_1 - u_0 + x_0 + \theta, u_2) \end{array}$$

Sinon, la décharge suivante est celle du neurone 1, et a lieu au temps $t_0 + x_0 + \theta + u_1$ en $(0, u_0 - x_0 - u_1)$ et $(u_2, u_0 - x_0 - u_1 + \theta)$ après le saut.

iii) Si $u_2 > u_0 - x_0 - u_1 + \theta$, la prochaine décharge de 2 a lieu à $t_0 + x_0 + \theta + u_1 + u_0 - x_0 - u_1 + \theta = t_0 + u_0 + 2\theta = t_3$ avec

$$\begin{array}{c|c} t_3^- & t_3^+ \\ \hline (u_2 - u_0 + x_0 + u_1 - \theta, 0) & (x_0 - u_0 + u_1 + u_2, u_3) \end{array} \quad \text{après le saut.}$$

Sinon, la prochaine décharge est celle du neurone 1, a lieu au temps $t_0 + x_0 + \theta + u_1 + u_2$ en $(0, u_0 - x_0 - u_1 - u_2 + \theta)$ et $(u_3, u_0 - x_0 - u_1 - u_2 + 2\theta)$ après le saut.

On voit que le temps τ séparant deux décharges consécutives du neurone 2 s'écrit

$$\tau = U_0 + K\theta \quad (\text{IV.3.1})$$

où U_0 est une v.a. de loi \mathcal{F} et K une v.a. entière.

On voit que

$$\mathbb{P}(K > 0) = \mathbb{P}(U_0 > x_0 + \theta)$$

$$\mathbb{P}(K > 1) = \mathbb{P}(U_0 > x_0 + \theta, U_1 < U_0 - x_0)$$

$$\begin{aligned} \mathbb{P}(K > k) &= \mathbb{P}(U_0 > x_0 + \theta, U_1 < U_0 - x_0, U_2 < U_0 - U_1 + \theta - x_0, \dots, U_k < U_0 - U_1 - \dots - U_{k-1} + (k-1)\theta - x_0) \\ &= \mathbb{P}(U_0 > x_0 + \theta, U_1 < U_0 - x_0, U_1 + U_2 < U_0 - x_0 + \theta, \dots, U_1 + U_2 + \dots + U_k < U_0 - x_0 + (k-1)\theta) \end{aligned}$$

Si l'on pose $U' = U - \theta$ ($EU' = EU - \theta$)

$$\begin{aligned} \mathbb{P}(K > k) &= \mathbb{P}(U_0 > x_0 + \theta, U'_1 < U'_0 - x_0, U'_1 + U'_2 < U'_0 - x_0, \dots, U'_1 + U'_2 + \dots + U'_k < U'_0 - x_0) \\ &= \mathbb{P}(U'_0 > x_0, U'_1 < U'_0 - x_0, \dots, U'_1 + U'_2 + \dots + U'_k < U'_0 - x_0). \end{aligned} \quad (\text{IV.3.2})$$

Or $U'_1 + \dots + U'_k$ est une marche aléatoire telle que $EU'_1 = EU - \theta = \theta_0 - \theta$ (2 neurones).

On retrouve le fait que $K \rightarrow +\infty$ lorsque $EU' < 0$, c'est-à-dire quand $\theta > \theta_0$.

Au contraire si $\theta < EU = \theta_0$, $EU' > 0$, et

$$\mathbb{P}(K = k) = \mathbb{P}((U'_0 > x_0) \text{ et } k \text{ est le premier entier strictement positif tel que } U'_1 + \dots + U'_k > U'_0 - x_0) \quad (\text{IV.3.3})$$

La loi de K est donc complètement déterminée, conditionnellement à la valeur x_0 du potentiel du neurone 1 à l'instant immédiatement antérieur à la décharge du neurone 2 (considéré).

On a

$$\begin{aligned} E_{x_0}(K) &= \sum_k \mathbb{P}(K > k) \\ &= \sum_k \int_{x_0 + \theta}^{+\infty} f(u_0) du_0 \int_0^{u_0 - x_0} f(u_1) du_1 \dots \int_0^{u_0 - x_0 - (k-2)\theta} f(u_{k-1}) du_{k-1} \int_0^{u_0 - x_0 - (k-1)\theta} f(u_k) du_k \end{aligned}$$

Lorsque $K = k$, l'intervalle τ prend la valeur $u_0 + k\theta$ avec $k > 0$ lorsque $u_0 > x_0 + \theta$.

En $t_{k+1}^- = t_0 + u_0 + k\theta$, on obtient l'état $(u_1 + \dots + u_k - u_0 + x_0 - (k-1)\theta, 0)$ et comme on est en régime stationnaire, on peut écrire

$$U_1 + \dots + U_k - U_0 + X_0 - (K-1)\theta \underset{\text{loi}}{\sim} X_0 \quad \text{ou encore}$$

$$U'_1 + \dots + U'_K - U'_0 + X_0 \underset{\text{loi}}{\sim} X_0$$

En particulier

$$E(U'_1 + \dots + U'_K) = EU'_0 = \theta_0 - \theta$$

et comme K est un temps d'arrêt,

$$(EK)(EU') = (EU'_0) \quad \text{d'où} \quad EK = 1$$

d'où

PROPOSITION IV.3.4. (Cas de 2 neurones)

En régime stationnaire, l'intervalle inter-spike moyen est $EU + \theta$ (quand $\theta < \theta_0 = EU$).

Ce résultat peut aussi se retrouver comme au paragraphe suivant.

4. INTERVALLE INTER SPIKE D'UN NEURONE QUAND

$\theta < \theta_0$. CAS GENERAL. EQUATION LIMITE.

Soit τ le temps de retour à 0 du site i :

$\tau = U + \theta \sum_{j \in \mathcal{V}(i)} N_j(\tau)$, et donc en initialisant par $\nu = (\mathcal{F})^{\otimes N}$,

$$T_{i,k} = U_0^i + \dots + U_{k-1}^i + \theta \sum_{j \in \mathcal{V}(i)} N_j(T_{i,k}).$$

Mais l'ergodicité du processus entraîne que $\frac{T_{i,k}}{k}$ converge p.s. vers

une limite s_i ($0 < s_i < \infty$). Comme $\frac{U_0^i + \dots + U_{k-1}^i}{k} \xrightarrow{\text{p.s.}} E(\mathcal{F})$

et que $N_i(T_{i,k}) = k$, il vient

$$E(\mathcal{F}) \frac{N_i(T_{i,k})}{T_{i,k}} + \theta \sum_{j \in \mathcal{V}(i)} \frac{N_j(T_{i,k})}{T_{i,k}} \xrightarrow{\text{p.s.}} 1.$$

Mais $\frac{N_i(t)}{t}$ converge p.s. ; en notant $\frac{1}{\tau_i}$ sa limite, les équations limites sont

$$1 = E(\mathcal{F}) \frac{1}{\tau_i} + \theta \sum_{j \in \mathcal{V}(i)} \frac{1}{\tau_j}, \quad \forall i \quad (\text{IV.4.1})$$

Ces convergences ont lieu en fait quelle que soit la mesure initiale. τ_i est donc l'espérance du temps de retour à 0 du site i , en régime stationnaire.

En particulier dans le cas homogène ($|\mathcal{V}(i)| = v$, ne dépendant pas de i),

$$\tau_i = E(\mathcal{F}) + \theta v \quad (\text{IV.4.2})$$

PROPOSITION IV.4.3. (Cas de N neurones, avec voisinages homogènes).

En régime stationnaire ($\theta < \theta_0 = E(\mathcal{F})/|V|$), l'intervalle inter-spike moyen est $E(\mathcal{F}) + |V|\theta$.

V. DIVERGENCE : CAS DE DEUX NEURONES.

Dans ce paragraphe $\theta > \theta_0 = E(\mathcal{F})$ ($N = 2$, $|V| = 1$).

1. ETUDE DE $(X_2 - X_1)^{T_\ell}$.

Rappelons qu'en toute situation les T_ℓ sont finis p.s.

Dans le cas de deux neurones, c'est le signe de $(X_2 - X_1)^{T_\ell}$ qui indique la nature de la décharge suivante : le neurone 1 (resp. 2) décharge si $(X_2 - X_1)^{T_\ell} > 0$ (resp. < 0).

Considérons tout d'abord le processus à temps discret $Z^\ell = (X_2 - X_1)^{T_\ell}$ (et non $X^{T_\ell} - \min(X^{T_\ell})$ comme au § III).

On a $Z^{\ell+1} = Z^\ell + \text{signe}(Z^\ell)(\theta - U_{\ell+1})$ où $U_{\ell+1}$ est une v.a. de loi \mathcal{F} , indépendante de Z^ℓ : en effet, la différence $(X_2 - X_1)^t$ est constante, égale à Z^ℓ , pour $T_\ell \leq t < T_{\ell+1}$ et à l'instant de la décharge $T_{\ell+1}$, on ajoute θ à la plus grande des deux composantes et une v.a. $U_{\ell+1}$ de loi \mathcal{F} à l'autre (qui est nulle à cet instant).

La situation est donc particulièrement simple : l'espérance du saut de Z^ℓ est du signe contraire à Z^ℓ si et seulement si $E(\mathcal{F}) - \theta > 0$

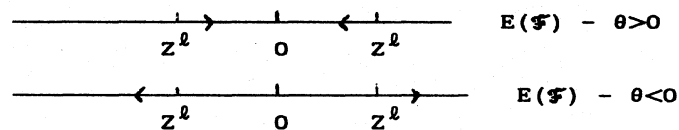


Figure 8.

On sait d'après le § IV que si $\theta < \theta_0$, (X^{T_ℓ}) et (X^t) sont des processus ergodiques.

Reste à étudier le comportement de (X^{T_ℓ}) et (X^t) lorsque $\theta > E(\mathcal{F})$.

THEOREME V.1.1. Cas de deux neurones, $\theta > E(\mathcal{F})$.

Lorsque $\theta > E(\mathcal{F})$:

- i) $|X_2 - X_1|^{T^l}$ et $|X_2 - X_1|^t$ tendent vers $+\infty$ ps, et plus précisément
 ii) Pour presque toute trajectoire ω , quand $t \rightarrow +\infty$, $(X_2 - X_1)^t$ converge vers $+\infty$ ou vers $-\infty$.

Démonstration

i) Notons $\mu = E(\mathcal{F})$, $\sigma^2 = \text{Var}(\mathcal{F})$. Pour $U \sim \mathcal{F}$, on pose $V = \theta - U$, $EV = \theta - \mu = \alpha > 0$, $\text{Var} V = \sigma^2$. On note encore $Z^l = (X_2 - X_1)^{T^l}$. On sait que $Z^{l+1} = Z^l + \text{sgn}(Z^l) (\theta - U_{l+1})$ avec U_{l+1} indépendante de Z^l , et de loi \mathcal{F} .

De $Z^{l+1} = Z^l + \text{sgn}(Z^l) \cdot V_{l+1}$, on obtient

$$\begin{aligned} |Z^{l+1}| &= |\text{sgn}(Z^l) (|Z^l| + V_{l+1})| \\ &= ||Z^l| + V_{l+1}| \end{aligned}$$

$|Z^l|$ est Z^l -mesurable. On note E^l l'espérance conditionnelle relativement à Z^l .

$$\begin{aligned} \text{Alors } E(|Z^{l+1}| / Z^1, \dots, Z^l) &= E^l(|Z^{l+1}|) = E^l(|Z^l| + V_{l+1}|) \\ &> |E^l(|Z^l| + V_{l+1})| = |Z^l| + \alpha > |Z^l| \end{aligned}$$

Donc $|Z^l|$ est une sous-martingale positive.

$$\text{Comme } E(|Z^{l+1}|) = E(E^l|Z^{l+1}|)$$

$$\text{on a } E(|Z^{l+1}|) > E(|Z^l|) + \alpha.$$

Donc la suite $E(|Z^l|) > l\alpha$, est croissante et tend vers $+\infty$.

D'autre part, la sous-martingale $(|Z^l|)$ se décompose (Décomposition de Doob) en la somme d'une martingale intégrable (M_l) et d'un processus croissant (A_l) , soit

$$|Z^l| = M_l + A_l.$$

Cette décomposition est unique si l'on prend $A_0 = 0$ et $M_0 = |Z^0|$.

Ces variables sont définies par

$$\begin{aligned} A_{\ell+1} - A_{\ell} &= \Delta A_{\ell} = E^{\ell} (|Z^{\ell+1}| - |Z^{\ell}|) \\ &= E^{\ell} (|Z^{\ell+1}|) - |Z^{\ell}| \\ &> \alpha, \end{aligned}$$

et $M_{\ell+1} - M_{\ell} = \Delta M_{\ell} = |Z^{\ell+1}| - E^{\ell} (|Z^{\ell+1}|)$.

On a donc $A_{\ell} > \ell\alpha$ et $A_{\ell} \xrightarrow[\text{p.s.}]{} +\infty$ quand $\ell \rightarrow +\infty$.

Reste à étudier le comportement de la martingale M_{ℓ} .

La sous-martingale M_{ℓ}^2 , intégrable puisque la loi \mathcal{F} est de carré intégrable, se décompose aussi en $M_{\ell}^2 = m_{\ell} + B_{\ell}$, où (m_{ℓ}) est une martingale intégrable et B_{ℓ} un processus croissant.

$$\begin{aligned} \text{On a } E^{\ell} (\Delta M_{\ell}^2) &= E^{\ell} ((|Z^{\ell+1}|)^2) - (E^{\ell} (|Z^{\ell+1}|))^2 \\ &= E^{\ell} (|Z^{\ell} + \text{sgn}(Z^{\ell}) V_{\ell+1}|^2) - (E^{\ell} |Z^{\ell} + \text{sgn}(Z^{\ell}) V_{\ell+1}|)^2 \\ &< E^{\ell} (|Z^{\ell} + \text{sgn}(Z^{\ell}) V_{\ell+1}|^2) - (E^{\ell} (Z^{\ell} + \text{sgn}(Z^{\ell}) V_{\ell+1}))^2 \end{aligned}$$

puisque $|EX| \leq E|X|$.

$$\begin{aligned} &= (Z^{\ell})^2 + 2 (\text{sgn } Z^{\ell}) \cdot Z^{\ell} \cdot E V_{\ell+1} + E [(V_{\ell+1})^2] \\ &\quad - (Z^{\ell})^2 - 2 (\text{sgn } Z^{\ell}) \cdot Z^{\ell} \cdot E V_{\ell+1} - E [(V_{\ell+1})^2] \\ &= \text{Var } V = \sigma^2. \end{aligned}$$

Comme $B_{\ell+1} - B_{\ell} = E^{\ell} (M_{\ell+1}^2) - M_{\ell}^2 = E^{\ell} (\Delta M_{\ell}^2)$

on obtient $B_{\ell} \leq \ell \sigma^2$.

Alors (voir NEVEU, § VII.2),

sur $(B_{\infty} < +\infty)$, M_{ℓ} converge presque sûrement vers une limite finie

sur $(B_{\infty} = +\infty)$, $M_{\ell} / \sqrt{B_{\ell}} \ln(B_{\ell}) \xrightarrow[\text{p.s.}]{} 0$ quand $\ell \rightarrow +\infty$, et donc

$$M_{\ell} = o(\sqrt{\ell} \ln(\ell)) \text{ p.s.}$$

Dans les deux cas, la sous-martingale $|Z|_{\ell}$ tend presque sûrement vers $+\infty$, puisque $A_{\ell} \rightarrow +\infty$ p.s. avec $A_{\ell} > \ell\alpha$ et que

$$M_{\ell}/A_{\ell} \xrightarrow[\text{p.s.}]{} 0.$$

Comme $(X_2 - X_1)^t$ est égale à Z^{ℓ} pour

$$T_{\ell} < t \leq T_{\ell+1}, \quad |X_2 - X_1|^t \xrightarrow[\text{p.s.}]{} +\infty \text{ quand } n \rightarrow +\infty.$$

ii) Reste à montrer que la probabilité que $Z^l = (X_2 - X_1)^{T_l}$ change de signe infiniment souvent est nulle.

Or $|Z^l| = A_l \left(1 + \frac{M_l}{A_l}\right)$ avec $A_l > l\alpha$ et $\frac{M_l}{A_l} \rightarrow 0$. Donc

il existe β ($0 < \beta < \alpha$) et $L(\omega)$, tels que $\forall l > L(\omega)$, $|Z^l| > l\beta$.

La probabilité d'un changement de signe, quand $|Z^l| > l\beta$, est majorée par

$$P(|V_{l+1}| > l\beta) < \frac{E(V_{l+1}^2)}{l^2 \beta^2}$$

qui est le terme général d'une série convergente. Donc d'après le lemme de Borel-Cantelli, Z^l ne peut changer de signe infiniment souvent : il n'y a pas d'oscillations.

Il en est de même pour $(X_2 - X_1)^t$. Mais alors pour chaque trajectoire, il existe un temps $T(\omega)$ après lequel $(X_2 - X_1)^t$ garde un signe constant, et tend donc soit vers $+\infty$, soit vers $-\infty$ d'après le i). \square

2. COMPORTEMENT DE X^t .

Reste à montrer que lorsqu'une des composantes de (X^t) tend vers $+\infty$, l'autre converge en loi.

PROPOSITION V.2.1. *Cas de deux neurones.*

Quand $\theta > \theta_0 = E(\mathcal{F})$, sur presque toute trajectoire, une des deux composantes tend vers $+\infty$, l'autre converge en loi vers la loi propre définie par la loi \mathcal{F} .

Démonstration. Pour chaque trajectoire, d'après V.1.1, une des deux composantes tend vers $+\infty$, puisque l'autre est positive.

Soit ω une trajectoire sur laquelle $X_1^t \rightarrow +\infty$. Soit $\tau_A(\omega)$ le temps d'entrée dans $(X_1(\omega) > A)$ et E_k l'événement

$U_1 + U_2 + \dots + U_k - k\theta < A - \theta$. Tant que E_1, E_2, \dots, E_k sont réalisés, c'est le neurone (2) qui décharge.

On a $\mathbb{P}(\bigcap_{k=\tau_A(\omega)}^{+\infty} E_k) > \varphi(A)$ où φ est une certaine fonction de A avec $\varphi(A) \rightarrow 1$ quand $A \rightarrow +\infty$.

En effet $\mathbb{P}(\bigcup_{k \geq \tau_A(\omega)} C E_k) \leq \sum_k \mathbb{P}(C E_k)$. Mais $\exists k_0, \exists C$ tel que $k \geq k_0$ entraîne

$$\mathbb{P}\left(\frac{U_1 + \dots + U_k - kE(\mathcal{F})}{k} > \theta - E(\mathcal{F}) + \frac{A - \theta}{k}\right) \\ \leq \mathbb{P}\left(\frac{U_1 + \dots + U_k - kE(\mathcal{F})}{k} > \theta - E(\mathcal{F})\right) \leq C \exp(-kh(\theta - E\mathcal{F}))$$

où h est la transformée de Cramer de la loi \mathcal{F} , qui existe du côté gauche puisque U est bornée à gauche.

Pour $k < k_0$,

$$\mathbb{P}(U_1 + \dots + U_k - k\theta > A - \theta) \leq \varphi_k(A)$$

pour une certaine fonction φ_k .

$$\text{Alors } \mathbb{P}(\bigcup_{k \geq \tau_A(\omega)} C E_k) \leq k_0 \max_k \varphi_k(A) + C \sum_{k_0}^{+\infty} e^{-kh(\theta - E\mathcal{F})}$$

Donc pour tout ε , il existe K tel que $k_0 > K \Rightarrow$ le deuxième terme $< \frac{\varepsilon}{2}$ et il existe A_0 tel que $A > A_0 \Rightarrow$ le premier terme avec $k_0 = K$ est inférieur à $\frac{\varepsilon}{2}$. D'où la propriété annoncée. Donc sur

$E_A^\infty = \bigcap_{k \geq \tau_A} E_k$, la loi limite de (X_2^t) conditionnellement à E_A^∞ tend vers la loi propre de X_2 , puisque E_A^∞ est asymptotiquement de probabilité 1. \square

On trouvera dans l'Appendice 1, les résultats de simulations menées pour différentes lois \mathcal{F} , et différentes valeurs de θ .

VI. DIVERGENCE. CAS GENERAL.

$$\theta > \theta_0 = E(\mathcal{F})/|V|.$$

1. STABILITE.

On constate dans les simulations que θ_0 est sous-estimé : ceci s'explique si l'on considère la démonstration des propositions III.3.1. et III.3.2 : on a majoré le nombre de décharges des voisins d'un neurone i par la situation où ils déchargeraient selon leur loi propre (c'est-à-dire avec la fréquence $\frac{1}{E(\mathcal{F})}$ au lieu de $\frac{1}{E(\mathcal{F})+\theta|V|}$).

Lorsque $\theta > \theta_0$, lorsqu'il y a divergence, apparaissent des "cartes" de neurones inhibés et de neurones actifs : certaines différences tendent vers $+\infty$. On trouvera dans l'Appendice 2 les résultats de simulations menées à partir d'états initiaux tirés suivant la loi $\mu = \mathcal{F}^{\otimes n^2}$ et différentes configurations divergentes, obtenues suivant la forme de V et la valeur de θ .

Une "carte" stable de neurones actifs et inhibés (c'est-à-dire de sites où $X_j^t \rightarrow +\infty$ et de sites où X_i^t reste borné) définit une partition du réseau en deux sous-ensembles E et F (actifs, inhibés) tels que à partir d'un certain temps T , les neurones qui déchargent appartiennent tous à E .

On cherche un critère de "stabilité asymptotique" d'une carte donnée.

On suppose que, pour tout i de E , $V_E(i) = \{i' \in E / i \in V(i')\}$ a le même cardinal v et que pour tout j de F , $W_E(j) = \{i \in E / j \in V(i)\}$ est non vide et a un cardinal w_j .

Remarquons tout de suite que si $|W_E(j)| = 0$ pour un certain site, nécessairement j devient actif au bout d'un temps fini, puisqu'il décroît de manière déterministe sans jamais être inhibé.

D'après les résultats démontrés au § IV, si

$$\theta v < E(\mathcal{F})$$

il y a récurrence du processus restreint à E et en notant $N_i(t)$ le processus de comptage au site i (qui compte les décharges du

neurone i , avant l'instant t), pour toute mesure initiale μ sur E , et pour tout site i ,

$$\frac{N_i(t)}{t} \xrightarrow[t \rightarrow +\infty]{\text{p.s.}} \frac{1}{E(\mathcal{F}) + \theta v}$$

Alors pour $j \in F$, on définit

$$Y_j(t) = \theta \sum_{i \in W_E(j)} N_i(t)$$

$$\frac{Y_j(t)}{t} \xrightarrow[t \rightarrow +\infty]{\text{p.s.}} \frac{\theta w_j}{E(\mathcal{F}) + \theta v}$$

Alors si

$$\forall j \in F, \theta(w_j - v) > E(\mathcal{F})$$

il existe t_0 (dépendant de μ) tel que

$$\mathbb{P}^\mu (\forall j \in F, Y_j(t) > t) \geq 1 - \varepsilon$$

En tirant les valeurs initiales X_j^0 sur les sites $j \in F$ suivant une loi ν telle que $X_j^0 > 2t_0$, $\forall j$, (définissant ainsi une mesure $\mu + \nu$ sur $E \cup F$), on assure qu'aucune décharge n'aura lieu dans F avant le temps t_0 (pour $t < t_0$, $X_j^t > 2t_0 - t_0$, $\forall j \in F$).

Alors pour tout $\varepsilon > 0$, il existe une mesure initiale $\mu + \nu$ sur $E \cup F$, telle qu'avec probabilité $(1 - \varepsilon)$, $(X_i^t)_{i \in E}$ est un processus récurrent et $(X_j^t)_{j \in E} \rightarrow +\infty$ quand $t \rightarrow +\infty$.

Comme tout état est atteignable en $N+1$ transitions, (lemme III.2.1 et corollaire III.2.4.) on obtient la

PROPOSITION VI.I.I.

Soit une partition E, F du réseau, avec

$$V_E(i) = \{i' \in E / i \in V(i')\} \quad \text{pour } i \in E,$$

$$W_E(j) = \{i \in E / j \in V(i)\} \quad \text{pour } j \in F,$$

et

$$|V_E(i)| = v, \quad |W_E(j)| = w_j > 0, \quad \forall j \in F.$$

Alors à condition que $\theta v < E(\mathcal{F}) < \theta(w_j - v)$, $\forall j \in F$, il y a une probabilité positive que le processus $(X_i^t)_{i \in E}$ soit récurrent et que $(X_j^t) \xrightarrow[t \rightarrow +\infty]{\text{p.s.}} +\infty$ quand $t \rightarrow +\infty$, pour tout $j \in F$.

REMARQUES.

1) Si E est un ensemble sur lequel $(X^t)_E$ est récurrent ($\sup_{i \in E} |V_E(i)| \theta < E(\mathcal{F})$), alors E "absorbe" tout point j de F tel que $w_j \theta < E(\mathcal{F})$. Une situation (E, F) stable est nécessairement saturée pour

$$\sup_{i \in E} |V_E(i)| \theta < E(\mathcal{F})$$

(VI.1.2)

et vérifie donc

$$(\inf_j |W_E(i)| - \sup_i |V_E(i)|) \theta > E(\mathcal{F})$$

(VI.1.3)

(cela implique en particulier $2 \sup_i |V_E(i)| < \inf_j |W_E(j)|$).

Il y aura bascule et instabilité si l'une de ces deux conditions n'est pas réalisée.

Le cas $|V| \theta < E \mathcal{F}$ correspond à $w_j = v = |V|$, et à la récurrence sur E tout entier.

2) La largeur d'une bande de neurones de F est limitée par la taille du voisinage V : tout site de F doit être inhibé par au moins un site de E : $w_j > 0 \forall j \in F$ et $w_j > v$.

2. EXEMPLES DE SITUATIONS LIMITEES.

Le réseau sous-jacent est un tore, dont on figure une image locale - qu'il faut imaginer périodique sur le tore. Bien entendu tout translatée d'une situation limite en est une.

On note \times les points de E ,
 \circ les points de F .

On note $\theta_1 = \sup_j \frac{E\mathcal{F}}{w_j - v}$, $\theta_2 = \frac{E\mathcal{F}}{v}$ et (E, F) est stable pour $\theta_1 < \theta < \theta_2$.

a) $|V| = 4$

$V = \begin{matrix} & & \circ & \\ & \circ & \times & \circ \\ & & \circ & \end{matrix}$

$$1) \begin{pmatrix} \times & \circ & \times & \circ \\ \circ & \times & \circ & \times \\ \times & \circ & \times & \circ \\ \circ & \times & \circ & \times \end{pmatrix}$$

$$v = 0 \quad \theta_1 = \frac{E(\mathcal{F})}{4} \quad \theta_2 = +\infty$$

$$w_j = 4$$

$$2) \begin{pmatrix} \times & \circ & \circ & \times \\ \circ & \times & \circ & \circ \\ \circ & \circ & \times & \circ \\ \times & \circ & \circ & \times \end{pmatrix}$$

$$v = 0 \quad \theta_1 = \frac{E(\mathcal{F})}{2} \quad \theta_2 = +\infty$$

$$w_j = 2$$

COMMENTAIRE : Si le tore étudié est $T = \mathbb{Z}/2m \times \mathbb{Z}/2m'$, la forme particulière des voisinages fait que le réseau se partitionne en damier :

$T = E+F$ avec $V(i) \subset E$ pour $i \in F$ et vice-versa. Et cependant, on trouve une situation limite (la 2)) qui n'est pas un damier. Noter que pour $\theta > \frac{E(\mathcal{F})}{2}$, les deux types sont des états finaux.

b) $|V| = 6$

$$V = \begin{array}{cccccc} & & & & & \circ \\ & & & & & \circ \\ & & & & & \circ \\ & & & & & \circ \\ & & & & & \circ \\ & & & & & \circ \end{array}$$

1) $\begin{pmatrix} \circ & \circ & \times & \times & \circ & \circ \\ \times & \times & \circ & \circ & \times & \times \\ \circ & \circ & \times & \times & \circ & \circ \\ \times & \times & \circ & \circ & \times & \times \end{pmatrix}$ $v = 1$ $w_j = 5$ $\theta_1 = \frac{E(\mathcal{F})}{4}$ $\theta_2 = E(\mathcal{F})$

2) $\begin{pmatrix} \times & \circ & \circ & \times & \circ & \circ & \times \\ \circ & \times & \circ & \circ & \times & \circ & \circ \\ \circ & \circ & \times & \circ & \circ & \times & \circ \\ \times & \circ & \circ & \times & \circ & \circ & \times \end{pmatrix}$ $v = 0$ $w_j = 3$ $\theta_1 = \frac{E(\mathcal{F})}{3}$ $\theta_2 = +\infty$

3) $\begin{pmatrix} \times & \circ & \circ & \circ & \times & \circ & \circ & \circ \\ \circ & \times & \circ & \circ & \circ & \times & \circ & \circ \\ \circ & \circ & \times & \circ & \circ & \circ & \times & \circ \\ \circ & \circ & \circ & \times & \circ & \circ & \circ & \times \end{pmatrix}$ $v = 0$ $w_j = 2$ $\theta_1 = \frac{E(\mathcal{F})}{2}$ $\theta_2 = +\infty$

4) $\begin{pmatrix} \times & \circ & \circ & \circ & \circ & \times & \circ & \circ \\ \circ & \times & \circ & \circ & \circ & \circ & \times & \circ \\ \circ & \circ & \times & \circ & \circ & \circ & \circ & \times \\ \circ & \circ & \circ & \times & \circ & \circ & \circ & \circ \end{pmatrix}$ $v = 0$ $w_j = 1 \text{ ou } 2$ $\theta_1 = \frac{E(\mathcal{F})}{1}$ $\theta_2 = +\infty$

c) $|V| = 8$

1) $\begin{pmatrix} \circ & \circ & \circ & \circ \\ \times & \times & \times & \times \\ \circ & \circ & \circ & \circ \\ \times & \times & \times & \times \end{pmatrix}$ $v = 2$ $w_j = 6$ $\theta_1 = \frac{E(\mathcal{F})}{4}$ $\theta_2 = \frac{E(\mathcal{F})}{2}$

2) $\begin{pmatrix} \circ & \circ & \circ & \circ \\ \circ & \times & \circ & \times \\ \circ & \circ & \circ & \circ \\ \circ & \times & \circ & \times \end{pmatrix}$ $v = 0$ $w_j = 2 \text{ ou } 4$ $\theta_1 = \frac{E(\mathcal{F})}{2}$ $\theta_2 = +\infty$

On a vérifié (Appendice 3) que ces exemples correspondent effectivement à des situations "stables" pour les valeurs convenables de θ .

De plus, on constate (Appendice 2) qu'elles sont effectivement situations limites quand on initialise le processus (X^t) au hasard. Resterait à calculer comment se fait le choix entre les différentes cartes stables lorsqu'il en existe plusieurs pour une même valeur de θ .

VII. CONCLUSION

Nous avons proposé un modèle du cortex cérébelleux, pour étudier les cellules de Purkinjie reliées seulement par les collatérales inhibitrices. Pour les valeurs du délai θ , (qui modélise l'inhibition d'un neurone due à la décharge d'un voisin), correspondant à la réalité (θ de l'ordre de 3 ou 4 ms, pour EU de 100 ms) on observe l'établissement d'un régime stationnaire en accord avec la théorie. Une évaluation du seuil θ_0 en dessous duquel le régime stationnaire s'installe est faite, bien qu'il s'agisse d'une minoration.

Lorsque l'inhibition est forte ($\theta > \theta_0$) le réseau diverge. L'intérêt est que s'établit alors une carte formée de bandes alternées de neurones actifs et inhibés. La forme et l'orientation de ces bandes dépend de la forme du voisinage considéré, de sa largeur et de la valeur de θ . Lorsqu'on franchit certaines valeurs de θ , on assiste à un phénomène de bifurcation : de nouvelles situations stables apparaissent sans que pour autant disparaissent les précédentes (voir spécialement le cas à 6 voisins). On aurait pu penser qu'autour de chaque neurone actif se serait établie une zone d'inhibition latérale, ces zones isolant les uns des autres les neurones actifs. Au contraire on voit qu'en augmentant l'inhibition, on fait "éclater" ces zones isolées, en provoquant la formation de moirures tout-à-fait semblables à celles qu'on observe dans les différents tissus du cortex lorsqu'on les soumet à des stimuli : cartes d'orientation préférentielle, ou de dominance oculaire dans le cortex visuel par exemple. Notre modèle met donc de façon particulièrement claire en évidence le fait que c'est l'inhibition qui cause ces différenciations.

Il reste à tenir compte du développement des fibres parallèles et modéliser le système C (cf. § I.2), pour déterminer les réponses du réseau à des stimuli extérieurs. C'est l'objet de l'article suivant.

Je tiens à remercier R. AZENCOTT et J. BRETAGNOLLE pour m'avoir encouragée et guidée au cours de fréquentes discussions.

REFERENCES

- AXELRAD H., BERNARD C., GIRAUD B., MARC M.E. (1985), CRAS PARIS, t. 301, Série III, n° 10.
- AXELRAD H., BERNARD C., COTTRELL M., GIRAUD B. (1987). The use of an artificial neural network to analyse the informational transfer properties of a simplified model of the cerebellar cortex in IEEE. Proceedings of the First International Conference on neural networks (à paraître).
- BLASDEL G.G., SALAMA G. (1986) . Nature, vol. 321, p. 579.
- CHAUVET G. (1986). Biol. Cybern. 55, p. 201-209.
- COX D.R., LEWIS P.A.W. (1968). Statistical analysis of series of events. Methven and Coltd, London.
- CREPEL F., DELHAYE-BOUCHAUD N., DUPONT J.L., SOTELO C. (1980). Neuro-science 5, p. 333-347.
- ECCLES J.C., ITO M., SZENTAGOTHAI J. (1967). The cerebellum as a neuronal machine. Springer, Berlin Heidelberg New York.
- GIRAUD B., BERNARD C., AXELRAD H. (1985). CRAS, Paris t. 301, Série III, n° 11.
- ITO M. (1984). The cerebellum and neural control. Raven Press, New York.
- KOHONEN T., (1984). Self-organization and Associative memory, Springer, Berlin Heidelberg, New York.
- LANDOLT J.P., CORREIA M.J. (1978). IEEE Trans. BME25, p. 1-12.
- MARR D. (1969). A theory of cerebellar cortex. J. Physiol. 202, p.437-470.
- NEVEU J. (1972). Martingales à temps discret. Masson, Paris.
- NUMMELIN E. (1984). General irreducible Markov chains and non-negative operators. Cambridge University Press.
- PALAY S.L., CHAN-PALAY V. (1974). Cerebellar cortex. Springer, Berlin Heidelberg, New York.
- REVUZ D. (1975). Markov chains, North Holland.
- RUMELHART T.E., Mc CLELLAND J.L. (1986). Paralel distributing processing, Tome 1 et 2, MIT Press, Cambridge.
- SAMPATH G., SRINIVASAN S.K. (1977). Stochastic models for spike trains of single neurons, Springer, Berlin, Heidelberg, New York.

APPENDICE 1 - CAS DE DEUX NEURONESLoi \mathcal{U} uniforme

$\theta_0 = E(\mathcal{U})$	θ	Intervalle inter-spike moyen théorique	Intervalle inter-spike moyen observé	Var. empirique	Fréq. moyenne x 100 unités de temps	Corr. des fréquences
3	0	3	2.9	1.98	34.8	-0.1
3	1	4	3.5	4.81	27.8	-0.2
3	2	5	4.1	15.35	23.2	-0.9
3	3	diverge $X_1 \rightarrow \infty$				
5	0	5	4.9	7.07	20.4	0.2
5	1	6	5.6	11.84	17.6	-0.7
5	2	7	6.6	23.8	15.5	-0.7
5	3	8	7.2	47.9	13.8	-0.8
5	4	9	7.9	142.4	11.5	-0.9
5	5	diverge $X_1 \rightarrow \infty$				
10	0	10	9.7	32.06	10.5	-0.2
10	5	15	14.1	132.3	6.9	-0.7
10	7	17	15.5	360.1	6.1	-0.9
10	9	19	28.1	1301	5.5	-0.9
10	10	diverge $X_2 \rightarrow \infty$				
20	0	20	19.45	127.4	5.1	-0.2
20	5	25	24.7	228.2	4.2	-0.5
20	10	30	31.1	494.5	3.4	-0.6
20	15	35	36.1	1304	2.9	-0.9
20	20	diverge $X_1 \rightarrow \infty$				

Loi \mathcal{F} exponentielle

$\theta_0 = E(\mathcal{F})$	θ	Intervalle inter-spike moyen théorique	Intervalle inter-spike moyen observé	Var. empirique	Fréq. moyenne x 100 unités de temps	Corr. des fréquences	
3	0	3	3.0	5.1	32.4	-0.2	
3	1	4	3.8	13.3	26.4	-0.7	
3	2	5	4.9	63.2	21.0	-0.9	
3	3	diverge $X_1 \rightarrow \infty$					
4	0	4	4.1	10.6	24.1	-0.2	
4	1	5	5.0	19.7	20.5	-0.8	
4	2	6	6.3	55.8	17.1	-0.9	
4	3	7	7.3	154.5	14.5	-1.0	
4	4	8	9.9	1430	12.1	-1.0	
4	5	diverge $X_1 \rightarrow \infty$					
5	0	5	5.2	19.2	19.3	0.1	
5	1	6	6.2	33.8	16.7	-0.7	
5	4	9	9.0	255.4	10.7	-0.9	
5	5	diverge $X_1 \rightarrow \infty$					
10	0	10	10.7	107.5	9.1	-0.1	
10	5	15	15.7	348.5	6.5	-0.7	
10	7	17	16.3	644.2	5.6	-0.9	
10	9	19	19.5	2600	5.1	-0.9	
10	10	diverge $X_1 \rightarrow \infty$					
20	0	20	20.7	374.3	4.7	0.2	
20	5	25	26.6	810.4	3.7	-0.3	
20	10	30	32.7	1751	3.2	-0.3	
20	15	35	34.1	3852	2.6	-0.8	
20	20	diverge $X_1 \rightarrow \infty$					

Loi \mathcal{F} normale

$\theta_0 = E(\mathcal{F})$	θ	Intervalle inter-spike moyen théorique	Intervalle inter-spike moyen observé	Var. empirique	Fréq. moyenne x 100 unités de temps	Corr. des fréquences
3	0	3	2.8	1.1	39.1	0.2
3	1	4	3.2	2.7	32.4	-0.1
3	2	5	3.9	18.5	25.6	-1.0
3	3	diverge $X_1 \rightarrow \infty$				
5	0	5	4.8	1.1	21.9	0.1
5	1	6	5.6	2.4	19.2	0.1
5	4	9	7.8	86.3	13.9	-0.9
5	5	diverge $X_1 \rightarrow \infty$				
10	0	10	9.8	12.2	10.5	-0.2
10	5	15	13.6	57.3	7.5	-0.6
10	10	diverge $X_1 \rightarrow \infty$				

APPENDICE 2 - CAS DE n^2 NEURONES

EXEMPLES DE DIVERGENCES $\theta > \theta_0$.

Les exemples suivants correspondent à des réseaux de 40×40 neurones, pour des types différents de voisinages \mathcal{V} , dans le cas

$$\theta > \theta_0 \simeq \frac{E\mathcal{F}}{|\mathcal{V}|} .$$

Pour chaque exemple, on écrit l'état du processus (X^t) au bout de 2000 itérations, la fréquence observée des neurones - on remarque que prédominent pour les neurones non inhibés des valeurs des fréquences proches de la fréquence propre, inverse de l'espérance de la loi \mathcal{F} -, l'intervalle inter-spike moyen observé des neurones.

Enfin on a choisi un seuil (intervalle inter-spike supérieur à 4 fois l'intervalle moyen propre) pour faire une carte des neurones actifs et inhibés.

Les résultats obtenus sont les mêmes si on change de loi \mathcal{F} .

On remarquera les bandes alternées de neurones actifs et de neurones inhibés, l'orientation et la largeur des bandes dépendant de la forme de \mathcal{V} .

Les configurations représentées ne sont pas forcément l'état limite (t infini), mais correspondent à une "réponse" quasiment stable du réseau neuronal, établie en un temps long par rapport à la durée de l'intervalle inter-spike moyen.

Dans tous les cas apparaissent des "formes typiques". Voir dans l'Appendice 3 une étude plus détaillée des cas $|\mathcal{V}| = 4, 6$ et 8 .

EXEMPLE 1-4 VOISINS

Le réseau est de taille 40 * 40

La forme du voisinage à 4 voisins est: x
 x . x
 x

La loi propre des intervalles inter-spikes est uniforme

Le paramètre d'inhibition teta vaut 6

L'intervalle inter-spikes moyen propre est 10

La fréquence moyenne de décharge propre pour 100 unités de temps est: 10

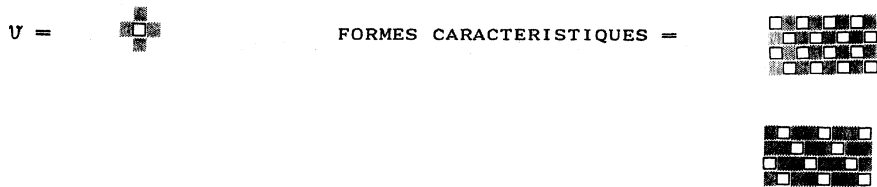
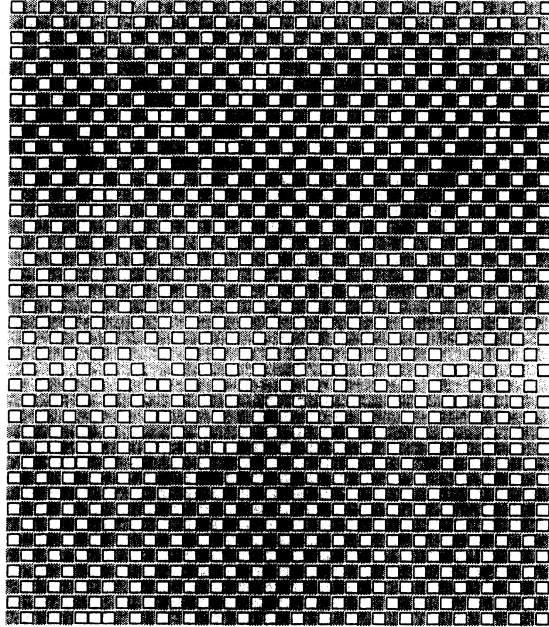
Etat du système au bout de 2000 itérations

```

61155 41313 91192 121259 51165 31316 91210 171042 141270 71250 91333 121216 151258 141469 41238 61435
820 61812 22117 12093 162025 182116 12238 71972 152005 62160 112140 32133 62066 22164 12128 132076 823
8 555 191101 11923 131879 112050 102025 22104 141977 12040 111847 42017 22082 112063 71891 11732 62046
4 477 7 76 128 13 669 21048 12013 52005 132095 82068 31078 10 926 52129 22103 111249 6 711 13 823 410
689 31720 6 766 11 638 7 80 424 12017 111992 42204 4 788 13 108 7 365 141918 11 584 34 18 253 1 434 9 5
61172 32013 21807 152000 9 300 297 32048 132079 31031 8 571460 13 337 16 8 15 232 111455 161709 141508
23 7 466 132030 72295 72110 6 196 564 122063 11586 8 591360 42148 4 922 5 856 192004 102168 52229 619
427 3 188 309 142126 122189 81865 13 171236 141953 1 711020 72140 112275 51789 51895 52067 82045 12195
41902 11 503 342 172055 132132 21480 15 141881 1 620 218 142170 102296 112128 91904 11765 91947 92238 1118
1124 22235 10 389 301 162102 82050 12 725 60 111433 15 41866 12153 102061 21943 101726 31829 112147 51844
11919 32124 18 200 442 32147 42040 2 328 1 61 9 808 122207 82104 52192 91994 141784 61906 21913 5 2
692 61475 42033 12 31130 72085 122011 16 697 4 773 61908 32116 112220 72270 81927 42021 61998 6 250 4
4 461 4 562 4 992 311 21930 32127 52072 71593 61838 12084 82008 52257 122189 92049 42243 12 249 424
10 957 10 218 7 278 91318 22126 12165 82077 22010 131977 72005 112042 42248 32161 12178 1 357 335 1222
949 11770 131134 31226 152149 82101 32004 42037 12012 82020 131958 72058 112147 92142 13 417 295 32266
11708 162016 31666 162071 52161 32083 21983 132134 12069 22098 31989 92202 22057 7 982 11 112156 120
415 41543 52116 42002 32117 192150 92085 22163 12241 32078 72030 172075 42133 11772 9 171787 11937
9 363 21195 82192 102094 82052 12255 12185 82144 22223 22092 92048 82154 112146 3 945 88 32082 319
8 808 30 13 762 121939 121985 132183 142141 122075 112215 32228 92050 32099 32316 162026 3 312 510 81773
1198 141421 11 63 28 6 768 21871 32114 162025 142012 32232 92235 142166 122159 22252 52004 3 356 279 816
92114 101810 1 948 1 100 189 41689 121994 21907 42101 102268 22251 12193 112206 52168 132012 2 610 351
1277 82053 122067 11807 3 524 158 141979 31946 11955 62162 51857 31808 91940 42087 12127 92078 2 756 3
12133 142069 22147 61957 3 605 291 32014 132010 32116 11674 8 643 16 663 111255 102059 92087 102152 2 1
1323 52057 22155 62189 32133 19 253 665 42077 172132 31535 12 207 229 9 318 46 16 574 41966 92078 91772
52127 71985 102231 42213 101996 6 71534 22295 91702 9 242 757 41937 91348 18 74 220 12092 62032 1 6
1238 11927 21729 81855 22075 14 916 492 22098 71982 8 146 743 32256 32172 91793 8 273 380 62151 21843
21922 41274 41156 151680 121811 13 131563 62170 3 400 362 172217 22162 42039 11881 3 260 601 62187 211
1165 31261 14 426 6 939 81992 2 512 479 112208 161267 27 92077 22083 52045 12009 31935 15 246 500 142193
101729 10 479 80 316 71582 111680 3 92021 81993 13 61459 132046 52062 62041 22185 142009 17 207 319 42
1226 161119 13 293 10 925 22073 6 747 347 41998 11 133 560 12136 62025 162129 42220 112257 82104 1 307 309
92017 141024 5 953 21881 81858 11 277 223 3 51 234 62065 32094 72121 62125 102197 172171 42122 1 318 3
1275 121992 21633 71872 41993 81792 5 874 5 797 102005 102080 91999 82026 22167 82245 82062 142114 10 4
82048 12180 62136 11965 82104 52017 121672 102081 81906 51730 111840 21956 42312 52275 11982 62286
1032 42152 102185 92066 122083 22129 12086 32010 11979 61490 61506 62016 92059 142195 21997 12028 721
81970 72135 42109 51915 21968 42103 42055 32043 101679 91342 61804 32167 32119 42108 132115 12246
1089 22176 92256 61853 91882 152036 21969 92034 121923 41427 61473 82070 132146 62121 72255 112095 121
152218 52133 111831 171706 82089 82003 72112 142155 31999 101635 11901 152052 172228 142130 32126 52142
1239 12200 101642 161360 31828 22134 132135 82148 142204 141990 81813 32061 12193 92158 52180 142137 421
32215 21895 11250 11137 52109 82098 22190 152163 112108 72076 101965 82201 22133 42123 62065 42183
318 11270 2 824 1 532 16 830 131301 11306 181180 61311 21374 91211 71265 41228 21151 81256 61290 111

```


ETAT FINAL (2000 ITERATIONS)



Les zones *grises* représentent les neurones inhibés

Les zones *blanches* représentent les neurones actifs.

EXEMPLE 2-6 VOISINS

Le réseau est de taille 40 * 40

La forme du voisinage à 6 voisins est:

```

      x
    x x . x x
      x
  
```

La loi propre des intervalles inter-spikes est uniforme

Le paramètre d'inhibition tête vaut 6

L'intervalle inter-spikes moyen propre est 10

La fréquence moyenne de décharge propre pour 100 unités de temps est: 10

Etat du système au bout de 2000 itérations

```

131329 313 7 5891036 16 21475 41 11384 259 41018 764 2 1421234 8 17 817 864 5 3 766 807 9 101060 622 6 1
1162 1 482264 4 71985 382 1714761090 8 8041847 2 22281 17 1412211552 7 414771460 6 312961372 11 32015 4
31971 200 131676 819 7 9981467 7 3212265 17 181747 825 3 4662001 12 91628 795 5 4 8931685 1 515041352 3
866 3 4341909 5 1592169 16 62222 34 315131525 8 452272 6 61786 716 13 425 58 41 401 9 6911957 1 102195 4
51429 185 82116 304 62036 645 2 9701939 6 32036 464 114421060 713541259 61450 427 121889 784 15 9131758 3 2
157 111348 789 811771279 8 7222033 1 1016461207 2 2802105 2 3322114 1 5951396 15 9661415 5 271477 743 3111511
783 184 1114191082 7 6691963 2 914441213 5 21994 656 1212051144 2 432 24 184 500 18 455 207 7 170 31515 540
813081586 12 3552230 3 41667 611 7 10 7731689 10 112379 199 910781074 1 11405 954 21221 237 131500 3 1716
1391 12 92393 45 71923 593 5 401 100 27 310 7 8501766 8 672410 19 31707 838 212861471 1010481323 1612421021
41982 753 1711571561 8 4582024 7 2461306 412341155 1812911458 11 6471865 13 2282135 15 2962285 6 1792123 14 22421
1455 1 542328 16 62296 8 101894 822 81136 859 11979 365 111997 773 314171184 1510301291 9 9631462 5 8851478
62242 353 171847 780 8 7642024 7 21615 259 2 605 2 6041724 10 112313 40 172240 158 42094 279 82178 297 320
339 512191234 17 4402087 13 514071515 4 1 602 93 877 2 1118021028 15 5451897 4 3261935 14 3702105 6 3812178
329 273 9 4102024 7 114141480 4 511271016 13 8 1411861329 4 21938 716 71025 643 151698 277 111809 692 3 4
1 5302196 9 514751034 4 313301321 8 3 567 603 165 14 1312401437 5 121404 244 1 378 1 9101107 15 10 928 8
1586 9 1716021161 5 141423 461 9 114051282 8 15 148 209 743 9 313091249 6 2 5421031 82 11 501 256 27 364
611191741 6 52185 255 5 263 991 164 14 101672 177 12 396 10 2841810 2 610911207 15 413341050 11295 447 319
1420 13 82149 542 15 9811962 3 12 3251241 35 21137 775 181563 802 4 6762072 12 121855 640 512361210 312961378
217681110 3 4802225 5 713731602 6 712531043 9 3691963 8 92010 454 110251886 3 4242102 2 4022238 1 24220
1412 6 502415 14 1612401506 5 916351220 1 12208 18 214341280 9 141968 790 1613331307 912841238 91115 941
71916 746 11 7861936 11 913791553 5 32228 210 1011211798 3 61561 853 3 72305 63 51503 990 7 9231072 3 4
1205 1 32327 30 712151552 4 21905 900 7 4472132 5 52036 418 1 211424 156 8 5331325 8 415191169 11 186 9
21857 328 1 9471463 3 813691313 10 62267 12 91924 691 6 5271933 18 131359 797 2 358 27 451 541 412221459
135 1211031305 2 471289 466 6 61963 490 1210091748 2 452261 13 172245 601 112391101 111486 230 41946 802 11 6
265 267 3 2021811 2 7 83 571 580 12 202164 5 22293 121 51829 913 712111250 9 8261797 1 922272 13 11762 6
7 5671909 18 61570 433 2 521 2 718 962 12 8021827 3 4122069 3 3882118 2 182350 4 101538 680 12 3031621 17
1292 6 111935 607 513181032 121631 207 82217 221 31998 205 52182 222 82216 346 11 9041075 18 611051367 4 3817
1 8681591 9 2722148 7 2992080 2 3522177 5 3302205 12 3632043 15 3931514 5 887 969 3 426 34 715 831 1610111725
171 326 210981058 112691156 114661293 1814931369 1513291250 5 542 220 72 409 21129 826 21584 259 132169 307 1019
302 131199 788 121692 245 32265 325 121991 861 714061354 9 9361314 13 410851097 6 4321894 1 1872104 15 2652235
111705 8 201407 12 651857 1 402309 3 72058 357 714331105 4 5521690 16 72293 41 12175 299 142224 476 321
249 11284 750 91969 360 22083 564 110271689 2 162233 11 161756 943 2 9441390 810031070 1111891348 211551265
361 194 1013161375 810431287 9 3581989 11 21966 511 9 9381414 11 11270 640 9 439 13 842 822 10 8481656 12 52819
16 3201900 2 8711269 2 5821645 18 81777 808 710681202 7 337 1 319 348 4 1201400 25 8 3292170 10 82006 546
40 14 10 533 6 2671551 2 61870 448 512321185 511971118 181393 293 102217 298 15 5821525 11 3 7821755 6 317
327 1 16 675 576 2 32151 407 1913221240 813581220 114151363 1614681152 311761163 15 347 13 528 593 12 5651901
615401117 3 71988 460 313051171 513611326 111831291 312791310 9 7462008 8 1911437 141104 198 32168 259 410
1467 1 22027 574 511031192 412771246 712771204 311441160 7 5801473 7 1 852 397 42 413 91014 944 7 7211857
1317411161 610871229 212651242 812381390 511741242 7 9671166 4 372 4 135 600 7 41233 413 4 7792147 11 614
414 12 2241192 10 1831152 6 2431233 3 2671342 14 2141141 9 176 954 6 533 316 4 953 472 2 271445 4 31002 604
  
```

Fréquence de chaque neurone en nb pour 100 unités de temps - le 0 indique les neurones inhibés

```

11 0 0 10 0 0 6 6 0 0 11 0 0 10 0 0 6 6 0 0 6 8 0 0 5 8 0 0 9 0 0 9 2 0 5 7 0
0 11 0 0 7 6 0 0 10 0 0 10 0 0 2 8 0 0 6 5 0 0 6 6 0 0 6 6 0 0 6 0 0 10 0 0 11 0 0 10 0
10 0 0 10 0 0 10 0 0 10 0 0 8 5 0 0 10 0 0 6 7 0 0 7 6 0 0 7 6 0 0 8 5 0 0 10 0 0 11 0
0 10 0 0 10 0 0 9 2 0 4 8 0 0 8 4 0 3 7 0 0 9 0 0 0 10 0 0 7 6 0 0 11 0 0 10 0 0 11 0
7 0 1 10 0 0 10 0 0 11 0 0 6 7 0 0 11 0 0 10 0 0 11 0 0 10 0 0 10 0 0 10 0 0 10 0 0 7 5 0
2 8 0 0 10 0 0 10 0 0 6 7 0 0 10 0 0 10 0 0 9 0 2 0 0 10 0 0 1 0 9 0 0 10 0 0 9 0 9 0 10 0
0 1 9 0 0 10 0 0 7 6 0 0 6 8 0 0 10 0 0 6 8 0 0 10 0 0 10 0 0 1 0 9 0 0 10 0 0 2 7 0 1 9 0
10 0 0 11 0 0 7 6 0 0 7 6 0 0 6 7 0 0 11 0 0 5 7 0 0 10 0 0 10 0 0 8 0 4 3 0 8 0 0 10 0 10 0
0 7 4 0 3 8 0 0 10 0 0 0 10 0 0 10 0 0 10 0 0 9 0 0 9 0 11 0 0 10 1 0 8 2 0 8 1 0 10 0 11 0
11 0 0 10 0 0 10 0 0 10 0 0 9 0 0 10 0 0 10 0 0 10 0 0 10 0 0 10 0 0 9 0 0 10 0 0 10 0 10 0
0 11 0 0 7 5 0 2 9 0 0 10 0 0 10 0 0 11 0 0 4 2 9 0 0 10 0 0 10 0 0 9 0 0 10 0 0 10 0 3 0 8
11 0 0 10 0 0 10 0 0 10 0 0 7 7 0 0 7 0 7 0 0 8 4 0 2 9 0 0 10 0 0 10 0 0 10 0 0 8 0 0 10 0
0 10 0 0 10 0 0 6 7 0 0 6 7 0 0 4 6 0 0 7 6 0 0 6 7 0 0 10 0 0 10 0 0 11 0 0 12 0 0 10 0
0 10 0 0 6 8 0 0 6 8 0 0 7 6 0 0 4 4 6 0 0 6 7 0 0 6 7 0 0 10 0 0 11 0 0 9 0 1 0 10 0 10 0
11 0 0 6 6 0 0 5 7 0 0 6 6 0 0 6 6 0 0 6 6 0 0 6 7 0 0 6 6 0 0 6 6 0 0 8 4 0 0 6 7 0 7 5 0
0 6 8 0 0 9 2 0 0 7 6 0 0 7 6 0 0 7 6 0 0 6 7 0 0 6 7 0 0 6 7 0 0 6 7 0 0 2 0 9 0 0 6 7 0 10 0
11 0 0 7 6 0 1 10 0 0 0 6 7 0 0 6 7 0 0 11 0 0 6 7 0 0 11 0 0 6 7 0 0 10 0 0 10 0 0 10 0 0
0 6 7 0 0 10 0 0 7 6 0 0 7 6 0 0 10 0 0 9 0 0 10 0 0 11 0 0 6 7 0 0 10 0 0 10 0 0 10 0 0
10 0 0 11 0 0 6 6 0 0 7 6 0 0 10 0 0 8 5 0 0 10 0 0 5 0 0 10 0 0 11 0 0 11 0 0 10 0 0 10 0 9
0 9 2 0 3 8 0 0 5 7 0 0 7 6 0 0 3 0 3 9 0 0 6 7 0 0 3 0 3 9 0 0 6 7 0 0 10 0 0 10 0 0 10 0
11 0 0 10 0 0 7 6 0 0 7 6 0 0 10 0 0 7 6 0 0 10 0 0 7 6 0 0 10 0 0 10 0 0 10 0 0 6 0 6 0 10 0 10 0
0 8 3 0 3 9 0 0 7 6 0 0 11 0 0 7 6 0 0 10 0 0 7 6 0 0 10 0 0 3 9 0 0 11 0 0 0 8 2 0 4 9 0
9 0 0 11 0 0 6 7 0 0 8 4 0 3 8 0 0 10 0 0 6 7 0 0 11 0 0 8 0 0 10 0 0 11 0 0 7 6 0 0 10 0 10 0
0 10 0 0 9 2 0 0 6 7 0 0 10 0 0 4 8 0 0 10 0 0 4 8 0 0 10 0 0 10 0 0 7 6 0 0 10 0 0 3 8 0 0
0 10 0 0 6 7 0 0 7 6 0 0 10 0 0 8 4 0 2 8 0 0 10 0 0 10 0 0 6 7 0 0 10 0 0 9 2 0 8 0 0 8 4 0 10 0
10 0 6 7 0 0 11 0 0 10 0 0 10 0 0 10 0 0 9 0 0 10 0 0 9 0 0 11 0 0 11 0 0 11 0 0 8 3 0 8 0 0 11
10 0 0 10 0 0 10 0 0 11 0 0 10 0 0 11 0 0 10 0 0 11 0 0 10 0 0 10 0 0 11 0 0 10 0 0 11 0 0 6 7 0
0 11 0 0 10 0 0 10 0 0 11 0 0 11 0 0 11 0 0 10 0 0 7 5 0 0 9 0 11 0 0 10 0 11 0 0 11 0 0 6 7 0
10 0 0 5 0 7 0 10 0 0 7 6 0 0 10 0 0 8 4 0 2 8 0 0 10 0 0 5 8 0 0 5 7 0 0 10 0 0 11 0 0 10 0
0 10 0 0 10 0 0 10 0 0 10 0 0 10 0 0 10 0 0 10 0 0 10 0 0 10 0 0 10 0 0 10 0 0 11 0 0 10 0 10 0
0 10 0 0 10 0 0 9 0 0 9 0 0 11 0 0 9 0 1 0 10 0 0 1 0 10 0 0 0 11 0 0 6 7 0 0 10 0 10 0 11 0
10 0 7 0 6 0 0 7 6 0 0 10 0 0 10 0 0 11 0 0 10 0 0 11 0 0 10 0 0 11 0 0 10 0 0 11 0 0 7 6 0 2
0 11 0 0 6 6 0 0 11 0 0 10 0 0 10 0 0 10 0 0 10 0 0 10 0 0 10 0 0 10 0 0 10 0 0 11 0 0 6 7 0 0
10 0 0 7 5 0 0 10 0 0 10 0 0 10 0 0 10 0 0 10 0 0 10 0 0 10 0 0 10 0 0 10 0 0 10 0 0 10 0 0
10 0 6 7 0 0 10 0 0 10 0 0 10 0 0 10 0 0 10 0 0 10 0 0 4 7 0 0 4 7 0 0 4 7 0 0 10 0 0 6 6 0 10 0
10 0 0 11 0 0 11 0 0 10 0 0 11 0 0 10 0 0 9 0 2 0 8 2 0 0 8 2 0 0 10 0 0 8 5 0 0 8 5 0 2 9 0
0 11 0 0 10 0 0 10 0 0 10 0 0 10 0 0 10 0 0 9 0 0 10 0 0 10 0 0 10 0 0 9 1 0 5 7 0 0 7 6 0 0 10 0

```

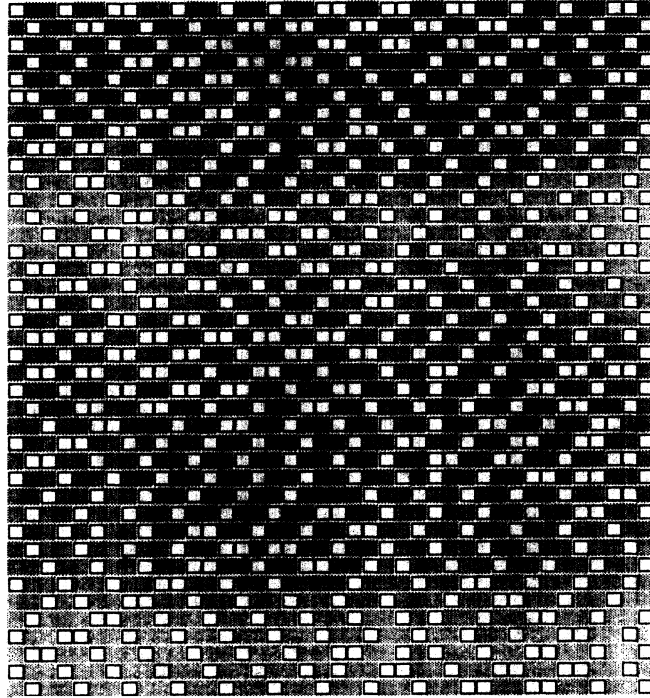
Intervalle inter-spike moyen de chaque neurone

```

9 * 9 * 14 15 * 83 9 * * 14 17 * 39 11 * * 14 14 * * 16 12 * * 17 11 * * 10 * * 10 46 * 17 13 *
* 9 * 9 * 13 15 * * 9 * 9999 * 11 18 * * 15 13 * * 13 15 * * 12 16 * * 12 17 * 115 9 * * 9 * 8 * 9
* 9 * 9 * 9750 * 10 38 * 24 11 * * 12 22 * 29 12 * * 10 * 136999 * 9 * * 9 * * 13 15 * * 8 * * 9 * 8 * 8
12 * 78 9 * * 9 * * 9 * 15 13 * * 14 14 * * 9 * * 9 * * 8 * * 9 * * 9 * * 9 * * 13 17 *
39 12 * 9 * * 9 * * 14 14 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 10710 * 250 9
* 57 10 * * 9 * 14 14 * * 16 12 * * 9 * 10 * 37 * * 9 * 750 53 * 10 * 9 * 50 12 * 57 10 * 9
* 12 23 * 28 11 * * 9 * * 15 13 * 250 9 * * 19 14 * * 9 * * 12 * 24 30 * 11 * * 9 * * 9 * 9
* 8 * 9 * 13 20 * 42 10 * * 9 * * 9 * * 9 * * 10750 * 8750 * 983 * 12 42 * 11 57 * 10 * * 9 * 8
* 9 * 9 * 9 * 14 13 * * 13 15 * * 22 21 15 * * 13 16 * * 9 * * 9 * * 9 * * 8 * * 9 * * 11
* * 9 * 16 12 * 13 14 * * 19 12 * * 15 14 * * 14 14 * * 13 15 * * 16 13 * * 9 * * 16187 15 * * 12 21 * * 15 14 * * 13 17 *
8 * 15 14 * * 10 35 * 375 14 15 * * 14 14 * * 13 15 * * 13 15 * * 15 13 * * 9 * * 9 * * 9 * * 9 *
8 * 14 16 * 71 9 * 13 15 * * 9 * * 10 * * 8 * * 15 13 * * 8 * * 15 13 * * 9 * * 9 * * 9 * * 9 *
9 * 8 * 15 14 * * 15 14 * * 17 12 * * 13 15 * * 9 * * 11 18 * * 9 * * 8 * * 8 * * 9 * * 9 * * 10
10 34 * 26 11 * * 13 15 * * 13 15 * * 13 28 * 31 10 * * 13 16 * * 18 11 * 300 9 * * 9 * * 9 * 9
8 * 9 * 13 15 * * 13 15 * * 8 * * 13 16 * * 9 * * 9214 * * 9 * * 33 10 * * 8 * * 13 15 * * 23 10 * * 9
11 33 * 28 10 * * 13 15 * * 8 * * 27 12 * * 9 * * 14 14 * * 8 * * 11 * 27 * * 9 * * 9 * * 9 * 9
10 * 9 * 16 14 * * 15 13 * * 9 * * 9166 * 15 13 * * 9 * * 10 * * 9 * * 9 * * 9 * * 9 * 9
* 9 * 10 * 24 11 * * 15 13 * * 22 12 * * 9 * * 9 * * 14 13 * * 10 * * 25 12 * * 11 25 * * 9
9 * 14 14 * * 14 * 14 * * 9 * * 9 * * 9 * * 8 * * 34 11 * * 10 35 * 500 9 * * 14 15 * * 8
13 18 * * 8 * 9 * * 9 * * 8 * * 9 * * 8 * * 9 * * 12 * 9 * * 8 * * 8 * * 9 * * 9 * * 9 * 9
9 * 9 * 9 * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * 9
* 8 * 9 * 9 * 750 9 * * 8 * * 13 15 * * 12 23 * 37 11 * * 9 * * 14 14 * 166 9 * 999 8 * * 8 * * 9 * * 8
9 * 107 17 * 12 * * 9 * * 9 * * 10115 * 18 12 * * 9 * * 9 * * 9 * * 15 12 * * 9 * * 9 * 9
* 9 * 9 * 9 * 9 * * 9 * * 15 14 * * 12 16 * * 9 * * 10 * * 9 * * 8 * * 8 * * 9 * * 9 * 9
* 500 13 * 16 * * 12 16 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * 9
9 * 9 * 12 17 * * 9 * * 9 * * 9 * * 8 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * 9
* 16 13 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * 9
9 * 9 * 9 * 8 * * 9 * * 8 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * * 9 * 9

```

ETAT FINAL (2000 ITERATIONS)



$v =$ 

FORMES CARACTERISTIQUES =



Dans ce cas, (voisinage elliptique) on notera que les bandes inhibées sont à 45 % aux bords du carré, mais s'inclinent à 30 % vers le centre.

EXEMPLE 3-8 VOISINS

Le réseau est de taille 40 * 40

La forme du voisinage à 8 voisins est:
 x x x
 x . x
 x x x

La loi propre des intervalles inter-spikes est uniforme

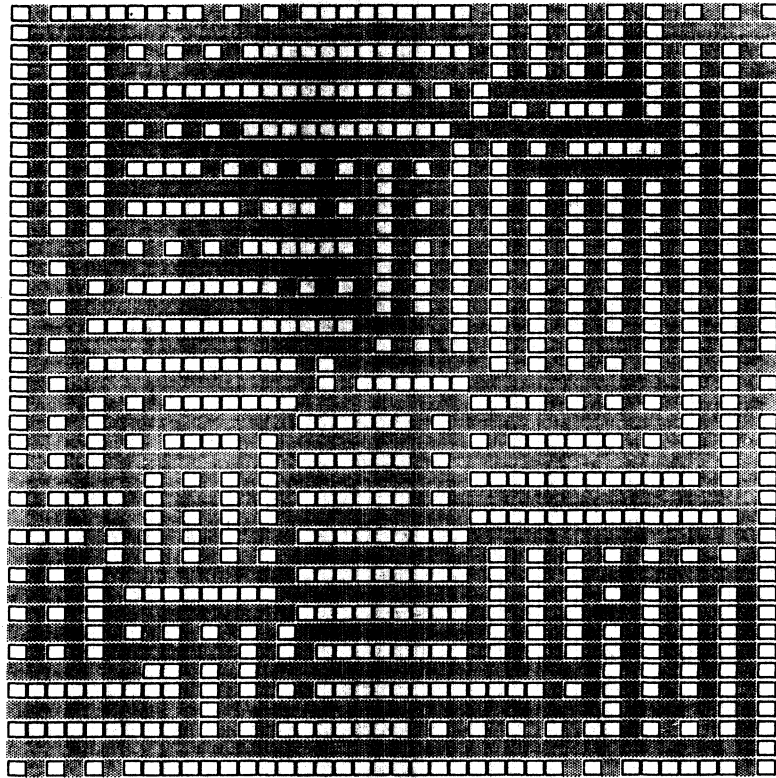
Le paramètre d'inhibition tэта vaut 5

L'intervalle inter-spikes moyen propre est 10

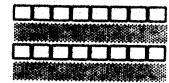
La fréquence moyenne de décharge propre pour 100 unités de temps est: 10

Etat du système au bout de 2000 itérations

11	18	10	2	11	4	3	8	16	1	13	13	7	3	16	11	16	10	20	2	10	12	24	6	46	8	197	8	220	3	208	5	22
231053	23	916	145	816	513	761	745	647	849	643	898	729	827	736	736	739	674	667	863	5381090	254	971	9	898	16	918	6	849	2	91		
14	372	7	142	14	8	9	1	1	12	3	10	13	12	18	5	2	6	1	6	1	9	13	12	219	1	748	5	897	4	936	15	74
26	992	22	925	23	913	125	947	543	901	524	825	519	940	590	909	644	856	597	832	501	927	574	191	135	7	220	10	335	1	385	6	53
4	450	11	546	6	182	5	19	8	4	13	11	7	15	3	2	11	9	5	5	5	27	5	24	3	480	444	489	689	219	804	114	92
8	811	12	760	7	842	621181	3271164	415	945	472	811	607	828	709	769	660	827	438	924	168	388	7	2	3	3	7	3	2	7	4		
7	567	20	651	7	412	5	60	3	70	3	12	6	33	11	38	12	46	11	13	10	14	3	2	112	412	622	699	721	581	601	283	76
11	827	1	698	9	779	601149	649	999	713	624	847	570	921	454	922	512	915	336	916	254	695	4	183	1	37	10	21	8	5	7	2	
19	667	13	747	4	203	13	24	14	4	23	16	33	1	36	12	31	5	66	3	55	2	383	11	936	331097	371102	94	812	45	78		
18	704	4	656	3	774	262	836	876	5201156	1861359	2271279	631299	1721273	81197	191011	3	735	7	489	8	215	1	154	11	19							
12	693	3	673	20	135	6	11	18	12	17	5	44	11	12	3	20	2	207	3	433	13	581	8	817	25	903	161136	311093	5104			
10	578	6	613	3	919	2581085	511	919	735	750	975	568	898	3051173	2041067	31061	22	939	14	734	10	673	6	642	12	576	5	51				
4	957	11	688	28	116	7	38	1	24	10	37	9	6	11	4	4	7	178	2	633	7	707	3	744	11	797	11	884	2	861	7	69
10	437	6	716	61173	3251088	4171012	669	938	741	744	838	4041188	122	852	4	729	10	824	13	755	1	853	4	654	8	661	14	76				
31237	44	758	2	12	6	2	7	9	5	9	6	18	11	19	9	612	91054	4	894	3	718	8	638	5	712	6	813	1	68			
12	428	2	644	821073	571	778	714	719	868	739	807	707	829	497	927	40	599	7	365	7	613	6	732	5	858	6	827	8	790	7	93	
111232	40	637	4	25	2	11	14	24	14	13	2	4	10	8	8	4	497	1021153	8	964	20	748	8	400	5	532	15	606	1	48		
1	352	4	475	1341026	701	759	779	629	866	794	619	925	342	641	241	22	4	5	21	1	186	6	714	21037	16	963	24	939	6108			
101151	13	602	9	24	1	22	14	13	23	8	10	6	10	111	2	488	280	677	856	317	774	27	379	4	178	1	120	8	185	11	21	
4	409	3	567	131020	294	979	512	915	595	871	469	907	59	404	7	72	10	19	8	26	5	8	421	216	915	94	891	241021	38113			
311201	6	835	16	113	2	16	3	16	9	13	5	9	11	343	66	949	3141118	2991062	254	230	14	9	15	12	14	10	32	1	4			
3	409	6	613	7	989	651073	318	924	561	548	717	370	148	1	8	7	8	2	3	8	14	356	150	910	745	617	844	3951036	229119			
261148	11	964	9	484	11	1	20	5	18	6	12	12	333	133	688	711	5151009	3211311	128	667	5	9	17	13	11	6	3	2	5			
3	126	8	137	9	137	29	410	661	365	909	101026	14	616	9	7	13	6	3	15	4	8	661	1011124	453	863	710	663	835	436	98		
42	986	161	754	269	14	35	3	71	13	143	8	287	5	712	1371108	481	804	820	4771113	129	530	2	48	1	26	11	17	12	9			
4	15	2	1	19	2	455	461113	181117	121057	52	491	5	8	10	2	16	5	14	4	438	175	994	554	711	701	763	679	744	61			
73	824	728	5011145	73	753	9	332	5	373	4	472	2	687	951088	475	855	792	619	960	311	377	5	3	8	7	35	8	16	4			
7	6	12	3	110	13	765	311142	21002	341184	50	520	11	5	23	4	14	8	15	14	8	261	355	857	55	845	215	927	447	87			
2281137	460	548	611	15	207	8	242	4	202	9	189	8	503	861067	4401041	638	786	530	724	110	277	1	162	10	235	10	85	4	4			
14	79	14	32	5	16	142	291	570	204	839	671075	74	511	10	34	2	20	1	28	13	13	5	513	261027	26	988	281074	12103				
371364	18	988	8	316	4	5	4	9	19	10	16	16	356	48	938	366	994	549	957	3571093	185	851	2	671	4	625	9	316	7	21		
8	12	10	224	2	825	111009	59	859	301	765	340	97	15	5	5	1	10	9	19	1	8	8	269	11	723	13	667	33	881	15	94	
421375	271046	19	466	4	112	16	128	1	208	12	138	19	588	509	848	690	739	808	4211048	1691142	21133	11062	9	641	4	43						
9	13	14	36	6	377	3	574	11	890	101003	13	633	5	76	15	45	4	6	5	10	15	12	16	7	267	13	330	3	746	9	89	
971247	302	742	141	425	12	186	15	293	16	573	14	915	211096	3301074	630	850	798	731	880	499	913	289	845	105	908	18	664	33	73			
2	30	4	50	16	96	5	148	12	226	8	443	3	205	1	24	3	30	4	37	5	17	1	16	3	9	5	4	53	2	338	4	61
1291103	595	740	567	537	312	464	13	532	8	854	10	998	142	910	459	814	726	823	667	988	4821103	3791046	388	953	514	468	349	176	56			
12	14	11	9	9	8	2	22	11	129	10	37	11	9	2	10	13	15	3	7	1	13	9	13	2	15	8	8	12	9	13	13	
781236	4241124	404	859	382	858	280	786	208	864	413	892	509	768	634	665	747	794	821	862	5381053	4941142	2851154	2081040	195	545	21						
1	18	4	52	11	58	3	33	3	11	3	18	4	13	13	14	8	9	2	3	8	2	18	3	29	16	12	12	7	10	3	5	

ETAT FINAL (2000 ITERATIONS) $v =$ 

FORMES CARACTERISTIQUES =



EXEMPLE 4-12 VOISINS

Le réseau est de taille 40 * 40

La forme du voisinage à 12 voisins est:

```

      x
     x x x
    x x . x x
     x x x
      x
  
```

La loi propre des intervalles inter-spikes est uniforme

Le paramètre d'inhibition teta vaut 5

L'intervalle inter-spikes moyen propre est 10

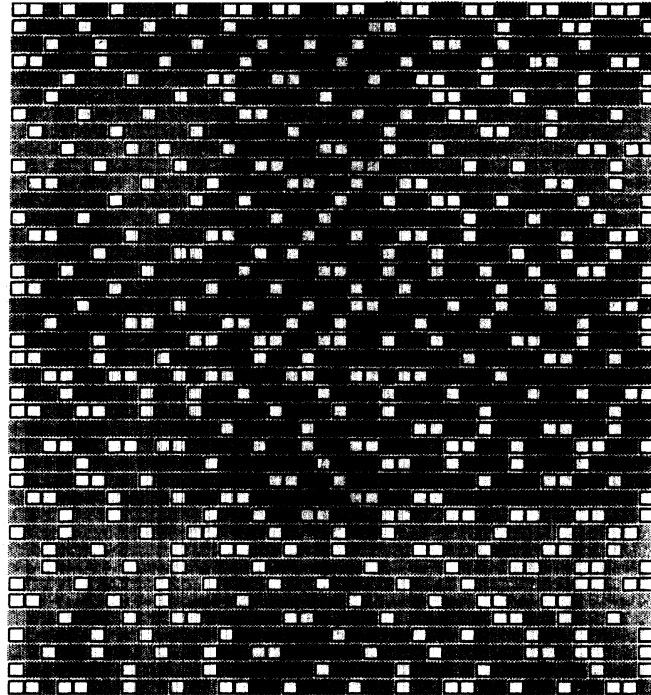
La fréquence moyenne de décharge propre pour 100 unités de temps est: 10

Etat du système au bout de 2000 itérations

```

4 5 389 8 3 189 9 6 629 16 8 572 71 6 22 539 5 15 590 37 6 22 264 279 1 5 376 393 5 8 623 130 1
4 553 508 27 6 556 433 2 12 728 616 180 72 6 617 118 93 421 11 24 618 418 5 8 731 433 7 30 369 438 1 18 61
677 387 3 662 470 9 262 249 10 753 686 4 1921037 10 2 631 510 5 154 573 20 195 39 335 794 77 6 627 901 1 18 44
4 14 664 428 144 6 172 713 408 3 761 545 493 195 11 174 381 27 281 632 4 114 994 507 1 7 870 310 9 7 765 887
14 245 392 3 376 677 6 18 751 642 602 744 6 7 909 936 1 12 728 254 626 869 1 56 873 63 4 824 63 1 414 349 20
271 288 3 275 458 50 14 100 741 789 6 809 279 16 317 158 29 852 9 11 633 679 35 43 339 118 9 18 307 967 299 10 20
6 52 7941256 7 35 932 770 1 438 683 21 464 477 7 9 659 329 68 615 318 10 781 656 4 651007 10 15 492 333 248 28
10 2 335 503 700 621 3 69 444 374 943 10 9 565 89 714 187 4 15 881 312 70 346 9 10 198 95 492 691 4 31574 8
355 567 286 1 887 953 48 8 609 9 93 902 433 15 189 596 128 280 877 13 51301 58 8 896 407 5 387 692 244 374 7
10 220 532 748 88 8 800 172 266 72 6 627 229 287 844 10 41060 293 294 573 5 19 412 125 27 869 407 8 6441096 7 4
446 11 6 228 593 48 686 46 6 993 704 8 7 454 404 459 993 13 14 934 543 13 246 826 6 11 848 446 10 77 52 499 88
70 64 114 18 26 700 6 620 626 428 271 6 415 755 3 568 493 99 432 25 3 689 259 6 444 835 5 111547 2 1 798 37
5 431174 31 4 566 564 746 796 2 188 452 281 52 632 104 1 598 619 5 248 589 24 28 209 104 122 34 16 75 406 424
544 3 2 941 551 651 690 6 751 786 454 658 6 7 913 80 43 12 3 555 694 6 1461192 35 4 963 352 2 251 386 16 83
214 135 133 631 601 16 702 685 6021003 12 111044 572 60 44 756 5 12 610 257 266 44 4 793 358 13 97 247 270 9 1 60
8 633 757 12 651 691 521 863 1 56 988 259 467 587 7 85 537 409 766 10 7 653 348 18 472 532 9 635 382 7 572 456 2
9 7 637 430 706 595 4 347 548 95 725 403 11 3021177 162 1 586 480 147 108 13 319 586 2 322 810 14 82 437 3 129 6
329 18 421 764 4 506 431 198 924 10 12 589 739 127 114 2 586 98 8 4941128 29 15 914 274 223 679 12 223 777 14 411 37
513 491 9 22 444 622 750 4 11 788 324 301 879 7 10 546 345 11 190 427 14 202 379 410 379 4 73 913 525 18 151 195 10
15 227 180 452 209 6 367 399 7 241 785 4 160 612 598 2 6 236 760 4 9 898 582 8 322 756 1 344 255 4611004 15 1
13 16 490 167 3 14 134 515 148 3 24 503 13 178 546 8 191 557 8 105 509 4 118 74 187 233 369 509 16 607 631 509 88
584 394 14 7 4961326 6 8 918 9 41 733 2 20 149 697 707 3 13 785 788 10 61594 26 4 478 617 829 620 1 629 90
12 265 173 3 151 279 256 155 7 349 63 31 437 679 11 348 332 193 467 6 11 534 352 32 1471023 224 9 607 809 718 121
3 10 584 930 11 81178 500 3 349 708 7 286 699 641 420 2 397 979 46 10 944 428 10 137 622 173 478 242 2 510 508 4
316 9 5 62 144 224 24 23 235 24 34 826 517 5 632 819 478 120 6 713 647 12 29 579 792 4 101269 123 9 10 325 92
512 287 9 2 933 686 1 271483 13 11 681 512 649 238 4 708 519 26 412 778 6 7 658 130 364 464 6 381072 12 23 24
5 152 435 111 15 52 413 248 12 218 846 26 2 553 495 5 153 285 527 8 4 845 835 8 11059 666 11 55 687 372 18 1
1 14 553 474 6 31420 123 15 311 497 30 218 23 221198 14 5 843 267 10 215 210 127 369 1 2 880 517 7 36 982 36
568 2 9 416 491 156 8 1531194 31 1 2091122 28 15 230 527 81 9 433 376 15 1 955 650 34 35 783 267 6 12 70 8
59 228 86 6 271 644 8 197 575 10 213 62 17 560 401 21 2131163 13 4 667 114 349 116 10 762 745 1 91195 401 5
13 541042 5 20 137 883 387 1 261109 19 17 307 551 22 50 4 644 309 11 187 546 5 111 629 121 29 200 99 26 226 24
41 11 10 919 615 1 364 609 453 469 4 1381225 3 1 702 705 15 3761064 8 77 293 631 707 3 13 988 701 38 9 837 57
675 151 13 466 590 654 474 10 583 789 7 164 319 219 255 14 210 208 6 13 883 552 18 288 477 279 293 13 4 680 476 49
9 100 896 450 4 592 701 294 27 7 963 523 3 3331353 64 13 785 424 12 497 741 582 426 13 582 657 44 14 298 500 2 2
14 4 321 327 554 205 5 806 353 10 361 411 606 610 2 241 376 31 116 395 483 15 565 492 438 128 6 181 906 8 20 871 38
99 155 263 1 424 946 14 7 161061 22 14 885 505 30 318 670 9 7 943 393 426 420 16 280 238 27 15 179 323 12 42 67
18 15 250 247 132 3 642 102 21 104 139 79 380 15 510 235 26 12 289 390 1 583 482 195 178 5 814 28 3 460 595 3 1
580 184 8 717 675 15 2111103 49 9 283 542 33 457 869 4 7 715 391 15 478 793 15 3831201 103 9 421 529 6 235 556 1
10 238 160 39 11 261 107 1 76 390 482 15 15 49 349 13 41 288 60 3 12 522 478 157 1 152 204 359 26 4 21 330 14
4 12 697 8 2 404 49 2 196 276 11 14 728 18 3 410 377 16 64 825 14 11 344 329 9 110 327 18 22 666 14 9 35
  
```


ETAT FINAL (2000 ITERATIONS)



FORMES CARACTERISTIQUES =



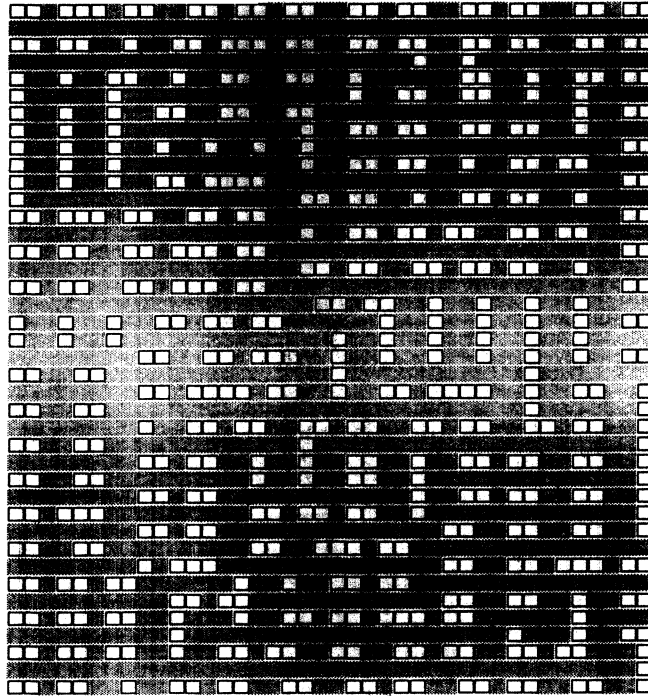
Fréquence de chaque neurone en nb pour 100 unités de temps -le 0 indique les neurones inhibés

Table with 10 columns and 60 rows of numerical data representing neuron frequency. The data shows varying values across rows, with many zeros indicating inhibited neurons.

Intervalle inter-spike moyen de chaque neurone-l'étoile indique les neurones inhibés

Table with 10 columns and 60 rows of numerical data representing the mean inter-spike interval. The data shows varying values across rows, with some asterisks indicating inhibited neurons.

ETAT FINAL (2000 ITERATIONS)



$v =$



FORMES CARACTERISTIQUES =



EXEMPLE 6-24 VOISINS

Le réseau est de taille 40 * 40

La forme du voisinage à 24 voisins est:
 x x x x x
 x x x x x
 x x . x x
 x x x x x
 x x x x x

La loi propre des intervalles inter-spikes est uniforme

Le paramètre d'inhibition tête vaut 5

L'intervalle inter-spikes moyen propre est 10

La fréquence moyenne de décharge propre pour 100 unités de temps est: 10

Etat du système au bout de 2000 itérations

```

12 67 146 3 6 54 22 15 122 29 8 72 83 2 12 47 12 5 9 311 23 17 129 63 7 1 7 18 16 8 13 74 1
9 388 958 161 409 873 397 292 716 379 11 903 665 12 4971131 281 57 379 742 35 17 595 834 95 32310881054 768 658 964 972 5
41211771980 907 94518111307 86916821298 36316461436 31211882001 951 66314531904 813 58317151774 610 86415431374 762 65113411344 44
11 88 804 9 5 496 222 5 718 700 13 615 503 6 61 664 4 16 288 637 25 18 544 562 17 7 115 70 4 4 370 242 1
5 28 562 10 14 492 181 14 422 422 11 570 602 4 226 756 7 7 72 69 3 7 76 208 9 10 6791114 444 4341629 737 2
51 8611526 231 5381430 770 961296 759 1014091465 29313372168 814 78814751435 836 69413481833 954 59815931731 447 4371659 763 3
7712721599 216 9301467 402 6201843 938 29814491129 33 5761306 304 31910281130 693 537 8961266 367 5 581 504 12 8 610 68
8 414 410 8 340 354 6 49 804 26 22 631 264 6 17 249 28 2 13 9 14 5 6 84 11 12 437 857 158 4341652 734 5
36 672 670 17 477 492 12 104 631 9 8 692 943 367 7611741 965 374 7871204 576 204 8781570 771 43417321583 487 7791714 495 1
26618302002 38316081953 55511612076 880 50713871494 417 56917301254 608110018441077 68613271907 750 4181194 545 8 10 426 20
13 9631165 9 7491137 94 4371298 578 219 527 565 5 11 184 118 11 34 564 3 9 88 301 11 18 448 155 11 4411388 441 23
15 68 229 8 31 244 17 15 91 29 1 11 785 205 3010981068 10 116 946 7 10 529 985 39 1301548 889 23612291961 612 42
3311911679 366 9951876 998 56614811416 551 69218911052 22716801688 288 9502259 948 54917632005 639 7211767 672 10 395 796 4
611321614 301 8361698 880 45813871511 614 4331106 185 14 425 548 9 1571175 89 13 696 612 2 6 608 53 17 81 721 4
13 17 283 10 16 138 4 15 10 163 10 15 669 297 1 709 784 16 73 578 10 13 134 176 17 10 653 312 9 4061553 668 37
28 8931467 214 6271653 976 313 9331399 382 1161377 832 211821247 11 8071348 5 5581238 793 862 94710181279 777 3371395 701 1
17914872029 462 86519851118 44211431865 800 4681446 777 1 842 911 1211751657 42311621464 451 815 817 210 758 477 7 606 394
9 362 631 10 4 391 3 7 36 379 3 6 929 600 3211261127 10 632 624 9 240 226 10 199 203 12 314 318 11 884 882 4
12 319 816 4 821142 430 6 478 973 3 13 974 401 6 886 890 13 755 746 10 842 844 10 909 903 1312681255 7916001595 3
23112792127 787 76922581505 47416022030 495 4511299 172 6 633 525 1112251221 7718181817 20218671859 21016741671 1011041062 1
8 3761149 47 6 848 354 13 506 858 31 4431490 619 49118881271 14814991263 31 942 940 8 772 769 4 489 481 1 451 430
6 5 460 2 10 473 213 11 275 299 11 15 847 153 2511342 727 17 776 530 10 395 387 3 421 434 9 702 691 213041284 14
532 81617381094 42912291311 79513971328 343 684 542 9 12 383 15 11 742 504 1813561377 6717511763 26418401817 26916281571 1
367 60613541120 419 831 993 720 796 700 44112381236 41910861891 836 38218941324 12015091473 1712371232 2 998 970 5 685 641
5 5 21 89 13 4 15 92 14 4 16 566 552 33910471699 698 2321344 730 16 473 435 2 110 116 5 163 163 15 662 630
694 703 9431048 179 6 4201227 523 390 772 551 12 14 27 45 1 11 333 19 7 294 314 12 695 704 41215911171 41718141399 16
697 699 9531536 770 23710081813 671 50315081312 332 8181242 783 858 89910701122105411751111 8851247 804 6151498 810 671032 547
11 8 7 596 220 8 315 943 8 8 612 718 16211491540103312881346 948 987 915 322 154 322 40 9 9 210 3 2 317 5 1
648 655 675 891 216 8 171063 63 40 852 442 3 254 262 4 172 307 5 28 326 16 9 450 146 18 7381323 371 7411795 878 26
649 663 687 890 408 5 1891627 670 40118981547 813401334 15 9881113 14 6551397 283 43417871098 32214251779 373 7351464 530 6
8 10 9 75 727 372 2881285 442 710641054 1416771676 14718781994 34314242067 447 5281585 566 18 371 377 12 10 120 2 1
171 9901013 264 945 625 2 769 398 7 256 255 7 327 319 1 663 663 11 309 615 1 9 381 5 1 813 883 24713841571 372 26
32511651162 18 520 184 9 347 933 13 836 960 162 421 365 5 439 218 9 329 638 4 3281149 282 931353 909 24213851566 382 61
2 195 191 1 295 717 32 3641320 330 7131183124610531003 81415431015 43617701542 40115201759 448 252 962 5 2 98 36 7 7
2812971592 14711991892 552 2181203 211 4 27 501 10 9 5961051 474 3401390 678 6 885 604 6 2291340 216 47618521266 138 80
2913091588 23 4241175 6 1 624 301 16 28 665 12 15 273 14 3 3 523 16 4 594 292 14 2241342 204 47918381259 134 82
1 43 350 17 18 431 11 11127 914 85 9541707 88 351025 672 6 711334 406 5414021128 16 361 799 15 14 220 14 4 1
4410751718 382 81620901106 41818081349 37 9781748 327 32715461415 482 79122411129 31315881604 61 8281743 646 54915281325 412 57
1410091640 389 7141724 682 1 998 541 2 161 703 4 14 376 319 9 15 859 141 4 433 678 8 4021364 654 55815351312 410 22
13 6 302 7 3 254 7 1 396 103 13 153 703 10 14 165 138 8 6 545 62 7 237 474 6 15 277 8 4 10 47 7

```

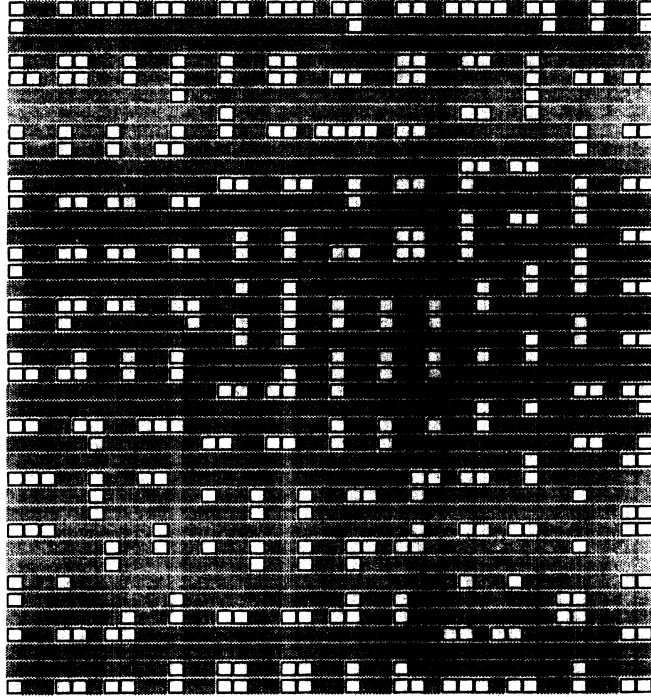
Fréquence de chaque neurone en nb pour 100 unités de temps -le 0 indique les neurones inhibés

Table with 40 columns and 49 rows of numerical data representing neuron frequencies. The first row starts with '7 1 1 6 2 3 6 2 0 5 6 0 1 7 4 0 3 7 3 0 3 5 0 0 7 4 1 3 4 4 5 0 3 7 0 0 8 0 0 7'.

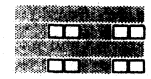
Intervalle inter-spike moyen de chaque neurone-l'étoile indique les neurones inhibés

Table with 40 columns and 49 rows of numerical data representing mean inter-spike intervals. The first row starts with '13 53 78 15 48 28 15 40136 18 16 * 53 13 23 * 28 13 27 * 28 17 * * 12 22 57 32 24 20 19 * 30 14 * * 11 * * 14'.

ETAT FINAL (2000 ITERATIONS)



FORMES CARACTERISTIQUES =



EXEMPLE 7-48 VOISINS

Le réseau est de taille 40 * 40

La forme du voisinage à 48 voisins est: x x x x x x x
x x x x x x x
x x x x x x x
x x x . x x x
x x x x x x x
x x x x x x x
x x x x x x x

La loi propre des intervalles inter-spikes est uniforme

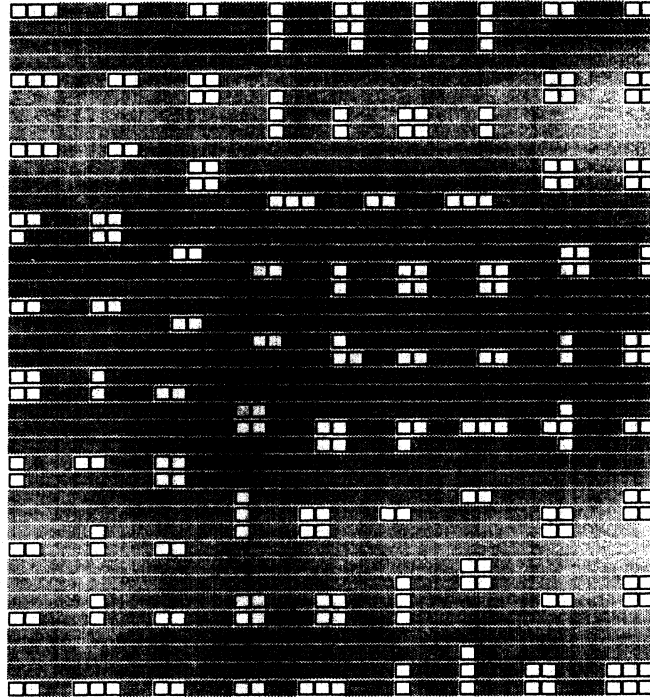
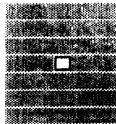
Le paramètre d'inhibition teta vaut 5

L'intervalle inter-spikes moyen propre est 10

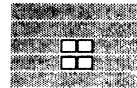
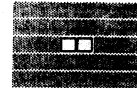
La fréquence moyenne de décharge propre pour 100 unités de temps est: 10

Etat du système au bout de 2000 itérations

Table with 40 columns and 40 rows of numerical data representing the state of the system after 2000 iterations.

ETAT FINAL (2000 ITERATIONS) $v =$ 

FORMES CARACTERISTIQUES =



APPENDICE 3 - CAS DE n^2 NEURONES

$E(\mathcal{F})=60$

DIVERGENCE :

ETUDE DETAILLEE DE DIFFERENTES CARTES STABLES POUR $|V| = 4, 6$ et 8 .

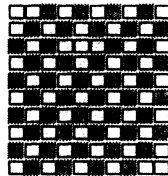
I.1) $|V| = 4$

ETAT INITIAL ALEATOIRE

$\theta = 25 > E(\mathcal{F}) / 4$

ETAT INITIAL										ETAT FINAL									
85	31	25	42	52	109	105	61	117	22	54	4	569	55	305	40	521	22	492	28
34	70	76	44	108	56	87	71	41	99	47	1120	17	1119	6	623	86	940	6	249
10	13	85	18	16	36	33	77	68	44	394	61	1410	10	681	3	754	19	844	37
108	56	38	85	38	93	57	58	67	55	64	1531	85	1085	53	768	7	1103	16	396
70	4	45	99	115	26	81	70	3	46	445	49	1495	74	1165	60	1206	79	1207	5
33	109	47	114	74	66	52	20	88	104	36	1409	1	1239	25	1316	20	1327	12	544
109	18	27	53	97	17	108	25	107	77	559	13	1302	38	972	15	1408	1	1432	28
85	41	100	1	97	8	38	80	37	40	66	1384	46	948	3	876	46	1405	19	519
103	105	29	103	58	68	76	54	102	111	728	40	1142	4	833	78	1132	45	1027	81
13	63	3	75	15	64	67	50	19	53	41	513	1	268	27	330	41	362	13	67

CARTE



Presque tous les neurones sont isolés.

Deux remarques:

- 1) Ne peuvent subsister des amas de m neurones que si $t\theta < E(F)/m$.
- 2) La largeur d'une bande de neurones inhibés est limitée par la portée du voisinage considéré: tout neurone inactif doit être voisin d'un neurone actif, sous peine de bascule dans le sous-ensemble des actifs.

2) $|V| = 4$

$\theta = 70$

ETAT INITIAL ALEATOIRE

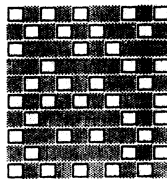
ETAT INITIAL

85	31	25	42	52	109	105	61	117	22
34	70	76	44	108	56	87	71	41	99
10	13	85	18	16	36	33	77	68	44
108	56	38	85	38	93	57	58	67	55
70	4	45	99	115	26	81	70	3	46
33	109	47	114	74	66	52	20	88	104
109	18	27	53	97	17	108	25	107	77
85	41	100	1	97	8	38	80	37	40
103	105	29	103	58	68	76	54	102	111
13	63	3	75	15	64	67	50	19	53

ETAT FINAL

84	5171	4	4692	24	2869	415	50	5187	7
5104	54	4866	34	7068	6	2847	2551	10	3139
2	2633	605	2358	31	4896	50	2837	448	885
348	2816	18	2705	208	403	2817	15	5767	6
2760	25	7215	16	4975	30	2631	3180	23	4976
57	7769	79	4904	58	4996	22	2467	2428	8
5389	3	3207	433	592	1027	2448	31	4267	802
28	5461	410	54	3487	80	978	5710	85	4410
5243	6	2719	483	533	1078	2906	6	7832	14
12	5133	11	2555	80	2614	3	4980	16	2253

CARTE



Pas de paires (sauf sur le bord, moins inhibé).

3) $|V| = 4$

$\theta = 28$

ETAT INITIAL EN BANDE

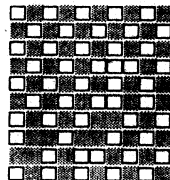
ETAT INITIAL

600	31	600	42	600	109	600	61	600	22
600	70	600	44	600	56	600	71	600	99
600	13	600	18	600	36	600	77	600	44
600	56	600	85	600	93	600	58	600	55
600	4	600	99	600	26	600	70	600	46
600	109	600	114	600	66	600	20	600	104
600	18	600	53	600	17	600	25	600	77
600	41	600	1	600	8	600	80	600	40
600	105	600	103	600	68	600	54	600	111
600	63	600	75	600	64	600	50	600	53

ETAT FINAL

51	38	1036	77	560	12	82	105	392	81
4	1170	43	1208	3	43	672	31	78	171
448	70	1652	20	308	414	7	753	85	92
87	787	89	353	115	19	1792	79	756	97
90	76	420	91	39	1123	37	1264	43	285
106	13	98	20	812	66	1848	16	1260	46
94	427	96	757	5	1254	30	1161	45	433
700	6	1260	62	1792	76	1484	40	1456	8
6	961	17	1323	14	1428	109	1215	9	715
84	85	924	87	1456	13	1176	53	1344	12

CARTE



Les bandes se désagrègent et le réseau tend vers la quinconce limite .

4) $|V| = 4$

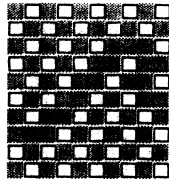
$\theta = 70$

ETAT INITIAL EN BANDE

ETAT INITIAL									
600	31	600	42	600	109	600	61	600	22
600	70	600	44	600	56	600	71	600	99
600	13	600	18	600	36	600	77	600	44
600	56	600	85	600	93	600	58	600	55
600	4	600	99	600	26	600	70	600	46
600	109	600	114	600	66	600	20	600	104
600	18	600	53	600	17	600	25	600	77
600	41	600	1	600	8	600	80	600	40
600	105	600	103	600	68	600	54	600	111
600	63	600	75	600	64	600	50	600	53

ETAT FINAL									
2870	21	4760	61	4620	6	4690	45	5250	109
107	5980	110	6584	27	5056	35	3597	10	4399
4340	20	6790	40	6790	23	1960	737	2170	11
92	5546	47	6765	17	1943	2590	33	1120	2535
4410	21	6860	14	3010	2175	25	4370	840	68
43	6579	13	2594	2520	32	6370	101	3150	344
2940	7	3290	2253	31	5927	14	5935	113	2557
189	911	2940	30	7210	1	6370	14	6300	15
101	3565	94	7063	67	6608	9	5404	67	4481
2450	61	5320	43	5670	37	4900	6	4970	66

CARTE



Plus aucune paire, les isolés actifs suivent la loi F.

Quinconce presque parfaite.

5) $|V| = 4$

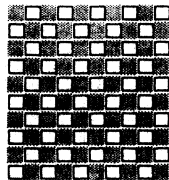
$\theta = 30$

ETAT INITIAL A L'EQUILIBRE NUMERO 1

ETAT INITIAL									
600	31	600	42	600	109	600	61	600	22
34	600	76	600	108	600	87	600	41	600
600	13	600	18	600	36	600	77	600	44
108	600	38	600	38	600	57	600	67	600
600	4	600	99	600	26	600	70	600	46
33	600	47	600	74	600	52	600	88	600
600	18	600	53	600	17	600	25	600	77
85	600	100	600	97	600	38	600	37	600
600	105	600	103	600	68	600	54	600	111
13	600	3	600	15	600	67	600	19	600

ETAT FINAL									
550	25	1390	73	1480	3	1540	90	1330	46
69	2560	8	2290	18	2560	13	2140	13	1540
1870	77	2680	19	2470	24	2350	38	2410	16
22	2890	1	2470	80	2710	37	2350	100	1720
1720	13	2680	10	2680	29	2620	3	2560	62
29	2410	35	2530	10	2920	14	2410	90	1660
1270	30	2410	58	2650	13	2800	21	2650	11
42	2320	49	2530	51	2590	22	2950	61	1930
1450	60	2740	67	2380	27	2710	13	2950	70
68	1780	14	1690	53	1540	54	1750	10	670

CARTE



Stabilité complète de la carte de départ. L'inhibition est supérieure au seuil calculé.

6) $|V| = 4$
 $\theta = 70$

ETAT INITIAL A L'EQUILIBRE NUMERO 1

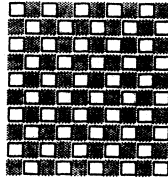
ETAT INITIAL

85	600	25	600	52	600	105	600	117	600
600	70	600	44	600	56	600	71	600	99
10	600	85	600	16	600	33	600	68	600
600	56	600	85	600	93	600	58	600	55
70	600	45	600	115	600	81	600	3	600
600	109	600	114	600	66	600	20	600	104
109	600	27	600	97	600	108	600	107	600
600	41	600	1	600	8	600	80	600	40
103	600	29	600	58	600	76	600	102	600
600	63	600	75	600	64	600	50	600	53

ETAT FINAL

36	5600	57	5670	14	5600	80	5390	66	2800
5390	101	7840	29	7350	3	8260	24	7140	30
35	7700	27	7420	24	7560	19	7560	3	5180
5670	49	7840	91	7490	12	7490	21	7630	33
21	8260	4	7350	9	7490	46	7560	4	5460
6300	40	7700	48	7700	25	7840	101	7630	14
69	8470	32	7630	22	8260	24	8050	62	5180
5810	72	8120	71	8120	60	8610	53	7980	33
41	8190	78	7840	38	8470	81	8470	24	5810
3570	8	5250	25	5320	83	5950	64	6510	21

CARTE



Stabilité complète de la carte de départ. L'inhibition est supérieure au seuil calculé.

7) $|V| = 4$
 $\theta = 25$

ETAT INITIAL A L'EQUILIBRE NUMERO 2

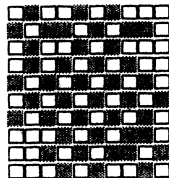
ETAT INITIAL

85	600	600	42	600	600	105	600	600	22
600	70	600	600	108	600	600	71	600	600
600	600	85	600	600	36	600	600	68	600
108	600	600	85	600	600	57	600	600	55
600	4	600	600	115	600	600	70	600	600
600	600	47	600	600	66	600	600	88	600
109	600	600	53	600	600	108	600	600	77
600	41	600	600	97	600	600	80	600	600
600	600	29	600	600	68	600	600	102	600
13	600	600	75	600	600	67	600	600	53

ETAT FINAL

46	425	53	4	1300	13	1326	74	674	10
950	37	13	1623	6	5275	80	2003	19	556
49	79	1030	92	4232	10	4800	41	1428	28
104	1465	28	4350	10	4450	41	4100	57	796
925	17	4275	87	4456	27	4225	20	4575	60
62	4700	38	4325	13	3842	1	4550	10	1528
1181	69	3500	36	4500	58	3998	41	2310	13
75	2483	14	3850	46	4525	26	1162	27	14
1169	14	2647	8	4700	14	1881	14	39	750
20	1304	73	1250	63	1025	43	71	750	21

CARTE



La carte se désagrège lentement, l'inhibition est trop faible.

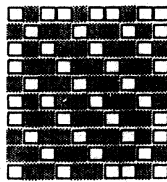
8) $|V| = 4$ $\theta = 35$

ETAT INITIAL A L'EQUILIBRE NUMERO 2

ETAT INITIAL									
85	600	600	42	600	600	105	600	600	22
600	70	600	600	108	600	600	71	600	600
600	600	85	600	600	36	600	600	68	600
108	600	600	85	600	600	57	600	600	55
600	4	600	600	115	600	600	70	600	600
600	600	47	600	600	66	600	600	88	600
109	600	600	53	600	600	108	600	600	77
600	41	600	600	97	600	600	80	600	600
600	600	29	600	600	68	600	600	102	600
13	600	600	75	600	600	67	600	600	53

ETAT FINAL									
3	1715	21	44	770	18	94	875	58	238
1470	11	1540	315	15	1365	455	64	1890	63
28	1330	30	770	910	46	1295	1085	95	1365
24	805	770	38	805	1295	82	1050	980	32
1085	87	1120	805	1	700	1050	81	1225	16
50	1400	62	840	700	20	665	1015	105	630
33	595	840	21	805	665	6	700	420	6
665	56	1050	805	50	770	700	20	1505	10
53	1225	4	1330	770	64	420	1225	5	1750
79	59	1470	6	35	420	57	96	1435	2

CARTE



Stabilité, seulement troublée par l'inhibition plus faible des bords.

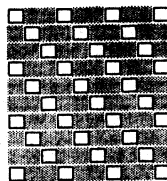
9) $|V| = 4$ $\theta = 70$

ETAT INITIAL A L'EQUILIBRE NUMERO 2

ETAT INITIAL									
85	600	600	42	600	600	105	600	600	22
600	70	600	600	108	600	600	71	600	600
600	600	85	600	600	36	600	600	68	600
108	600	600	85	600	600	57	600	600	55
600	4	600	600	115	600	600	70	600	600
600	600	47	600	600	66	600	600	88	600
109	600	600	53	600	600	108	600	600	77
600	41	600	600	97	600	600	80	600	600
600	600	29	600	600	68	600	600	102	600
13	600	600	75	600	600	67	600	600	53

ETAT FINAL									
9	4060	770	15	2940	1050	40	3500	840	74
4060	27	3220	2940	49	3780	3500	6	3500	840
980	3220	14	2870	3780	24	4340	3500	12	3640
2	3500	2870	2	2800	4340	9	3430	3640	12
3500	9	3570	2800	38	2940	3430	13	3080	1260
1120	3570	11	3150	2940	54	2800	3080	52	2590
62	3570	3150	36	2870	2800	30	2940	2590	93
3570	60	3500	2870	14	2520	2940	35	3290	420
1750	3500	6	3360	2520	13	2660	3290	79	3500
71	1750	3360	53	1050	2660	53	1050	3500	71

CARTE



Stabilité complète de la carte de départ. L'inhibition est supérieure au seuil calculé.

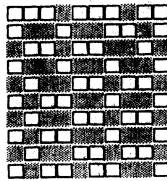
II. 1) $|V| = 6$ $\theta = 20$

ETAT INITIAL ALEATOIRE

ETAT INITIAL									
85	31	25	42	52	109	105	61	117	22
34	70	76	44	108	56	87	71	41	99
10	13	85	18	16	36	33	77	68	44
108	56	38	85	38	93	57	58	67	55
70	4	45	99	115	26	81	70	3	46
33	109	47	114	74	66	52	20	88	104
109	18	27	53	97	17	108	25	107	77
85	41	100	1	97	8	38	80	37	40
103	105	29	103	58	68	76	54	102	111
13	63	3	75	15	64	67	50	19	53

ETAT FINAL									
67	32	68	262	35	50	259	64	38	17
3	590	79	6	351	850	13	123	568	21
19	36	101	1298	67	69	1317	46	52	197
28	1056	57	24	427	433	36	202	1447	5
150	33	156	1659	80	11	2201	8	67	606
43	1469	29	37	732	97	46	279	1148	65
412	57	93	1433	47	51	1888	2	25	377
4	1101	65	60	917	114	36	360	434	20
393	50	4	843	24	6	1216	29	28	311
25	105	69	100	35	27	37	176	36	48

CARTE



Une majorité de paires de neurones actifs et voisins.

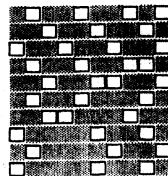
2) $|V| = 6$ $\theta = 50$

ETAT INITIAL ALEATOIRE

ETAT INITIAL									
85	31	25	42	52	109	105	61	117	22
34	70	76	44	108	56	87	71	41	99
10	13	85	18	16	36	33	77	68	44
108	56	38	85	38	93	57	58	67	55
70	4	45	99	115	26	81	70	3	46
33	109	47	114	74	66	52	20	88	104
109	18	27	53	97	17	108	25	107	77
85	41	100	1	97	8	38	80	37	40
103	105	29	103	58	68	76	54	102	111
13	63	3	75	15	64	67	50	19	53

ETAT FINAL									
985	70	1332	3142	9	421	412	26	1417	15
87	4120	1376	1	3208	456	37	490	40	2899
1086	8	3035	3418	77	70	1183	927	1568	6
8	3206	26	15	2788	2593	108	105	1967	455
1720	10	1545	3349	765	8	3431	4220	65	246
16	3809	2197	38	39	5366	402	3	3438	101
3209	104	54	4253	1447	10	2958	3275	36	1397
41	3891	1950	8	2697	2858	10	2980	2701	53
2053	33	2229	2853	13	2968	2926	64	1602	1811
45	1513	59	825	2215	15	1267	3450	8	203

CARTE



Moins de paires, les actifs sont plus souvent isolés.

3) $|V| = 6$

ETAT INITIAL ALEATOIRE

$\theta = 70 > E(\mathcal{F})$

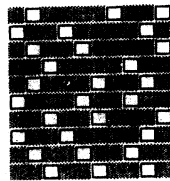
ETAT INITIAL

85	31	25	42	52	109	105	61	117	22
34	70	76	44	108	56	87	71	41	99
10	13	85	18	16	36	33	77	68	44
108	56	38	85	38	93	57	58	67	55
70	4	45	99	115	26	81	70	3	46
33	109	47	114	74	66	52	20	88	104
109	18	27	53	97	17	108	25	107	77
85	41	100	1	97	8	38	80	37	40
103	105	29	103	58	68	76	54	102	111
13	63	3	75	15	64	67	50	19	53

ETAT FINAL

1585	31	512	3082	292	279	17	2821	5257	84
14	5350	1856	103	3148	786	2637	451	18	5519
1305	89	5085	4388	28	535	553	3187	2968	49
1468	3586	39	465	2868	3623	743	2	3177	3095
2270	15	2315	2789	3155	36	3331	7800	46	566
73	2519	4837	354	33	7376	2462	77	2778	2724
5249	512	55	5403	4677	88	278	4745	3147	16
19	6021	5520	90	2927	2978	812	3190	4	3080
3143	35	4889	5383	14	4648	2486	73	5802	491
603	3243	4	3045	4665	23	3247	5470	53	853

CARTE



Plus aucune paire, les isolés actifs suivent la loi F.

4) $|V| = 6$

ETAT INITIAL A L'EQUILIBRE NUMERO 1

$\theta = 20 > E(\mathcal{F})/4$ et $< E(\mathcal{F})$

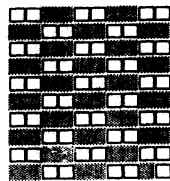
ETAT INITIAL

85	31	600	600	52	109	600	600	117	22
600	600	76	44	600	600	87	71	600	600
10	13	600	600	16	36	600	600	68	44
600	600	38	85	600	600	57	58	600	600
70	4	600	600	115	26	600	600	3	46
600	600	47	114	600	600	52	20	600	600
109	18	600	600	97	17	600	600	107	77
600	600	100	1	600	600	38	80	600	600
103	105	600	600	58	68	600	600	102	111
600	600	3	75	600	600	67	50	600	600

ETAT FINAL

89	85	720	480	85	69	600	520	19	62
360	740	70	41	1100	1060	7	61	500	40
6	118	1320	1180	37	90	1200	1120	19	41
240	660	113	89	1080	1280	30	34	540	200
45	57	1220	1220	32	47	1320	1080	31	4
300	560	38	18	1140	1400	3	34	620	63
26	73	1040	1020	77	35	1220	880	3	3
240	620	99	51	1000	1120	23	17	560	35
62	32	1020	920	60	5	1100	860	71	36
45	140	35	13	360	400	2	54	220	74

CARTE



Stabilité complète de la carte de départ. L'inhibition est supérieure au seuil calculé.

5) $|V| = 6$

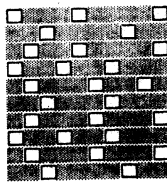
$\theta = 70 > E(\mathcal{F})$

ETAT INITIAL A L'EQUILIBRE NUMERO 1

ETAT INITIAL									
85	31	600	600	52	109	600	600	117	22
600	600	76	44	600	600	87	71	600	600
10	13	600	600	16	36	600	600	68	44
600	600	38	85	600	600	57	58	600	600
70	4	600	600	115	26	600	600	3	46
600	600	47	114	600	600	52	20	600	600
109	18	600	600	97	17	600	600	107	77
600	600	100	1	600	600	38	80	600	600
103	105	600	600	58	68	600	600	102	111
600	600	3	75	600	600	67	50	600	600

ETAT FINAL									
15	560	5040	1890	4	349	910	2800	217	43
2310	3780	15	284	5670	1120	287	6	770	5040
1515	19	5810	6440	66	766	3780	2940	378	28
12	7140	2031	73	6440	5740	15	438	4130	910
1570	17	3500	5670	425	27	8190	4480	10	1266
910	6930	25	424	2870	5670	64	384	4060	1330
3849	65	4690	7280	617	29	5250	3710	487	19
93	8610	2230	49	3290	7700	1	320	1190	2660
2323	50	3080	5390	269	28	3360	5040	225	21
910	3150	2	544	1120	3220	377	76	1260	2590

CARTE



Plus aucune paire, les isolés actifs suivent la loi F.

6) $|V| = 6$

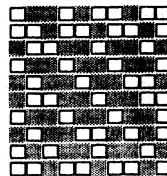
$\theta = 16 < E(\mathcal{F})/3$

ETAT INITIAL A L'EQUILIBRE NUMERO 2

ETAT INITIAL									
85	600	600	42	600	600	105	600	600	22
600	70	600	600	108	600	600	71	600	600
600	600	85	600	600	36	600	600	68	600
108	600	600	85	600	600	57	600	600	55
600	4	600	600	115	600	600	70	600	600
600	600	47	600	600	66	600	600	88	600
109	600	600	53	600	600	108	600	600	77
600	41	600	600	97	600	600	80	600	600
600	600	29	600	600	68	600	600	102	600
13	600	600	75	600	600	67	600	600	53

ETAT FINAL									
19	37	42	37	11	11	100	33	8	81
36	55	34	49	44	224	51	43	56	13
35	8	11	288	41	80	496	96	69	49
46	256	1	57	160	144	60	41	464	20
40	25	208	16	43	20	224	59	35	20
83	5	68	144	67	20	41	53	23	109
7	175	48	1	37	224	10	3	88	13
9	50	44	192	90	66	96	61	2	9
42	17	26	8	96	36	49	32	62	85
39	30	160	34	42	31	91	17	7	103

CARTE



La carte se désagrège lentement, l'inhibition est trop faible.

ETAT INITIAL A L'EQUILIBRE NUMERO 2

7) $|U| = 6$

$\theta = 25 > E(\mathcal{F})/3$

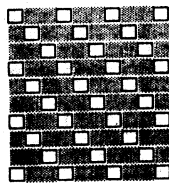
ETAT INITIAL

85	600	600	42	600	600	105	600	600	22
600	70	600	600	108	600	600	71	600	600
600	600	85	600	600	36	600	600	68	600
108	600	600	85	600	600	57	600	600	55
600	4	600	600	115	600	600	70	600	600
600	600	47	600	600	66	600	600	88	600
109	600	600	53	600	600	108	600	600	77
600	41	600	600	97	600	600	80	600	600
600	600	29	600	600	68	600	600	102	600
13	600	600	75	600	600	67	600	600	53

ETAT FINAL

9	1250	425	15	1000	250	40	1100	250	37
550	62	1075	975	49	1350	1125	6	350	300
300	250	73	1150	1275	93	1475	1400	12	400
8	1075	975	2	1000	1300	9	1250	1350	12
375	79	1150	1025	5	925	1050	13	200	375
300	375	11	1000	1000	54	925	975	52	1
24	1100	1050	36	825	875	30	725	700	17
350	60	1125	1025	14	825	900	35	300	25
525	350	6	950	850	13	900	950	71	375
53	550	1350	23	275	850	53	400	1225	71

CARTE



Stabilité, seulement troublée par l'inhibition plus faible des bords.

8) $|U| = 6$

$\theta = 50 > E(\mathcal{F})/2$

ETAT INITIAL A L'EQUILIBRE NUMERO 3

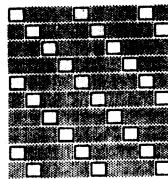
ETAT INITIAL

85	600	600	600	52	600	600	600	117	600
600	70	600	600	600	56	600	600	600	99
600	600	85	600	600	600	33	600	600	600
600	600	600	85	600	600	600	58	600	600
70	600	600	600	115	600	600	600	3	600
600	109	600	600	600	66	600	600	600	104
600	600	27	600	600	600	108	600	600	600
600	600	600	1	600	600	600	80	600	600
103	600	600	600	58	600	600	600	102	600
600	63	600	600	600	64	600	600	600	53

ETAT FINAL

13	1950	2250	500	43	2050	2050	250	63	1800
1950	1	2000	1750	2050	28	1700	1800	1800	74
400	2000	74	2000	1900	1700	15	1700	150	250
100	250	2000	86	1650	1800	1700	22	2050	150
21	1900	1500	1650	20	1800	1900	2050	48	2000
1900	97	1900	2250	1800	7	2450	1950	2000	22
100	1900	57	2100	2100	2450	14	2200	600	100
300	600	2100	16	2300	2200	2200	67	1800	200
71	2050	2050	2300	74	2050	2000	1800	20	2150
2050	15	350	2050	2050	71	300	2200	2150	62

CARTE



Stabilité complète de la carte de départ. L'inhibition est supérieure au seuil calculé.

9) $|U| = 6$

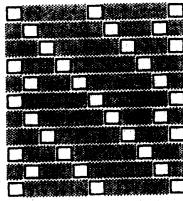
ETAT INITIAL A L'EQUILIBRE NUMERO 4

 $\theta = 70 > E(\mathcal{F})$

ETAT INITIAL										
85	600	600	600	600	109	600	600	600	600	600
600	70	600	600	600	600	87	600	600	600	600
600	600	85	600	600	600	600	77	600	600	600
600	600	600	85	600	600	600	600	67	600	600
600	600	600	600	115	600	600	600	600	46	600
33	600	600	600	600	66	600	600	600	600	600
600	18	600	600	600	600	108	600	600	600	600
600	600	100	600	600	600	600	80	600	600	600
600	600	600	103	600	600	600	600	102	600	600
600	600	600	600	15	600	600	600	600	53	600

ETAT FINAL										
24	3290	1050	1050	1050	33	3640	1050	910	2520	54
3290	4	3430	840	1260	3640	19	5250	2800	52	4130
2800	3430	37	3990	1260	1260	3780	10	5110	4340	67
28	4620	5530	31	3570	1330	980	3500	13	3430	2660
4340	78	2660	5250	47	3290	980	1050	3430	13	3500
22	5740	1120	980	3290	50	3430	980	1050	5110	95
3990	83	3780	1540	1050	3430	5	4760	2730	74	4340
2450	3780	104	3080	980	630	3080	40	4480	3850	20
44	3780	4480	36	3080	770	910	2940	36	3010	2520
3710	76	2590	4690	68	2940	910	770	3010	30	3360
40	2380	910	700	2940	37	700	700	1190	3360	46

CARTE



Stabilité, seulement troublée par
l'inhibition plus faible des bords.

ETAT INITIAL A L'EQUILIBRE NUMERO 1

III. 1) $|V| = 8$

$\theta = 20 < E(F)/2$

ETAT INITIAL

600	600	600	600	600	600	600	600	600	600
34	70	76	44	108	56	87	71	41	99
600	600	600	600	600	600	600	600	600	600
108	56	38	85	38	93	57	58	67	55
600	600	600	600	600	600	600	600	600	600
33	109	47	114	74	66	52	20	88	104
600	600	600	600	600	600	600	600	600	600
85	41	100	1	97	8	38	80	37	40
600	600	600	600	600	600	600	600	600	600
13	63	3	75	15	64	67	50	19	53

ETAT FINAL

12	180	51	75	31	28	34	23	220	52
1	137	45	51	3	82	46	83	54	56
320	1080	900	760	700	700	740	620	1180	440
84	36	85	105	1	39	14	8	37	25
480	1200	1180	1100	1060	940	920	820	1200	340
30	66	49	45	33	40	62	3	63	80
360	1140	1040	820	1060	920	1280	920	1240	340
6	36	58	43	53	4	18	70	90	40
260	1020	1020	800	1140	960	1320	940	1260	380
7	32	73	67	4	86	75	83	34	33

2) $|V| = 8$

$\theta = 40 > E(F)/2$

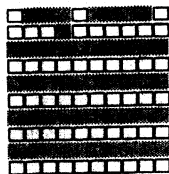
160	280	800	160	1440	12	1160	28	1200	4
67	79	389	66	948	57	647	84	721	48
1720	2680	2280	2360	2600	1800	2440	1280	3240	1440
9	650	19	520	28	115	36	64	507	13
1520	3680	1480	3120	1520	2120	2240	1680	3640	1440
34	709	54	434	62	84	132	14	768	15
1600	2840	2440	2320	2760	1480	2920	1520	3640	1200
10	2	740	9	617	7	519	12	502	49
1480	2920	2640	3000	2640	2840	2280	2600	2520	1520
36	783	93	1115	65	1104	60	610	58	54

3) $|V| = 8$

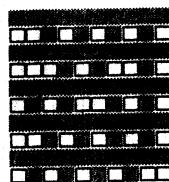
$\theta = 70 > E(F)$

840	980	1120	1120	3360	1050	2520	770	2030	630
44	450	611	40	2868	4	2007	4	1520	107
3010	5460	3500	5530	5320	3080	4760	3010	6580	3220
25	2536	85	2495	46	263	333	63	2757	17
3080	7420	3080	7490	3220	3290	3290	3220	7630	3290
14	2449	63	2664	47	697	502	12	2568	31
3290	5320	5530	5530	5670	3290	3080	3850	5950	3570
535	25	2930	9	2577	44	138	950	7	910
3360	5810	5810	6160	5880	6230	3430	6790	4130	4130
7	2893	3	3465	4	3734	28	3440	69	783

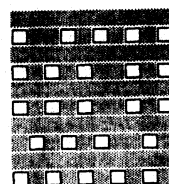
CARTE



Stabilité complète de la carte de départ. L'inhibition est supérieure au seuil calculé.



La carte se désagrège lentement,



Plus aucune paire, les isolés actifs suivent la loi F.

ETAT INITIAL A L'EQUILIBRE NUMERO 2

4) $|V| = 8$

$\theta = 20 < E(\mathcal{F})/2$

ETAT INITIAL										ETAT FINAL									
600	600	600	600	600	600	600	600	600	600	29	105	180	39	100	48	320	20	500	49
600	70	600	44	600	56	600	71	600	99	60	45	300	39	380	25	520	1	620	33
600	600	600	600	600	600	600	600	600	600	280	280	1040	74	1100	63	1180	27	1180	33
600	56	600	85	600	93	600	58	600	55	47	41	40	60	200	5	160	16	80	63
600	600	600	600	600	600	600	600	600	600	340	520	1040	20	1140	100	1160	29	1020	84
600	109	600	114	600	66	600	20	600	104	66	67	27	87	41	29	43	13	86	99
600	600	600	600	600	600	600	600	600	600	200	480	1000	160	1200	180	1180	89	980	61
600	41	600	1	600	8	600	80	600	40	5	36	75	34	2	36	112	75	33	26
600	600	600	600	600	600	600	600	600	600	240	420	1100	120	1380	160	1320	109	1160	35
600	63	600	75	600	64	600	50	600	53	51	52	17	2	40	2	93	48	92	18

5) $|V| = 8$

$\theta = 40 > E(\mathcal{F})/2$

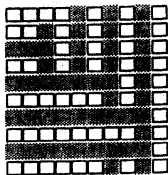
58	2	1920	27	2760	106	3080	15	3360	74
35	41	1920	8	2760	71	3080	32	3360	53
2320	2320	6800	1560	7000	1040	6920	1520	6880	1080
79	82	2520	22	2520	39	2160	76	2120	92
3560	3560	8360	2240	8560	2120	8360	2040	8640	2560
71	56	1720	76	1840	3	2080	63	2480	11
3200	3200	7680	2120	8600	2240	8360	2040	8800	2760
34	45	1720	3	2600	94	2240	80	2360	73
3120	3120	7480	2600	9200	2480	8720	2120	9000	2560
41	59	1600	60	2400	86	2160	60	2320	15

6) $|V| = 8$

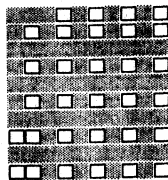
$\theta = 70 > E(\mathcal{F})/2$

1330	1330	3430	770	2800	630	2800	910	3360	1050
1330	42	3430	18	2800	59	2800	21	3360	11
4130	4130	8610	3220	7910	3430	7700	3080	7840	3500
1400	73	3850	89	3710	5	3710	81	3360	95
3710	3710	8260	3220	8260	3780	8540	3430	7770	3290
910	27	3080	16	3150	33	3710	16	3500	20
3010	3010	7700	3290	7980	3360	7840	3290	7700	3290
700	1	3220	8	3500	28	2940	12	3010	40
2940	2940	7840	3570	8190	3500	7700	3010	7420	3220
840	42	3290	6	3500	79	3570	49	3220	29

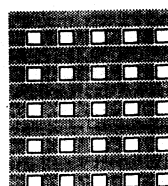
CARTE



La carte se désagrège lentement, l'inhibition est trop faible.



Stabilité, seulement troublée par l'inhibition plus faible des bords.

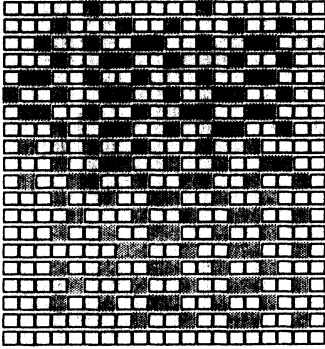


Stabilité complète de la carte de départ. L'inhibition est supérieure au seuil calculé.

ETAT FINAL (2000 ITERATIONS)

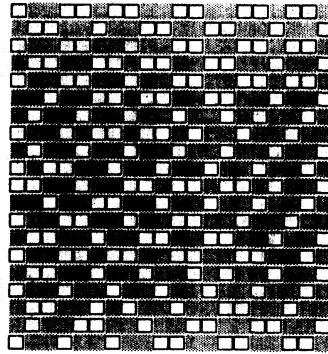
$$|v| = 6$$

$$E(\mathcal{F}) = 10$$



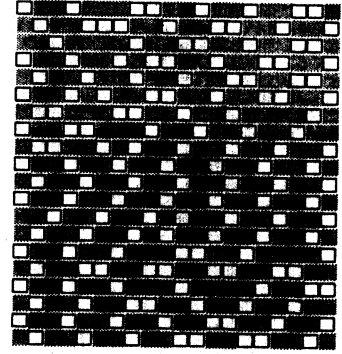
$$\theta = 3$$

Une majorité de paires de neurones actifs et voisins.

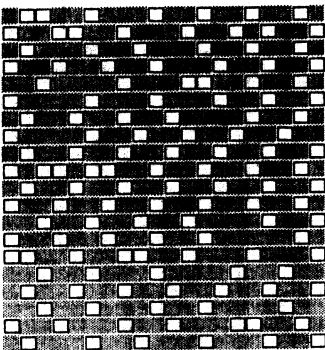


$$\theta = 5$$

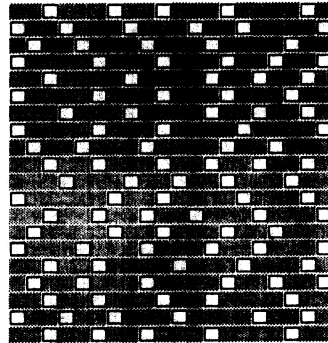
Moins de paires, les actifs sont plus souvent isolés.



$$\theta = 7$$



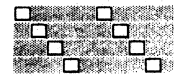
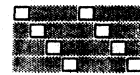
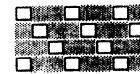
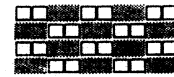
$$\theta = 10$$



$$\theta = 15$$

Plus aucune paire, les isolés actifs suivent la loi F.

FORMES CARACTERISTIQUES =



APPENDICE 4 - CAS DE n^2 NEURONES

REGIME STATIONNAIRE $\theta < \theta_0$.

On donne ici quelques exemples de processus convergents pour $\theta < \theta_0$: état, fréquences, intervalles inter-spikes moyens des neurones d'un réseau.

Pour quelques couples de neurones on a calculé les 10 premières auto-corrélations, et l'inter-corrélation entre les nombres de décharges observées dans des fenêtres de temps simultanées.

Les corrélations des neurones en interactions inhibitrices sont négatives, et d'autant plus grandes en valeur absolue que θ est grand.

Les corrélations des neurones qui n'interagissent pas sont en général faibles, comme le sont les auto-corrélations.

Intervalle inter-spike moyen de chaque neurone-(l'étoile indique les neurones inhibés)

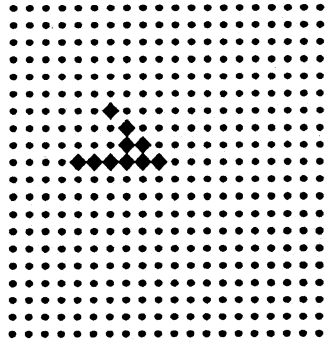
```

11 15 16 15 14 16 14 14 15 16 13 14 16 14 15 14 16 15 14 13
15 16 18 17 16 19 15 17 18 16 19 16 16 17 17 17 16 16 17 14
13 16 15 18 17 15 19 15 17 16 16 17 16 14 17 16 15 15 17 15
15 15 18 16 18 16 16 15 16 16 15 15 18 17 15 15 18 16 17 13
15 16 16 16 15 16 18 17 16 17 19 18 16 16 17 16 17 17 17 15
13 17 15 17 16 18 16 15 17 16 16 15 16 17 15 16 14 18 15 16
13 17 15 17 16 16 17 16 16 15 19 17 17 15 17 15 17 15 19 13
15 17 15 15 19 16 15 17 15 18 16 15 16 15 17 16 16 15 17 14
14 16 18 16 15 17 16 17 18 17 15 16 18 17 17 16 16 16 17 15
16 16 14 19 17 16 15 16 15 20 17 16 16 14 18 16 19 15 17 13
14 18 16 18 15 17 16 17 18 15 16 18 15 18 15 16 15 17 16 17
14 16 15 17 16 16 16 16 16 17 16 15 18 14 17 18 14 18 13
13 17 15 16 17 17 15 15 17 17 16 18 15 17 15 15 18 17 15 15
15 16 18 17 15 16 16 15 16 15 16 15 17 16 17 16 16 18 17 16
13 17 16 15 20 16 15 17 15 17 17 16 15 17 17 15 17 15 16 14
14 17 16 16 16 15 17 14 19 16 15 15 18 15 17 17 16 16 16 15
14 16 18 16 17 18 15 16 15 17 17 15 16 16 17 16 16 19 14
14 15 17 16 16 16 16 17 16 17 14 16 15 16 16 16 15 16 16 14
15 17 15 20 17 17 18 16 17 19 16 18 17 16 16 16 17 19 16 15
13 15 16 14 15 14 14 14 15 14 15 15 14 14 16 14 14 15 14 13

```

Tous les neurones sont actifs,leur intervalle moyen inter-spike est de l'ordre de $E(F)+4\theta$

Les neurones observés sont marqués



Intervalles moyens inter-spikes

Coordonnées	Intervalle	Variance
(10,10)	19.9	139.09
(10, 9)	15.4	107.74
(10, 8)	16.6	117.03
(10, 7)	16.0	149.35
(10, 6)	16.0	134.65
(10, 5)	17.6	128.22
(9, 9)	17.9	147.58
(9, 8)	17.5	121.71
(8, 8)	17.3	170.11
(7, 7)	17.7	181.39

10 premières autocorrélations des neurones

Coordonnées	Autocorrélations									
(10,10)	-.13	-.03	-.12	.01	.04	.10	-.09	.04	-.23	-.06
(10, 9)	.01	-.03	-.12	-.15	.05	-.09	.00	-.20	.07	-.19
(10, 8)	.04	-.01	-.04	-.05	-.02	.01	-.01	-.02	-.12	-.04
(10, 7)	-.19	.02	-.07	.00	.16	-.12	-.15	.06	-.05	-.14
(10, 6)	-.12	-.09	.16	.08	.06	.11	-.10	.11	.13	-.03
(10, 5)	-.03	-.22	.04	.19	-.03	-.21	.04	-.16	-.11	-.01
(9, 9)	.00	.13	.15	.13	.04	-.03	-.14	-.05	-.07	-.14
(9, 8)	.02	.00	-.03	-.04	.15	-.01	-.10	.08	-.07	.04
(8, 8)	-.06	-.11	-.10	.04	.06	-.14	.10	-.02	.08	.07
(7, 7)	.18	.05	-.21	-.04	.04	.08	-.08	-.08	-.05	-.06

Nombre moyen de spikes par fenêtre de 100

Coordonnées	Nombre	Variance
(10,10)	5.0	1.47
(10, 9)	6.5	2.35
(10, 8)	6.0	3.68
(10, 7)	6.3	3.35
(10, 6)	6.1	1.46
(10, 5)	5.7	1.48
(9, 9)	5.5	3.93
(9, 8)	5.7	1.80
(8, 8)	5.7	2.09
(7, 7)	5.7	5.48

Matrice d'intercorrélations des neurones observés

(10,10)	1.00									
(10, 9)	-.42	1.00								
(10, 8)	.11	-.21	1.00							
(10, 7)	.14	-.18	-.43	1.00						
(10, 6)	.00	-.08	.20	-.20	1.00					
(10, 5)	-.11	.31	-.18	.25	-.23	1.00				
(9, 9)	.15	-.13	-.40	.32	-.48	.00	1.00			
(9, 8)	-.32	.20	-.47	-.18	-.21	-.13	.12	1.00		
(8, 8)	.03	-.11	.25	.28	.20	.07	-.34	-.61	1.00	
(7, 7)	.06	-.38	.30	.08	.35	-.11	-.05	-.45	.29	1.00

(10,10) (10, 9) (10, 8) (10, 7) (10, 6) (10, 5) (9, 9) (9, 8) (8, 8) (7, 7)

REMARQUES:

- a) Les autocorrélations sont faibles dans l'ensemble, (rôle important de l'innovation due au tirage indépendant de la v.a. U de loi F après chaque décharge).
- b) Les coefficients de corrélations d'activité les plus importants correspondent presque tous aux corrélations négatives entre voisins directs, qui s'inhibent: par exemple (8,8)-(9,8), (10,10)-(10,9), (10,8)-(10,7), (10,8)-(9,8)...
- c) Sauf sur les bords, où les neurones sont moins inhibés (moins de voisins), l'intervalle moyen inter-spike est de l'ordre de $E(F)+4\theta$.

EXEMPLE 1 BIS---4 VOISINS---REGIME PRESQUE STATIONNAIRE

Le réseau est de taille 20 * 20

La forme du voisinage à 4 voisins est: x
 x . x
 x

La loi propre des intervalles inter-spikes est uniforme

Le paramètre d'inhibition t \acute{e} ta vaut 3 (>t \acute{e} ta z \acute{e} ro)

L'intervalle inter-spikes moyen propre est 10

La fr \acute{e} quence moyenne de d \acute{e} charge propre pour 100 unit \acute{e} s de temps est: 10

Etat du syst \acute{e} me au bout de 2000 it \acute{e} rations

13	9	12	6	6	13	7	11	8	15	14	5	9	4	2	17	16	17	11	16
2	5	4	2	5	14	14	15	13	7	6	3	3	9	10	18	9	13	2	13
8	10	1	3	10	13	21	5	1	14	2	6	2	1	26	9	8	10	15	2
6	8	3	5	4	14	3	5	1	14	9	13	7	32	7	48	5	6	16	5
4	4	12	9	13	1	7	8	1	7	25	3	40	4	52	13	35	7	11	9
3	12	3	30	3	39	3	24	3	34	3	15	12	52	10	20	3	17	12	3
8	4	39	13	25	3	39	19	24	8	50	3	22	10	20	11	15	14	9	2
10	32	1	29	16	25	2	22	2	45	8	36	10	10	7	15	12	5	8	11
7	1	68	15	29	13	48	8	26	17	36	14	20	12	9	16	17	8	2	8
2	28	13	62	2	99	14	78	8	14	5	14	10	13	11	15	8	12	5	15
13	2	21	9	86	2	81	8	66	1	11	5	10	2	5	15	7	14	8	11
2	17	8	13	1	58	6	61	18	15	16	6	7	4	3	6	2	2	5	7
1	1	14	5	44	8	71	7	41	7	14	4	16	3	11	16	8	1	6	12
1	9	3	8	5	50	1	60	15	23	15	11	16	5	7	12	13	9	10	7
12	7	15	13	34	2	40	6	55	1	10	2	12	3	10	4	5	5	4	23
6	9	1	13	17	105	3	41	7	69	2	36	10	6	4	9	1	13	24	9
7	3	10	2	70	7	27	4	34	2	33	14	1	3	9	18	8	8	2	8
2	10	3	3	7	18	5	8	4	5	4	10	8	7	5	5	5	11	12	13
9	4	7	13	3	6	7	1	11	3	6	4	10	4	15	16	9	13	9	3
6	8	8	9	9	6	4	10	2	6	5	5	13	15	1	7	4	12	9	13

Fr \acute{e} quence de chaque neurone en nb pour 100 unit \acute{e} s de temps-(1e 0 indique les neurones inh \acute{e} b \acute{e} s)

7	4	5	6	4	6	4	6	5	6	5	6	4	6	4	7	3	7	4	7
5	5	4	4	5	3	5	4	5	3	5	3	6	2	6	2	7	2	6	4
6	3	6	3	4	5	4	5	4	6	3	5	3	6	3	6	2	6	3	6
5	5	3	6	4	4	5	3	6	3	5	3	6	3	6	3	7	4	5	5
6	4	5	3	5	5	3	6	3	6	3	5	3	5	5	4	4	5	5	5
5	3	6	4	6	3	6	2	6	2	6	3	5	5	4	6	4	5	3	6
5	6	3	6	3	7	2	7	2	7	2	6	4	4	6	3	5	3	6	5
6	2	7	2	6	2	7	1	8	2	8	2	5	4	3	7	2	6	3	6
4	6	2	7	2	6	2	8	1	8	1	7	3	5	5	2	6	3	4	5
6	2	7	2	7	2	8	0	9	0	9	1	6	4	3	7	2	6	4	6
4	7	2	7	2	8	1	9	0	9	0	8	2	5	5	3	6	3	5	5
6	2	7	1	8	0	10	0	9	0	8	1	7	3	4	5	3	6	3	6
5	5	2	8	0	9	0	9	0	8	1	8	2	5	4	4	5	4	5	5
6	3	7	1	9	0	9	0	9	0	8	1	7	3	5	4	5	4	4	6
4	5	2	8	0	10	0	9	0	9	0	8	2	6	3	5	4	5	4	5
5	3	8	0	10	0	9	0	9	0	9	1	6	3	5	4	5	4	6	5
6	5	2	9	0	9	0	9	1	9	1	7	3	5	4	5	3	6	3	6
6	3	7	0	10	0	9	1	7	2	6	3	5	4	5	3	6	3	5	5
5	6	2	8	0	8	1	6	3	5	4	5	4	5	4	5	3	6	3	6
7	4	7	2	8	3	7	4	6	5	5	5	6	4	6	5	6	5	5	7

Intervalle inter-spike moyen de chaque neurone-(l'étoile indique les neurones inhibés)

```

13 20 16 16 20 15 20 15 18 15 19 16 21 14 22 13 27 13 24 12
20 18 25 22 18 27 17 25 18 26 16 25 16 35 14 36 14 37 15 23
15 29 14 26 21 17 22 17 25 15 25 17 27 15 31 15 37 15 27 15
18 17 30 15 21 24 17 27 16 29 18 25 16 26 15 26 14 25 19 19
16 22 18 25 19 18 27 14 32 15 25 18 25 18 19 20 24 20 19 16
16 30 16 22 16 29 15 35 15 33 15 26 18 19 23 16 23 17 26 15
19 15 29 16 32 13 37 12 41 13 37 15 20 22 15 29 17 30 16 18
14 33 14 35 14 38 13 54 12 48 12 36 17 21 26 13 40 14 26 16
21 14 35 13 38 14 50 11 76 12 67 13 31 17 18 33 14 29 20 18
14 41 13 40 13 41 12 * 10 * 10 67 14 22 26 14 34 16 22 15
21 13 47 12 46 12 82 10 * 10 * 12 33 19 17 29 15 27 19 19
14 35 13 59 12 * 10 * 10 * 11 63 14 25 21 19 26 15 29 15
18 17 38 12 * 10 * 10 * 11 73 12 42 16 22 20 18 24 17 19
14 27 14 90 10 * 10 * 10 * 11 65 13 26 16 24 19 21 22 15
20 16 45 11 * 9 * 10 * 10 * 12 41 15 27 18 22 17 22 17
17 28 12 * 9 * 10 * 10 * 10 57 14 26 17 22 17 24 15 20
16 19 41 10 * 10 * 11 86 10 51 13 27 18 23 16 30 15 31 16
16 26 13 * 9 * 11 63 13 40 15 27 18 21 19 25 15 32 17 18
19 15 41 12 * 12 59 15 31 17 23 18 25 18 23 18 27 15 25 16
13 22 13 33 11 25 13 21 15 19 17 16 16 22 15 20 16 19 17 14

```

Certains neurones sont inactifs, leur voisins ont leur fréquence propre, les autres sont ralentis.

La carte au bout de 2000 itérations est

```

x x x x x x x x x x x x x x x x x x x x x x
x x x x x x x x x x x x x x x x x x x x x x
x x x x x x x x x x x x x x x x x x x x x x
x x x x x x x x x x x x x x x x x x x x x x
x x x x x x x x x x x x x x x x x x x x x x
x x x x x x x x . x x x x x x x x x x x x x
x x x x x x x . x . x x x x x x x x x x x
x x x x x x . x . x . x x x x x x x x x x
x . x x x . x . x . x . x x x x x x x x x
x x . x . x . x . x . x x x x x x x x x x
x x x . x . x . x . x . x x x x x x x x x
x x x . x . x . x . x . x x x x x x x x x
x x x x . x . x . x . x . x x x x x x x x
x x x . x . x . x . x . x x x x x x x x x
x x . x . x . x . x . x . x x x x x x x x
x x . x . x . x x x x x x x x x x x x x x
x x . x . x . x x x x x x x x x x x x x x
x x x x x x x x x x x x x x x x x x x x

```

Les neurones observés sont les mêmes que dans l'exemple 1

Intervalles moyens inter-spikes

Coordonnées	Intervalle	Variance
(10, 10)	58.9	4902.61
(10, 9)	10.1	45.63
(10, 8)	46.9	1916.73
(10, 7)	12.0	76.76
(10, 6)	20.3	354.91
(10, 5)	13.0	309.89
(9, 9)	29.1	2336.91
(9, 8)	11.2	65.01
(8, 8)	51.4	14623.58
(7, 7)	29.1	1789.11

10 premières autocorrélations des neurones

Coordonnées	Autocorrélations									
(10,10)	-.27	-.19	.45	-.36	.10	-.69	-.64	-.71	-.71	-.71
(10, 9)	-.14	.02	-.02	-.02	-.04	.04	-.07	-.08	-.10	-.07
(10, 8)	.10	-.43	-.20	-.46	-.55	-.08	.43	-.22	.07	.76
(10, 7)	.35	.22	.10	.17	.26	.04	-.05	.07	.06	.15
(10, 6)	.10	-.01	.23	-.12	-.22	-.12	-.11	-.40	-.38	-.35
(10, 5)	.18	-.03	.10	.17	-.01	-.01	-.01	.02	-.05	.01
(9, 9)	-.19	.03	-.12	-.17	.00	-.11	-.23	-.22	-.23	-.24
(9, 8)	.13	.06	.01	-.06	.05	.15	.05	.06	-.05	-.01
(8, 8)	.27	.14	-.13	-.11	-.03	-.12	-.13	-.09	-.05	-.05
(7, 7)	.25	-.04	.01	.02	.09	.01	.12	.08	-.11	-.07

Nombre moyen de spikes par fenêtre de 100

Coordonnées	Nombre	Variance
(10,10)	.5	1.20
(10, 9)	9.9	3.88
(10, 8)	.8	1.50
(10, 7)	8.3	8.85
(10, 6)	2.4	8.66
(10, 5)	7.6	14.23
(9, 9)	1.3	6.74
(9, 8)	8.6	6.66
(8, 8)	1.8	4.76
(7, 7)	2.7	8.74

Matrice d'intercorrélations des neurones observés

(10,10)	1.00									
(10, 9)	-.56	1.00								
(10, 8)	.76	-.46	1.00							
(10, 7)	-.53	.20	-.74	1.00						
(10, 6)	.62	-.08	.72	-.75	1.00					
(10, 5)	-.42	-.02	-.59	.65	-.87	1.00				
(9, 9)	.89	-.58	.83	-.63	.55	-.45	1.00			
(9, 8)	-.66	.37	-.84	.68	-.63	.48	-.65	1.00		
(8, 8)	.56	-.38	.80	-.63	.68	-.55	.60	-.81	1.00	
(7, 7)	.58	-.45	.63	-.57	.68	-.52	.55	-.67	.88	1.00

(10,10) (10, 9) (10, 8) (10, 7) (10, 6) (10, 5) (9, 9) (9, 8) (8, 8) (7, 7)

REMARQUES:

- a) Ce cas n'est pas stationnaire, puisque $\theta = 3 > \theta_c = 10/4$. Mais la divergence est lente. On observe des autocorrélations plus grandes (rôle plus important de l'inhibition).
- b) Certains neurones ne déchargent plus, ceux qui sont "protégés" des autres par un entourage de neurones inhibés ont leur fréquence propre (10 décharges pour 100 unités de temps). Une quinconce (carte n 1) commence à se dessiner.
- c) Les coefficients d'intercorrélations tendent
 vers +1 pour les couples de neurones tous deux inhibés (par ex. (10,10), (10,8), (10,6), (9,9), (8,8), (7,7))
 vers -1 pour les couples de neurones inhibé-actif, (par ex. (10,10)-(10,9) ou (10,8)-(10,7)),
 alors qu'ils sont plus faibles entre neurones actifs à l'intérieur de la zone où a commencé la divergence : (10,9)-(10,7) ou (10,9)-(10,5) par exemple.
- d) La variance des estimateurs des intervalles moyens inter-spikes augmente avec θ .

EXEMPLE 1 TER---4 VOISINS---REGIME STATIONNAIRE

Le réseau est de taille 20 * 20

La forme du voisinage à 4 voisins est: x
 x . x
 x

La loi propre des intervalles inter-spikes est uniforme

Le paramètre d'inhibition téta vaut 4

L'intervalle inter-spikes moyen propre est 15

La fréquence moyenne de décharge propre pour 100 unités de temps est: 6

Etat du système au bout de 2000 itérations

4	9	1	10	2	25	5	3	6	14	14	4	24	9	14	4	1	21	21	10
12	13	11	26	5	23	2	14	4	9	4	16	11	1	14	2	6	10	10	14
20	19	16	10	15	6	14	5	6	11	18	14	7	10	2	23	3	17	4	5
5	13	6	6	11	16	2	11	13	5	11	5	20	26	16	6	10	7	11	7
4	4	11	6	18	4	8	11	7	10	17	22	9	11	8	10	7	2	12	6
4	2	11	3	11	2	7	14	16	23	8	7	4	6	17	1	4	5	11	8
12	14	21	6	11	17	19	15	17	21	4	11	6	14	17	8	16	7	19	18
11	4	28	5	15	2	5	17	7	16	5	28	1	17	6	17	15	11	10	15
22	45	10	26	5	14	1	12	13	5	9	18	8	12	26	5	9	15	7	13
16	1	15	3	23	10	15	15	21	10	9	8	10	11	13	15	6	3	14	13
6	13	27	23	4	16	8	10	16	11	24	10	3	5	11	11	2	13	10	19
9	13	13	7	18	21	28	11	5	7	15	7	13	8	13	12	1	8	8	8
13	21	16	13	7	7	9	22	8	9	21	18	18	6	11	13	17	7	3	11
12	20	18	17	9	4	1	10	21	15	12	20	15	13	18	6	13	14	9	23
2	6	17	4	26	14	11	17	11	11	11	7	6	2	4	10	2	17	7	16
17	21	14	15	19	17	1	11	5	5	6	18	17	1	13	22	18	3	9	14
15	5	23	13	18	19	27	23	13	15	23	8	11	6	7	16	25	21	7	9
13	11	5	8	9	9	9	15	26	8	5	7	16	23	8	24	15	6	6	23
4	3	2	1	2	1	7	2	6	12	13	17	18	6	13	11	8	14	1	8
26	6	10	4	5	19	22	3	4	23	8	4	5	11	22	6	12	16	7	15

Fréquence de chaque neurone en nb pour 100 unités de temps--(le 0 indique les neurones inhibés)

4	3	3	3	3	4	3	3	3	4	3	3	4	3	4	3	3	3	4	4
4	3	3	2	3	2	3	2	3	3	3	3	3	2	3	3	3	3	3	3
4	2	3	4	2	4	2	4	2	3	3	3	2	4	2	3	2	3	3	3
3	3	2	3	3	2	4	2	4	3	3	2	4	3	3	3	3	3	4	4
3	2	3	3	2	4	2	4	2	3	3	3	3	3	2	3	3	3	3	3
3	3	3	3	3	2	4	2	4	2	4	3	3	3	3	3	3	3	2	4
4	2	3	2	3	4	2	4	2	3	3	2	3	3	3	3	2	4	3	3
3	3	3	3	3	2	4	2	4	2	3	3	3	3	3	3	3	3	3	4
4	2	4	2	3	3	2	3	2	3	2	3	3	3	3	3	3	3	3	3
3	4	2	4	2	3	4	3	3	3	3	3	3	3	3	3	3	3	3	4
4	1	5	2	4	2	3	2	3	3	3	3	3	3	2	3	3	2	3	3
3	4	2	3	2	4	2	3	3	3	2	4	2	3	3	2	3	3	3	3
4	2	3	3	4	1	4	3	3	3	3	3	3	2	3	3	2	3	3	3
4	3	3	4	2	4	2	3	3	3	3	2	4	2	5	2	3	3	3	4
4	3	2	3	3	3	3	2	3	3	3	3	2	4	1	4	2	3	3	4
3	3	3	3	3	2	3	3	3	2	3	4	2	4	3	3	3	3	3	3
4	3	3	3	3	4	2	3	2	3	3	2	4	2	3	2	3	2	3	4
3	4	2	3	3	3	4	2	3	3	3	2	4	2	4	3	4	2	3	4
3	2	4	2	3	3	2	4	2	3	2	4	2	4	2	3	3	3	3	3
4	3	3	4	4	3	5	3	4	3	4	2	5	3	4	3	4	4	3	4

Intervalle inter-spike moyen de chaque neurone-(l'étoile indique les neurones inhibés)

```

22 25 26 26 25 26 24 26 25 30 24 25 27 23 25 22 25 25 27 21
25 27 30 34 29 34 27 33 27 28 29 30 29 29 34 29 33 25 27 27
24 35 26 24 38 20 39 22 37 27 27 30 26 37 22 38 30 37 28 27
25 25 40 26 26 37 24 50 24 31 30 30 34 25 32 26 30 29 31 24
26 34 27 32 33 25 47 21 38 24 33 27 28 28 32 28 37 27 25 26
26 27 32 26 28 38 21 41 24 40 24 27 31 30 27 33 27 26 38 24
22 34 26 36 31 22 47 21 48 21 38 27 26 33 29 31 26 35 25 31
33 25 32 25 32 35 22 45 22 46 21 35 28 27 29 30 31 27 31 23
23 46 21 35 25 28 35 26 36 26 38 29 29 33 26 31 29 31 29 25
28 22 47 23 38 27 24 29 25 31 27 25 32 28 30 30 31 28 29 24
21 54 19 47 23 35 27 37 27 31 31 26 30 31 26 35 25 31 35 28
28 22 41 25 41 20 40 26 32 28 27 38 24 40 26 27 38 27 31 26
23 33 27 33 20 61 21 33 27 25 33 26 32 25 38 30 26 35 26 28
23 32 27 24 35 23 38 27 32 33 29 34 24 36 19 38 27 31 30 24
25 27 35 28 27 30 29 35 29 27 30 26 38 24 51 22 34 27 30 25
26 30 26 30 26 36 29 30 26 31 35 29 24 44 21 31 25 31 30 26
25 29 26 30 31 22 36 25 33 28 25 26 43 21 38 27 34 28 33 25
26 24 35 26 32 32 25 38 25 32 26 37 21 43 24 32 25 34 30 22
26 35 23 40 28 30 37 21 38 25 46 22 50 20 38 28 33 30 25 33
20 26 27 24 23 32 18 30 23 30 22 34 20 32 22 27 24 25 29 20

```

Tous les neurones sont actifs

Les neurones observés sont les mêmes que dans l'exemple 1

Intervalles moyens inter-spikes

Coordonnées	Intervalle	Variance
(10, 10)	29.3	1086.45
(10, 9)	25.4	1047.02
(10, 8)	29.0	1107.91
(10, 7)	24.6	810.55
(10, 6)	27.1	2766.75
(10, 5)	35.9	2499.56
(9, 9)	34.8	1244.06
(9, 8)	26.8	403.67
(8, 8)	45.4	1811.26
(7, 7)	48.2	5219.69

10 premières autocorrélations des neurones

Coordonnées	Autocorrélations									
(10, 10)	-.05	.04	-.09	.01	.11	-.06	.04	-.17	.16	.05
(10, 9)	-.07	.02	.02	.04	.04	-.07	-.01	.02	-.07	-.06
(10, 8)	.35	.21	.06	.01	-.06	.13	-.13	.02	-.02	-.20
(10, 7)	.03	.21	.22	-.14	-.01	-.15	-.18	-.14	-.21	-.19
(10, 6)	-.09	-.04	-.05	.01	-.06	.01	.04	-.06	-.05	.08
(10, 5)	.21	-.11	-.16	-.12	-.14	-.15	-.13	-.04	-.04	-.24
(9, 9)	-.01	-.12	-.25	-.11	-.17	.00	-.01	.13	-.05	.03
(9, 8)	.07	.08	.05	.13	-.03	-.03	.02	-.22	.03	-.08
(8, 8)	.11	.07	.11	.00	-.11	-.02	.11	-.07	.07	.07
(7, 7)	-.02	-.18	.13	.01	-.04	.06	-.16	-.02	.08	.01

Nombre moyen de spikes par fenêtre de 100

Coordonnées	Nombre	Variance
(10,10)	3.2	5.32
(10, 9)	3.9	5.94
(10, 8)	3.4	4.35
(10, 7)	4.1	8.89
(10, 6)	3.7	6.11
(10, 5)	2.6	6.44
(9, 9)	2.7	2.40
(9, 8)	3.7	2.30
(8, 8)	2.2	2.38
(7, 7)	2.1	2.52

Matrice d'intercorrélations des neurones observés

(10,10)	1.00																				
(10, 9)	-.72	1.00																			
(10, 8)	.44	-.64	1.00																		
(10, 7)	-.15	-.25	-.65	1.00																	
(10, 6)	-.17	-.09	.55	-.80	1.00																
(10, 5)	.12	-.01	-.38	.62	-.75	1.00															
(9, 9)	.35	-.45	.34	-.22	.18	-.23	1.00														
(9, 8)	-.21	.30	-.50	.68	-.56	.44	-.54	1.00													
(8, 8)	-.34	.47	-.06	-.35	.17	-.22	-.13	-.38	1.00												
(7, 7)	.02	.02	.40	-.47	.25	-.29	-.01	-.36	.59	1.00											

(10,10) (10, 9) (10, 8) (10, 7) (10, 6) (10, 5) (9, 9) (9, 8) (8, 8) (7, 7)

REMARQUES:

- Ce cas ($\theta = 4$, $\theta_0 = 15/4$) apparaît encore comme "stationnaire" à la simulation, car la divergence est très lente.
- Les intervalles inter-spikes moyens sont de l'ordre de $E(F) + 4\theta = 29$, même si commence à se dessiner une inhibition en alternance. La variance est très grande, comme cela se produit quand θ augmente.
- Les autocorrélations sont faibles, comme en régime stationnaire.
- La matrice d'intercorrélations a les mêmes caractéristiques que celle de l'exemple 1-bis, même si la différenciation entre neurones actifs et inhibés ne s'est pas encore faite clairement.

Intervalle inter-spike moyen de chaque neurone-(l'étoile indique les neurones inhibés)

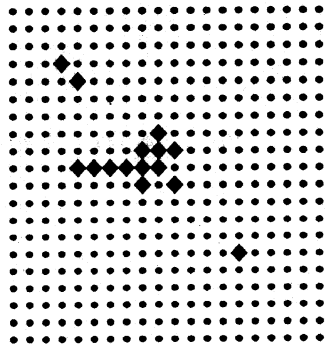
```

19 21 23 22 24 23 21 22 24 25 23 22 23 25 26 24 24 22 23 19
22 23 23 28 24 25 23 22 31 23 23 27 26 24 22 22 24 28 24 19
20 25 24 23 27 22 23 26 21 26 26 25 23 26 27 26 24 25 24 21
25 22 22 27 23 29 26 22 30 24 22 27 22 24 27 24 28 26 24 20
21 24 24 22 24 23 24 27 23 23 25 27 24 27 25 23 24 26 26 19
24 21 26 29 22 24 26 25 24 26 26 27 24 27 23 25 26 28 22 24
20 21 23 25 24 23 26 27 23 25 26 24 22 24 26 26 24 29 23 19
22 25 24 24 23 26 24 25 26 23 25 27 27 26 29 25 28 21 24 22
20 20 24 29 27 23 26 24 21 27 29 25 23 27 26 25 23 26 22 22
24 23 24 24 24 28 29 25 23 27 24 26 23 25 25 21 27 24 23 22
19 22 28 25 25 28 24 25 26 23 26 24 27 25 26 30 23 26 25 21
23 21 26 29 27 21 31 24 23 27 26 23 24 24 26 21 24 28 21 21
22 26 25 23 26 28 21 25 24 25 25 23 21 28 22 26 26 23 22 24
21 21 25 25 26 26 25 23 24 27 27 24 23 26 20 26 23 26 25 20
21 26 25 25 25 26 28 25 25 23 22 25 25 26 26 23 28 23 23
20 23 25 28 24 26 26 24 22 24 25 28 26 23 26 23 24 27 21 21
24 20 24 26 25 23 24 25 26 23 25 22 27 24 26 27 27 27 23 20
19 22 28 23 27 28 24 25 25 25 23 26 24 26 24 25 27 26 24 20
23 23 23 26 27 24 23 26 28 24 24 25 25 27 26 23 25 26 24 19
21 22 24 25 20 24 28 20 23 24 24 22 21 21 22 23 22 24 24 18

```

Tous les neurones sont actifs, leur intervalle moyen inter-spike est de l'ordre de $E(F)+6\theta$

Les neurones observés sont marqués



Intervalles moyens inter-spikes

Coordonnées	Intervalle	Variance
(10,10)	27.1	297.22
(10, 9)	23.3	268.44
(10, 8)	25.2	272.06
(10, 7)	29.4	239.67
(10, 6)	28.5	299.19
(10, 5)	23.7	217.70
(9,10)	27.3	321.80
(8,10)	23.8	286.56
(9, 9)	21.9	243.55
(4, 4)	27.1	297.76
(15,15)	26.7	269.49
(11,11)	26.1	291.78
(9,11)	29.0	313.39
(11, 9)	26.7	319.27
(5, 5)	23.7	324.04

10 premières autocorrélations des neurones

Coordonnées	Autocorrélations									
(10,10)	-.10	.00	-.03	.04	-.19	-.22	-.27	-.10	-.32	-.23
(10, 9)	-.12	-.01	-.03	-.18	.08	-.16	-.08	-.06	.07	.20
(10, 8)	-.05	-.06	.10	.19	.09	.02	.12	.16	.26	.10
(10, 7)	-.21	-.13	.16	-.27	.05	-.06	-.35	.11	-.36	-.38
(10, 6)	-.14	.03	-.09	.02	.28	-.25	-.19	-.01	-.06	.13
(10, 5)	-.07	.11	-.18	-.15	-.10	-.14	-.22	.14	.02	.22
(9,10)	.21	-.04	-.02	.00	.02	.24	.20	.05	-.11	-.18
(8,10)	-.07	-.12	.00	-.07	.04	.00	-.01	-.08	.06	-.06
(9, 9)	.06	-.08	-.14	-.12	.14	.02	.21	.01	-.11	-.03
(4, 4)	.08	-.08	.03	.21	.14	.04	-.08	.02	.02	.01
(15,15)	.11	-.05	-.05	-.29	-.10	.03	.07	.27	.06	-.18
(11,11)	.13	-.14	-.12	.01	.11	-.10	.28	.15	.15	-.10
(9,11)	-.14	.07	.00	.01	-.12	-.10	-.05	.10	-.28	.07
(11, 9)	.04	-.20	-.10	-.12	-.11	-.06	-.16	-.10	-.23	-.05
(5, 5)	-.05	-.20	.03	-.18	.01	.00	-.01	-.13	.05	.00

Nombre moyen de spikes par fenêtre de 100

Coordonnées	Nombre	Variance
(10,10)	3.6	1.51
(10, 9)	4.3	2.19
(10, 8)	4.0	1.37
(10, 7)	3.4	.77
(10, 6)	3.5	1.09
(10, 5)	4.2	2.03
(9,10)	3.6	1.71
(8,10)	4.2	1.85
(9, 9)	4.6	2.45
(4, 4)	3.7	1.90
(15,15)	3.7	1.04
(11,11)	3.8	1.43
(9,11)	3.4	1.51
(11, 9)	3.8	1.01
(5, 5)	4.2	2.03

Matrice d'intercorrélations des neurones observés

(10,10)	1.00																				
(10, 9)	-.26	1.00																			
(10, 8)	-.48	.06	1.00																		
(10, 7)	-.21	.24	-.10	1.00																	
(10, 6)	.38	.08	.00	.10	1.00																
(10, 5)	-.23	-.07	-.03	-.56	-.45	1.00															
(9,10)	-.10	.00	-.21	-.14	.26	.26	1.00														
(8,10)	-.48	.02	.23	.15	-.23	-.23	-.58	1.00													
(9, 9)	-.37	.00	.20	.13	-.28	-.04	-.08	.17	1.00												
(4, 4)	.20	.01	-.26	.50	.15	-.30	.08	.04	-.16	1.00											
(15,15)	.00	-.02	.13	.24	.03	-.44	-.35	.08	.39	.32	1.00										
(11,11)	.37	.06	-.08	.13	.13	-.20	.19	-.13	-.22	.25	-.26	1.00									
(9,11)	-.20	.32	.11	-.07	.11	-.01	-.30	.26	-.07	-.17	-.25	-.05	1.00								
(11, 9)	.18	-.18	-.49	.16	-.25	-.09	-.22	.19	-.29	.10	-.16	.01	.07	1.00							
(5, 5)	.19	.03	-.28	.24	-.16	.01	.26	-.04	.34	.11	-.23	.30	-.13	-.12	1.00						

(10,10) (10,9) (10,8) (10,7) (10,6) (10,5) (9,10) (8,10) (9,9) (4,4) (15,15) (11,11) (9,11) (11,9) (5,5)

Il s'agit d'un cas stationnaire ($\theta=2 < \theta_c=15/6$): intervalle moyen inter-spikes de l'ordre de $15+6\theta$, autocorrélations faibles, coefficients d'intercorrélations peu significatifs, les plus importants correspondant aux couples de neurones voisins qui s'inhibent: (10,10)-(10,8) ou (8,10)-(9,10) par exemple.

EXEMPLE 3---8 VOISINS---RESINE STATIONNAIRE

Le réseau est de taille 40 * 40
La forme du voisinage à 8 voisins est:
La loi propre des intervalles inter-splkes est uniforme
Le paramètre d'inhibition tête vaut 2
L'Intervalle Inter-splkes moyen propre est 20
La fréquence moyenne de décharge propre pour 100 unités de temps est: 5

Etat du système au bout de 3000 itérations

Table with 40 columns and multiple rows of numerical data, representing the state of the system over time and across different neurons.

Fréquence de chaque neurone en nb pour 100 unités de temps -(le 0 indique les neurones inhi(bés))

Intervalles moyens inter-spikes

Coordonnées	Intervalle	Variance
(20,20)	33.5	329.94
(19,20)	37.0	446.08
(20,19)	32.2	479.92
(20,21)	35.5	462.43
(21,20)	28.9	403.24
(19,19)	34.7	436.29
(10,10)	34.9	421.76
(20,10)	38.4	484.52
(5,5)	33.4	438.26
(15,15)	34.5	395.34
(1,1)	24.4	216.55
(35,35)	32.0	415.21
(30,10)	32.5	482.98
(5,35)	31.4	467.37
(35,20)	33.6	470.84

10 premières autocorrélations des neurones

Coordonnées	Autocorrélations									
(20,20)	-.02	-.05	-.13	-.22	.23	.16	-.09	-.03	-.12	.06
(19,20)	-.01	-.18	-.03	-.04	-.25	.01	-.18	-.12	-.06	-.32
(20,19)	-.09	.09	.08	-.07	-.05	-.04	-.08	-.01	.20	-.17
(20,21)	-.06	.08	-.06	-.15	-.09	.07	-.03	.03	.06	-.07
(21,20)	.04	-.12	-.14	-.21	-.22	.18	.10	-.09	.00	-.17
(19,19)	-.13	-.05	-.17	-.16	-.09	.03	-.06	-.19	-.07	-.34
(10,10)	-.06	.05	.05	.27	-.06	.02	.02	.20	-.14	.01
(20,10)	-.12	.09	-.20	-.07	.01	.13	.06	-.08	-.03	-.01
(5,5)	-.19	-.04	.12	.03	-.04	.15	.12	-.22	-.09	.08
(15,15)	-.06	.04	-.05	-.12	.07	.14	.19	.42	.21	-.02
(1,1)	-.03	-.10	-.03	-.13	.08	.03	-.13	-.10	.09	.03
(35,35)	.02	.02	.09	-.10	.02	-.11	-.06	-.15	-.26	-.11
(30,10)	.07	-.03	.01	-.17	-.06	-.06	-.03	-.14	-.05	-.28
(5,35)	-.10	.24	.10	.01	-.07	.31	.01	.07	.32	-.12
(35,20)	-.16	.03	-.10	.04	.02	.01	.14	-.20	.03	-.09

Nombre moyen de spikes par fenêtre de 100

Coordonnées	Nombre	Variance
(20,20)	3.0	.50
(19,20)	2.7	.92
(20,19)	3.1	.99
(20,21)	2.8	1.17
(21,20)	3.5	1.96
(19,19)	2.9	1.12
(10,10)	2.8	1.17
(20,10)	2.6	1.02
(5,5)	3.0	.86
(15,15)	2.9	1.27
(1,1)	4.1	1.10
(35,35)	3.1	1.77
(30,10)	3.1	1.64
(5,35)	3.1	1.41
(35,20)	3.0	1.89

Matrice d'intercorrélations des neurones observés

(20,20)	1.00																				
(19,20)	-.21	1.00																			
(20,19)	.20	.10	1.00																		
(20,21)	.00	-.13	.28	1.00																	
(21,20)	.04	-.41	-.15	-.19	1.00																
(19,19)	-.43	.10	-.02	.10	.04	1.00															
(10,10)	.33	.25	.31	.05	.07	-.18	1.00														
(20,10)	.10	.07	-.22	.15	.03	.05	-.01	1.00													
(5,5)	.33	.20	.19	.11	-.11	-.29	.00	.11	1.00												
(15,15)	.04	-.27	-.12	.03	.02	-.32	.00	-.05	-.03	1.00											
(1,1)	-.24	.07	.06	-.20	-.16	.14	.05	-.13	-.15	.01	1.00										
(35,35)	-.19	-.19	.26	.25	-.36	-.04	-.30	-.14	-.09	-.08	-.09	1.00									
(30,10)	.24	.10	.39	.37	.00	.01	.19	-.12	.45	-.29	-.25	-.07	1.00								
(5,35)	-.04	-.22	-.07	.14	.20	-.24	.14	-.13	-.03	.28	-.13	.15	-.03	1.00							
(35,20)	.00	-.06	.03	-.01	.16	.39	-.20	-.16	.00	.04	.30	-.15	-.10	.03	1.00						

(20,20)(19,20)(20,19)(20,21)(21,20)(19,19)(10,10)(20,10)(5,5)(15,15)(1,1)(35,35)(30,10)(5,35)(35,20)

REMARQUES:

Dans cet exemple à 8 voisins $\theta=2 < \theta_0=20/8$, l'inhibition est faible. Il y a peu de corrélations entre les sites, du même ordre de grandeur entre neurones proches ou neurones éloignés.

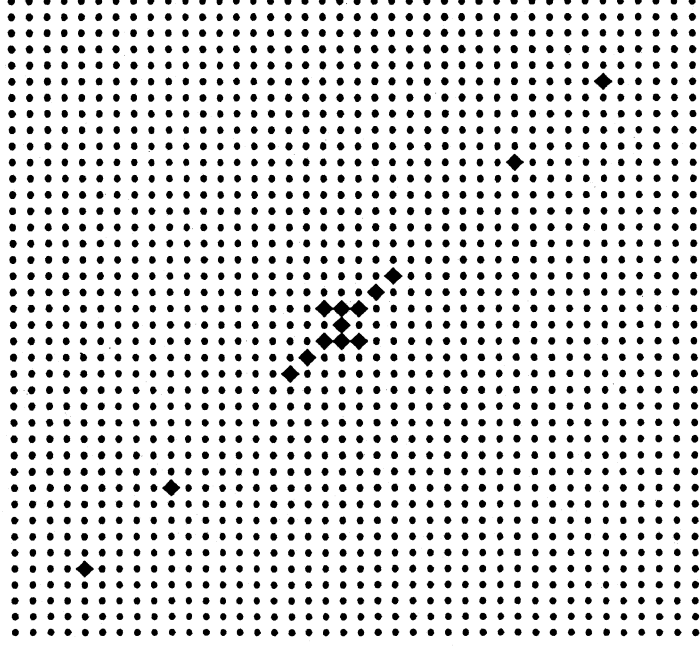
L'autocorrélation est peu significative et l'intervalle moyen inter-spike est de l'ordre de $E(F)+8\theta=36$.

Intervalle Inter-spike moyen de chaque neurone-(1'étoile) indique les neurones inhibés)

Table with 32 rows and 32 columns of numerical data representing mean inter-spike intervals for each neuron. The values range from approximately 27 to 55. The table is a lower triangular matrix where the diagonal elements represent the self-connections and the lower triangle represents the connections between different neurons. The top row starts with 31 and ends with 32. The bottom row starts with 32 and ends with 31.

Tous les neurones sont actifs, leur intervalle inter-spike est de 1'ordre de E(F)+12 θ

Les neurones observés sont marqués



Intervalles moyens inter-spikes

Coordonnées	Intervalle	Variance
(20, 20)	46.6	785.59
(20, 19)	46.0	844.45
(20, 21)	42.8	816.63
(21, 19)	44.5	863.46
(19, 21)	42.6	871.62
(19, 19)	44.2	849.65
(18, 18)	48.5	782.68
(17, 17)	45.0	588.11
(10, 10)	48.2	664.58
(5, 5)	39.0	790.09
(21, 21)	39.9	829.26
(22, 22)	51.3	1047.47
(23, 23)	43.9	886.43
(30, 30)	44.1	674.91
(35, 35)	50.5	932.65

10 premières autocorrélations des neurones

Coordonnées	Autocorrélations											
(20, 20)	.20	.14	.06	-.16	-.11	-.20	-.30	-.04	.25	.22		
(20, 19)	-.08	-.21	.46	-.06	-.22	.18	-.01	-.01	.11	-.05		
(20, 21)	-.18	-.05	.01	-.04	-.15	-.06	.11	-.14	.07	-.20		
(21, 19)	-.19	-.16	.19	.12	-.38	-.03	-.10	-.29	-.08	-.14		
(19, 21)	.13	-.05	-.10	-.01	.08	-.24	-.08	-.10	-.17	-.07		
(19, 19)	-.14	-.05	-.05	.18	.13	-.09	.08	-.01	.04	-.22		
(18, 18)	.03	.14	.05	.09	.17	.37	.03	.23	-.01	-.11		
(17, 17)	-.15	-.15	-.04	.12	-.24	-.03	-.07	-.58	-.25	-.39		
(10, 10)	-.14	.13	.06	.00	-.01	-.11	-.20	-.13	-.15	-.07		
(5, 5)	-.15	.08	-.19	-.23	-.20	-.08	.17	.04	.20	-.03		
(21, 21)	-.04	-.05	-.04	-.06	.16	-.18	.06	-.04	-.11	.17		
(22, 22)	.18	-.06	-.26	-.20	-.07	.10	.30	.16	-.08	-.15		
(23, 23)	.00	-.02	.23	.08	-.06	.06	.18	.22	-.07	-.07		
(30, 30)	-.29	-.09	-.09	.12	.07	.00	-.14	-.02	-.20	-.02		
(35, 35)	.19	.11	-.07	-.15	-.16	-.20	-.40	-.47	-.21	-.20		

Nombre moyen de spikes par fenêtre de 100

Coordonnées	Nombre	Variance
(20, 20)	2.1	1.31
(20, 19)	2.2	.58
(20, 21)	2.3	.58
(21, 19)	2.2	1.03
(19, 21)	2.3	1.50
(19, 19)	2.3	.92
(18, 18)	2.0	.68
(17, 17)	2.2	.74
(10, 10)	2.1	.64
(5, 5)	2.5	1.54
(21, 21)	2.5	1.61
(22, 22)	1.9	1.14
(23, 23)	2.3	1.35
(30, 30)	2.3	.49
(35, 35)	2.0	1.57

Matrice d'intercorrélations des neurones observés

(20, 20)	1.00																								
(20, 19)	-.31	1.00																							
(20, 21)	-.08	.02	1.00																						
(21, 19)	.07	-.05	.10	1.00																					
(19, 21)	-.25	.21	.04	.29	1.00																				
(19, 19)	-.06	-.32	.22	-.17	-.38	1.00																			
(18, 18)	-.08	-.12	-.02	.16	-.44	.49	1.00																		
(17, 17)	-.46	-.16	.06	.15	.41	-.07	-.16	1.00																	
(10, 10)	-.05	-.14	.14	-.02	.34	-.31	-.28	-.18	1.00																
(5, 5)	-.21	-.06	.06	-.11	.15	.12	-.19	.34	-.14	1.00															
(21, 21)	-.11	.02	-.20	-.28	-.06	.21	.29	.03	-.18	.41	1.00														
(22, 22)	-.26	.06	.07	.08	-.12	.12	.17	.33	.05	.11	-.13	1.00													
(23, 23)	.17	-.30	-.31	-.26	-.01	.09	-.27	-.28	.06	-.05	.10	-.16	1.00												
(30, 30)	.05	-.10	-.17	.27	-.02	-.12	.11	-.22	-.04	-.29	-.24	-.16	.08	1.00											
(35, 35)	-.15	.19	-.04	.00	.46	-.18	.00	-.07	.54	-.09	.31	-.29	-.07	.16	1.00										

(20, 20) (20, 19) (20, 21) (21, 19) (19, 21) (19, 19) (18, 18) (17, 17) (10, 10) (5, 5) (21, 21) (22, 22) (23, 23) (30, 30) (35, 35)

REMARQUES:

Les remarques à faire sont les mêmes que celles de l'exemple 3. On voit que l'intercorrélacion (faible) est répartie dans tout le réseau, elle ne diminue pas en fonction de la distance des sites.

EXEMPLE 5--14 VOISINS--REGIME STATIONNAIRE

Le réseau est de taille 40 * 40
La forme du voisinage à 14 voisins est:
XXXXXXXX
XXXXXXXX

La loi propre des intervalles inter-splkes est uniforme
Le paramètre d'inhibition téta vaut 2
L'intervalle inter-splkes moyen propre est 30

La fréquence moyenne de décharge propre pour 100 unités de temps est: 3
Etat du système au bout de 3000 itérations

Table with 40 columns and 3000 rows of numerical data representing system state over time.

Fréquence de chaque neurone en nb pour 100 unités de temps-(le 0 indique les neurones finités)

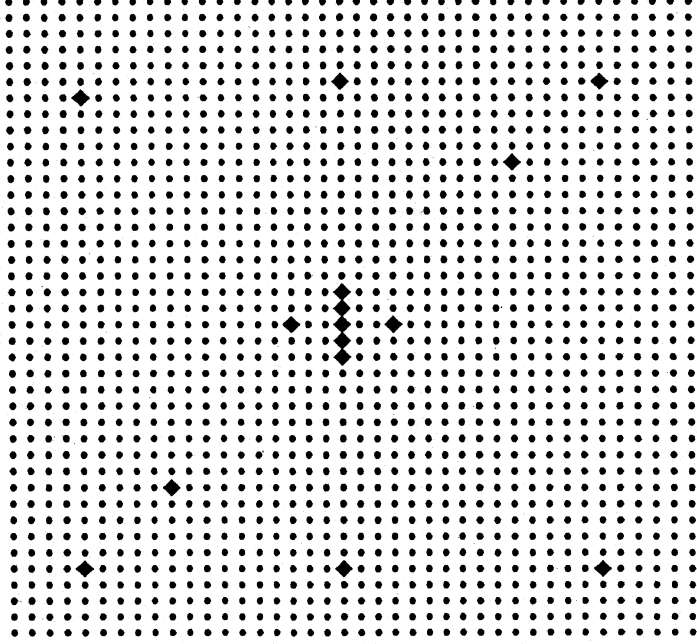
Table with 40 columns and 3000 rows of numerical data representing neuron frequency over time.

Intervalle Inter-spike moyen de chaque neurone-(*)étoile indique les neurones inhibés)

34	48	48	43	45	52	40	47	41	51	54	43	43	45	50	46	38	43	43	46	41	43	49	45	48	40	48	42	43	43	45	46	43	36								
42	50	57	62	56	59	49	54	53	56	59	49	54	60	56	52	54	60	53	54	52	58	58	49	62	50	42	51	58	52	54	51	57	58	53	58	59	50	41			
45	46	56	58	60	54	51	53	53	54	56	52	58	54	50	58	52	52	59	52	46	54	50	58	52	50	48	54	56	47	56	53	53	56	48	50	48					
36	52	62	62	56	58	54	54	54	52	56	50	49	57	57	57	58	52	50	45	49	49	52	62	57	58	54	52	54	51	49	54	56	52	56	46	38					
38	52	62	62	56	58	54	49	53	49	56	55	52	57	57	57	58	52	50	51	62	48	50	48	49	50	48	50	53	51	51	58	58	45	40							
43	50	51	57	46	51	52	52	56	53	58	45	48	58	56	56	56	56	53	50	54	46	58	60	50	53	52	58	56	52	57	56	48	45								
43	50	51	57	46	51	52	52	56	53	58	45	48	58	56	56	56	56	53	50	54	46	58	60	50	53	52	58	56	52	57	56	48	45								
40	47	58	54	52	58	57	51	56	49	50	48	51	51	54	51	53	54	57	51	54	53	50	60	52	54	48	52	52	52	56	62	45	47								
45	45	65	57	59	54	49	52	58	47	60	59	54	56	57	54	54	52	52	58	49	54	53	59	48	52	50	65	52	51	65	48	51	53	40							
43	50	66	48	52	47	57	50	54	51	59	52	54	53	47	54	58	49	51	49	54	58	46	66	66	52	52	57	45	51	50	54	53	50	39							
48	48	57	57	45	57	46	57	54	47	46	52	51	54	45	49	54	58	46	50	48	49	51	58	53	53	52	59	57	52	50	43	59	59	45	41						
35	52	57	54	57	45	57	46	52	51	54	45	49	60	52	51	53	66	50	49	58	50	48	49	51	58	53	53	52	59	57	52	59	59	45	41						
42	50	66	48	45	56	47	50	62	52	57	53	49	54	52	63	49	54	60	53	51	43	62	59	45	50	52	54	46	57	62	50	49	59	50	43						
43	53	54	60	50	52	58	51	58	50	51	48	59	58	47	51	48	59	58	48	52	56	53	58	50	59	59	54	46	48	51	49	50	57	48	43						
43	44	53	60	59	50	52	58	51	58	50	51	48	59	58	47	51	48	59	58	48	52	56	53	58	50	59	59	54	46	48	51	49	50	57	48	43					
45	48	57	60	59	47	59	52	50	57	52	50	51	54	57	46	57	48	54	77	48	54	77	48	54	77	48	54	77	48	54	77	48	54	77	48	45					
45	48	57	60	59	46	66	51	50	59	50	53	50	56	53	59	63	54	54	58	51	51	59	45	52	51	51	51	51	51	59	52	49	52	45	43						
42	48	60	52	47	51	59	47	50	50	57	52	48	36	53	50	58	53	51	59	52	52	46	52	58	59	36	56	58	51	49	50	57	60	54	60	45					
41	40	58	54	53	56	53	50	52	46	50	47	58	52	62	48	36	54	48	56	54	50	58	53	51	54	52	49	47	51	58	53	50	50	52	50	44					
36	51	63	49	50	56	50	44	49	70	56	50	54	60	51	56	50	54	50	54	50	56	58	45	48	59	51	46	58	49	49	51	60	53	48	48	51	66	90	43	47	50
45	50	62	57	54	59	51	53	62	52	50	51	57	50	54	50	58	58	50	56	58	52	63	52	51	48	59	59	52	50	45	58	60	50	43	47	50					
38	48	59	60	53	50	45	43	54	45	51	52	53	48	58	51	53	54	48	47	51	52	51	56	52	50	45	51	59	53	57	54	50	47	43							
44	46	66	52	57	50	60	59	62	54	52	59	51	58	58	51	50	53	44	48	47	51	52	51	56	58	60	50	52	50	52	50	62	62	48	42						
39	45	54	54	53	47	50	60	56	50	51	56	57	49	56	59	50	57	70	57	57	56	58	60	52	56	63	58	59	60	50	52	50	62	62	48	42					
44	52	58	53	51	56	57	52	50	54	45	48	50	58	53	50	58	54	52	56	54	52	50	60	59	56	44	51	47	50	52	57	45	44	54	50	58	51	38			
43	47	56	60	57	51	60	51	58	70	58	48	58	60	46	49	58	54	50	52	65	49	56	48	56	45	57	58	54	56	51	50	60	45	50	53	50	44	48			
46	50	53	59	42	51	60	50	59	49	50	45	49	51	52	56	50	57	58	51	57	52	65	54	59	51	59	48	51	51	50	60	45	50	53	50	44	48				
44	50	53	59	48	56	56	47	54	47	54	59	47	48	53	52	59	52	53	56	52	51	56	63	56	57	51	50	52	54	59	56	57	48	54	49	59	51	38			
44	52	63	51	63	53	45	57	52	50	50	45	47	54	50	51	52	57	51	45	47	50	63	48	47	48	54	52	50	58	53	50	60	54	56	41						
41	48	58	51	63	53	45	57	52	50	50	45	47	54	50	51	52	57	51	45	47	50	63	48	47	48	54	52	50	58	53	50	60	54	56	41						
43	50	50	52	54	52	43	60	53	54	52	45	47	54	50	51	52	57	51	45	47	50	63	48	47	48	54	52	50	58	53	50	60	54	56	41						
46	41	51	62	50	51	52	59	57	53	47	54	45	47	54	50	51	52	57	51	45	47	50	63	48	47	48	54	52	50	58	53	50	60	54	56	41					
46	41	51	62	50	51	52	59	57	53	47	54	45	47	54	50	51	52	57	51	45	47	50	63	48	47	48	54	52	50	58	53	50	60	54	56	41					
38	47	68	58	59	49	50	58	45	52	53	52	62	56	52	51	57	50	57	45	47	54	50	58	53	50	59	57	57	48	51	58	52	48	41							
41	47	50	60	50	51	53	52	57	56	52	46	62	70	45	52	50	54	51	58	58	60	57	53	57	53	54	69	49	60	53	56	56	48	44							
48	49	68	57	51	50	65	59	51	60	52	50	45	56	60	65	52	59	57	54	53	59	59	60	53	52	51	62	46	57	50	55	52	47	52	60	48	41				
38	44	44	44	45	45	42	43	45	45	41	54	43	41	44	40	48	41	43	44	42	41	38	49	47	50	41	41	40	46	45	42	43	45	46	45	42	43				

Tous les neurones sont actifs, leur Intervalle moyen Inter-spike est de l'ordre de E(f)+i4θ

Les neurones observés sont marqués



ANALYSE MATHÉMATIQUE D'UN MODÈLE DU CORTEX CÉRIBELLEUX :

II. EFFET DE L'ACTION INHIBITRICE DES COLLATÉRALES RECURRENTES EN PRÉSENCE DE STIMULI.

*Marie COTTRELL
U.A. C.N.R.S. 743 "Statistique Appliquée"
Laboratoire de Mathématique, Bât. 425
Université de Paris 11
F - 91405 ORSAY Cedex*

Nous analysons les réactions d'un réseau de neurones reliés par des liaisons inhibitrices, en présence de stimuli extérieurs.

L'état du système est dans ce cas un processus de Markov dont le comportement dépend des paramètres du système (niveau de l'inhibition, degré d'excitation).

Ce modèle du cortex cérébelleux intègre les excitations apportées par les fibres grimpantes ou les fibres moussues.

MOTS CLES :

**Processus de Markov
Réseaux de neurones
Cortex cérébelleux
Liaisons inhibitrices
Stimuli**

I - INTRODUCTION.

Dans cette deuxième partie, on prend en compte des stimuli extérieurs et on considère que certains neurones sont excités. Voir le système C, figure 2c du I.2, Partie I.

On modélise ainsi le fait que chez le jeune rat les fibres grimpan-
tes (présentes dès le 3ème jour) excitent directement quelques cellules
de Purkinjie (3 à 5) situées principalement sur un axe transversal
(CREPEL et al. (1970)). Au contraire, chez le rat adulte, chaque fibre
grimpante ne contacte qu'une cellule de Purkinjie.

On modélise aussi le fait que les signaux apportés par les fibres
moussues sont transmis par l'intermédiaire des fibres parallèles qui
contactent un grand nombre de cellules de Purkinjie, situées principa-
lement le long de l'axe longitudinal.

II - MODELE MATHEMATIQUE DU RESEAU DE NEURONES AVEC EXCITATION

1. LE MODELE - INTRODUCTION DES STIMULI.

Les notations générales sont les mêmes que dans la première partie (Analyse Mathématique du Cortex Cerebelleux I).

On considère maintenant que certains neurones sont excités, c'est-à-dire qu'ils sont amenés à décharger avec une fréquence propre supérieure à celles des autres neurones.

Pour cela la loi \mathcal{F}_{ij} de la v.a. U_{ij} avec laquelle est réinitialisé le neurone (i,j) au moment d'une décharge dépend du site (i,j) .

On suppose que les lois \mathcal{F}_{ij} appartiennent à une même famille de lois (uniforme, exponentielle, normale, dans nos exemples), mais ces lois diffèrent par leurs espérances : $E \mathcal{F}_{ij} < E \mathcal{F}_{i'j'}$ signifiant que le neurone (i,j) est plus excité que le neurone (i',j') .

On modifie alors tout simplement les modèles discrets et continus présentés dans la Partie 1 (II.1), en posant :

Si $X_{ij}^t = 0$, X_{ij}^{t+1} (modèle discret) ou X_{ij}^{t+dt} (modèle continu) est une réalisation d'une v.a. U_{ij} de loi \mathcal{F}_{ij} indépendante de X^t .

On suppose comme pour \mathcal{F} , que les lois \mathcal{F}_{ij} sont à support \mathbb{R}^+ , et sont d'espérances et de variances finies.

Les deux processus ainsi définis (discrets et continus) sont alors des processus de Markov. De même le processus échantillonné aux instants T_ℓ de décharge d'un neurone, noté (X^T_ℓ) est une chaîne de Markov.

La plupart des démonstrations et des résultats de la 1ère partie restent valables :

La chaîne (X^T_ℓ) est irréductible, apériodique et son comportement dépend de la valeur du paramètre d'inhibition θ .

Il faut donc évaluer dans ce nouveau contexte la valeur seuil θ_0 en-dessous de laquelle le système converge.

2. SYSTEME DE DEUX NEURONES.

Notons U , de loi \mathcal{F}_1 , (resp. V , de loi \mathcal{F}_2) la v.a. réinitialisant l'état du neurone (1) (resp. (2)) au moment de leur décharge. On suppose pour fixer les idées que $E \mathcal{F}_1 < E \mathcal{F}_2$. C'est donc le neurone (1) qui est stimulé.

Alors en posant $Z^k = (X_2 - X_1)^k$, comme dans la première partie, on a :

$$Z^{k+1} = Z^k + \mathbb{1}_{Z^k > 0} (\theta - U_{k+1}) + \mathbb{1}_{Z^k < 0} (V_{k+1} - \theta)$$

où U_{k+1} (resp. V_{k+1}) est une v.a. de loi \mathcal{F}_1 (resp. \mathcal{F}_2), indépendante de $(X^T)^k$.

Alors,

- Si $\theta < E \mathcal{F}_1 = \min(E \mathcal{F}_1, E \mathcal{F}_2)$, l'espérance du saut de Z^k est de signe contraire à Z^k .
- Si $E \mathcal{F}_1 < \theta < E \mathcal{F}_2$, l'espérance du saut de Z^k est toujours positif.
- Si $\max(E \mathcal{F}_1, E \mathcal{F}_2) = E \mathcal{F}_2 < \theta$, l'espérance du saut de Z^k est du même signe que Z^k .

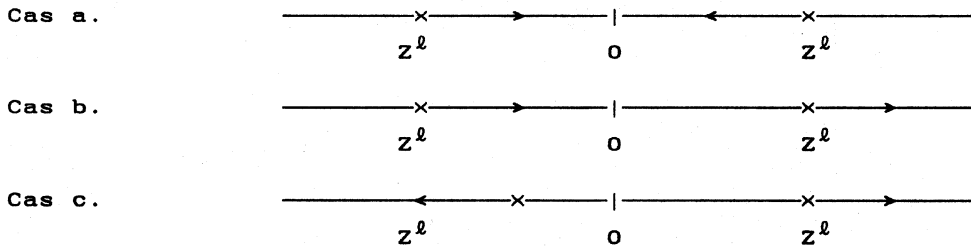


FIGURE 1.

On a alors :

PROPOSITION II.2.1. - Cas de 2 neurones.

On pose $\theta_0 = \min(E \mathcal{F}_1, E \mathcal{F}_2)$

$\theta_1 = \max(E \mathcal{F}_1, E \mathcal{F}_2)$

- i) Lorsque $\theta < \theta_0$, les processus X^T et X^t sont ergodiques.
- ii) Lorsque $\theta_0 < \theta < \theta_1$, $X_2^t \xrightarrow[t \rightarrow +\infty]{\text{p.s.}} +\infty$ et X_1^t converge en loi vers un processus de renouvellement de loi \mathcal{F}_1 .

iii) Lorsque $\theta > \theta_1$, les processus (X^t) et (X^{T^l}) divergent : sur presque toute trajectoire, une des deux composantes tend p.s. vers $+\infty$, l'autre converge en loi vers un processus de renouvellement de loi \mathcal{F}_1 ou \mathcal{F}_2 .

Démonstration

Les cas i) et iii) se déduisent des mêmes arguments que ceux utilisés dans le IV de la première partie.

Supposons $E \mathcal{F}_1 < E \mathcal{F}_2$ et donc $\theta_0 = E \mathcal{F}_1$, $\theta_1 = E \mathcal{F}_2$.

Dans le cas ii), il suffit d'observer que Z^l est une marche aléatoire d'incrément $V-\theta$, $\theta_1-\theta > 0$, tant que $Z^l < 0$, et d'incrément $\theta-U$, d'espérance $\theta - \theta_0 > 0$ quand $Z^l > 0$. \square

On trouvera dans l'appendice 1 les résultats numériques de simulations menées pour différentes lois \mathcal{F}_i et différents θ .

Dans le cas $\theta < \theta_0$, on peut calculer l'espérance de l'intervalle inter-spike d'un neurone, en régime stationnaire.

Reprenons les calculs faits au IV de la première partie. Il faut simplement noter différemment les v.a. tirées à chaque décharge, suivant qu'il s'agit du neurone (1) ou du neurone (2).

Soit K la v.a. entière mesurant le nombre de décharges (éventuellement nul) du neurone (1) entre t_0 et t_1 , instants de décharges consécutives du neurone (2).

Alors on a :

- au temps t_0 , l'état $(x_0, 0)$, et en t_0^+ , $(x_0 + \theta, V_1)$

Si $K = 0$ (si $V_1 < x_0 + \theta$)

- au temps $t_0 + V_1$, l'état est $(x_0 + \theta - V_1, 0)$

Si $K > 0$ (si $V_1 > x_0 + \theta$)

- au temps $t_0 + x_0 + \theta$, l'état est $(0, V_1 - x_0 - \theta)$ et ensuite $(U_1, V_1 - x_0)$

Si $K = 1$ (si $U_1 > V_1 - x_0$)

- au temps $t_0 + V_1 + \theta$, l'état est $(x_0 + U_1 - V_1, 0)$

Si $K > 1$ (si $U_1 < V_1 - x_0$)

- au temps $t_0 + x_0 + \theta + U_1$, l'état est $(0, V_1 - x_0 - U_1)$ puis $(U_2, V_1 - x_0 - U_1 + \theta)$

Si $K = 2$ (si $U_2 > V_1 - x_0 - U_1 + \theta$)

- au temps $t_0 + V_1 + 2\theta$, l'état est $(x_0 + U_1 + U_2 - V_1 - \theta)$

⋮

Donc

Si $K = k$ (si $U_k > V_1 - x_0 - U_1 - \dots - U_k - (k-1)\theta$)
 - au temps $t_0 + V_1 + k\theta$, l'état est $(x_0 + U_1 + U_2 + \dots + U_k - V_1 - (k-1)\theta)$.

L'intervalle inter-spike est donc de la forme

$$\tau = V_1 + K\theta$$

où K est une v.a. entière dont la loi est donnée par

$$P(K > k) = P(V_1 > x_0 + \theta, U_1 < V_1 - x_0, U_1 + U_2 < V_1 - x_0 + \theta, \dots, U_1 + \dots + U_k < V_1 - x_0 + (k-1)\theta)$$

(cela permet de redémontrer que le processus converge ssi $E\mathcal{F}_1 - \theta > 0$ et par symétrie $E\mathcal{F}_2 - \theta > 0$, sinon la v.a. K (et son homologue L , nombre de décharges de (2) entre deux décharges de (1)) tend vers $+\infty$).

Si l'on note Y_0 et Y_1 les états du neurone (1) aux instants t_0 et t_1 , on a :

$$X_1 = U_1 + U_2 + \dots + U_K - V_1 + X_0 - (K-1)\theta$$

et comme en régime stationnaire, $X_1 \sim X_0$ et on obtient puisque K est un temps d'arrêt

$$(EK)(EU - \theta) = EV - \theta$$

d'où

$$EK = \frac{EV - \theta}{EU - \theta}$$

Cela permet de conclure :

PROPOSITION II.2.2.

En régime stationnaire ($\theta < \theta_0$) pour le réseau de deux neurones, les intervalles inter-spikes moyens sont

$$E\mathcal{F}_2 + \frac{E\mathcal{F}_2 - \theta}{E\mathcal{F}_1 - \theta} \theta \quad \text{pour le neurone (2)}$$

$$E\mathcal{F}_1 + \frac{E\mathcal{F}_1 - \theta}{E\mathcal{F}_2 - \theta} \theta \quad \text{pour le neurone (1)}$$

(Voir l'appendice 1).

3. SYSTEME DE n^2 NEURONES

Les résultats de la première partie se généralisent :

PROPOSITION II.3.1.

$$\text{Si } \theta < \theta_0 = \min_{(i,j)} \left\{ \frac{E \mathcal{F}_{ij}}{|V|} \right\}$$

les processus (X^{T_ℓ}) et (X^ℓ) sont ergodiques.

Dans ce cas, on observe sur les simulations que tous les neurones sont actifs, déchargeant bien sûr avec une fréquence plus faible que leur fréquence propre. On trouvera dans l'appendice 2 différents exemples.

On a choisi de stimuler des groupes de neurones disposés, soit en plaques, soit en lignes (cellules grimpantes) par exemple. On remarque que l'intervalle moyen inter-spike de chaque neurone est environ égal à

$$E \mathcal{F}_{ij} + |V| \theta \quad \text{si le neurone } (i,j) \text{ est à l'intérieur de la zone excitée ou de la zone non excitée.}$$

$$E \mathcal{F}_{i'j'} + \frac{E \mathcal{F}_{ij} - \theta}{E \mathcal{F}_{i'j'} - \theta} \theta \quad \text{si le neurone } (i,j) \text{ est à la lisière d'une zone excitée et d'une zone non excitée.}$$

DIVERGENCE.

$$\text{Lorsque } \theta > \theta_1 = \max_{i,j} \left\{ \frac{E \mathcal{F}_{ij}}{|V|} \right\}, \text{ le processus } (X^t)$$

diverge, et se comporte comme dans le cas non stimulé : des bandes de neurones inhibés et actifs s'établissent, brisant le mur d'inhibi-

tion latérale autour de chaque neurone. Bien sûr, l'intervalle inter-spikes moyen de chaque neurone actif est à peu près égal à son intervalle moyen propre $E \mathcal{F}_{ij}$.

Enfin, lorsque $\theta_0 < \theta < \theta_1$, certains neurones sont inhibés, d'autres pas et l'on peut obtenir toute une variété de cartes.

Examinons quelques cas :

- 1) *Un seul neurone* (i_0, j_0) *est excité, $\mathcal{F}_{ij} = \mathcal{F}$ pour tout* $(i, j) \neq (i_0, j_0)$ *et* $E \mathcal{F}_{i_0 j_0} < E \mathcal{F}$. *Alors si* $E \mathcal{F}_{i_0 j_0} / |V| < \theta < E \mathcal{F} / |V|$, *on observe la divergence des seuls voisins de* (i_0, j_0) . *Cela dessine parfaitement la forme de* $V(i_0, j_0)$. *C'est l'inhibition latérale.*

Voir l'appendice 3.

- 2) *Une seule zone* D *(plaque ou ligne) de neurones est excitée, en prenant par exemple*

$$E \mathcal{F}_{ij} = E \mathcal{F}_1 \quad \text{si} \quad (i, j) \in D$$

$$E \mathcal{F}_{ij} = E \mathcal{F}_2 \quad \text{si} \quad (i, j) \notin D$$

avec $E \mathcal{F}_1 < E \mathcal{F}_2$. Alors si $E \mathcal{F}_1 / |V| < \theta < E \mathcal{F}_2 / |V|$, il y a

convergence à l'extérieur de D

divergence à l'intérieur de D

et au bord de D *on observe un périmètre de neurones particulièrement inhibés.*

Voir l'appendice 4.

III - CONCLUSION

Ce qui est intéressant dans le modèle étudié c'est que en renforçant l'inhibition, on brise l'inhibition latérale et on crée des bandes ou moirures de neurones actifs et inhibés en alternance.

Ces bandes alternées des tissus neuronaux que l'on observe dans de nombreux contextes sont donc la conséquence de l'inhibition et non de l'excitation de certains neurones.

Ces phénomènes s'apparentent à ce qui se produit lors de la formation des cristaux : ce sont les contraintes (pression, volume fixe, etc...) qui aboutissent à la formation de cristaux réguliers et de dispositions périodiques.

Les références de cette partie sont les mêmes que celles de la première partie.

APPENDICE 1 - CAS DE DEUX NEURONES

Le neurone 1 est excité

Loi \$ uniforme

$E(\mathcal{F}_1)$		$E(\mathcal{F}_2)$		θ	Intervalle moyen inter-spikes théoriques	Intervalle moyen inter-spikes observés et variances entre ()	Fréquences moyennes observées / 100 unités de temps	Corr. des fréquences	
3	6	0	3.4	6.5	3.2 { 3.0 }	5.6 { 9.9 }	34.5	17.7	0.0
3	6	1	3.5	14	3.1 { 3.5 }	7.6 { 24.3 }	31.1	13.1	-0.4
3	6	3	3	∞	2.9 { 2.0 }	14.3 { 141 }	32.5	6.9	-0.8
3	6	10	∞	∞	$X_1 \rightarrow \infty$	5.7 { 10.7 }			
5	10	0	5.4	10	4.8 { 6.6 }	9.9 { 32.5 }	20.7	10.1	0.1
5	10	1	5.75	12.35	5.1 { 9.0 }	12.3 { 46.5 }	19.5	8.1	-0.3
5	10	2	5.85	15.3	5.4 { 11.2 }	14.6 { 95.0 }	18.5	6.8	-0.4
5	10	3	5.85	20.5	5.4 { 13.2 }	20.7 { 256 }	18.5	4.8	-0.8
5	10	4	5.66	34	5.5 { 15.9 }	27.8 { 1269 }	18.0	3.5	-0.8
5	10	5	5	∞	4.8 { 7.0 }	$X_2 \rightarrow \infty$			
5	10	15	5	∞	$X_1 \rightarrow \infty$	9.9 { 31.3 }			
5	15	3	5.5	33	5.1 { 10.1 }	35.4 { 629 }	19.4	2.8	-0.7
5	15	4	5.36	59	5.1 { 10.0 }	59.3 { 4106 }	19.6	1.6	-0.8
5	15	5	5	∞	4.8 { 7.02 }	$X_2 \rightarrow \infty$			
5	15	25	5	∞	4.8 { 7.02 }	$X_2 \rightarrow \infty$			

Loi \$ exponentielle

$E(\mathcal{F}_1)$		$E(\mathcal{F}_2)$		θ	Intervalle moyen inter-spikes théoriques	Intervalle moyen inter-spikes observés et variances entre ()	Fréquences moyennes observées / 100 unités de temps	Corr. des fréquences	
3	6	0	3.4	6.5	3.0 { 5.5 }	6.5 { 29.6 }	33.3	15.1	0.1
3	6	1	3.5	14	3.4 { 8.5 }	8.4 { 52.3 }	29.8	11.8	-0.6
3	6	3	3	∞	3.6 { 12.5 }	11.8 { 245 }	27.9	8.3	-0.8
3	6	4	3	∞	3.3 { 10.3 }	38.5 { 11296 }	30.5	1.5	-0.8
3	6	10	∞	∞	$X_1 \rightarrow \infty$	$X_2 \rightarrow \infty$			
5	11	0	5.4	11	5.1 { 19.8 }	6.4	19.6	8.9	0.0
5	11	1	5.6	13.5	5.4 { 26.1 }	11.2 { 91.4 }	18.5	7.3	-0.4
5	11	2	5.75	17	5.4 { 25.6 }	13.7 { 120.4 }	18.4	5.3	-0.4
5	11	3	5.6	23	5.7 { 35.5 }	18.6 { 338 }	17.6	4.6	-0.7
5	11	4	5.6	39	5.8 { 38.2 }	21.5 { 695 }	17.3	3.1	-0.7
5	11	5	5	∞	5.8 { 46.5 }	31.1 { 2172 }	17.0	1.7	-0.8
5	11	6	5	∞	5.3 { 20.8 }	41.5 { 9333 }			
5	16	0	5.3	16	5.1 { 20.4 }	$X_2 \rightarrow \infty$	38.9	12.4	-0.1
5	16	1	5.4	19.7	5.6 { 23.6 }	15.8 { 181 }	17.7	5.7	-0.3
5	16	3	5.6	25.3	5.5 { 23.22 }	17.6 { 328 }	36.6	7.8	-0.7
5	16	3	5.5	33	5.5 { 25.7 }	25.3 { 704 }	36.1	5.8	-0.8
5	16	4	5.4	56	5.5 { 28.7 }	33.7 { 1110 }	37.5	2.7	-0.8
5	16	10	5.4	∞	5.3 { 21.7 }	75.1 { 5161 }			

APPENDICE 2 - CAS DE n^2 NEURONES

REGIME STATIONNAIRE

On donne quelques exemples de processus convergents pour $\theta < \min \frac{E(\mathcal{F}_{ij})}{|U|}$: état, fréquences, intervalles inter-spikes moyens des neurones d'un réseau, auto-corrélations et inter-corrélations de quelques neurones observés.

On peut faire les mêmes commentaires que dans l'appendice 4 de la partie I : corrélations négatives entre neurones voisins augmentant en valeur absolue avec θ , auto-corrélations faibles.

Les zones grises représentent les neurones totalement inhibés.

Les zones blanches représentent les neurones actifs et excités.

Les petits cercles marquent les neurones actifs ayant une activité propre normale.

EXEMPLE : 6 VOISINS -UNE PLAQUE DE NEURONES EXCITES

REGIME STATIONNAIRE

Le réseau est de taille 20 * 20

La forme du voisinage à 6 voisins est:
 x
 x x . x x
 x

La loi propre des intervalles inter-spikes est uniforme

Le paramètre d'inhibition teta vaut 4

L'intervalle inter-spikes moyen propre des neurones non excités est 60

L'intervalle inter-spikes moyen propre des neurones excités est 30

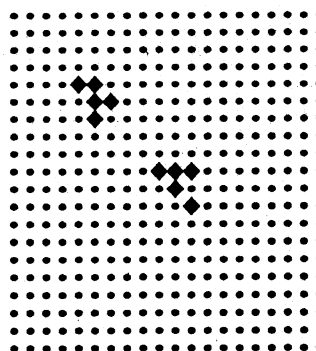
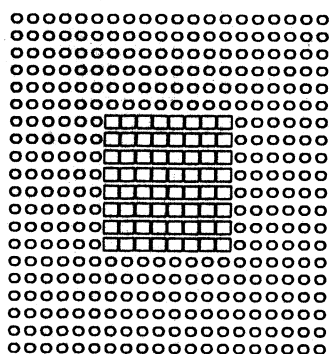
La plaque excitée est définie par: $6 < i < 15, 6 < j < 15$.

Etat du système au bout de 4000 itérations

8	101	80	48	45	28	42	34	28	14	9	91	63	114	63	19	4	41	4	54
105	11	79	38	13	37	66	87	50	18	44	85	105	5	5	70	2	60	12	14
14	108	70	9	59	24	67	97	69	74	30	44	54	96	34	35	9	26	68	40
37	8	94	9	14	76	68	54	6	30	13	88	90	12	51	34	33	8	23	79
30	66	26	11	21	36	38	34	91	34	18	89	52	12	26	83	28	21	96	9
33	6	79	48	105	51	7	61	25	7	22	72	54	26	65	3	14	7	99	28
85	66	8	28	41	30	29	34	6	57	6	14	30	31	27	61	12	38	5	7
10	26	45	99	20	22	4	5	14	13	1	10	3	3	25	9	29	89	41	51
50	27	11	19	17	18	29	3	48	12	2	2	10	12	18	6	31	3	37	38
28	53	64	26	6	15	20	23	19	57	11	12	11	20	21	22	1	41	24	56
36	34	53	7	70	75	39	2	13	14	21	11	6	10	29	6	12	29	20	14
14	35	82	60	44	3	25	17	33	38	25	38	35	5	44	10	39	9	36	30
18	46	2	42	10	44	46	49	13	25	36	6	38	13	46	19	17	27	87	62
18	67	16	31	24	55	9	26	43	12	8	12	10	27	62	25	39	46	24	44
22	90	35	25	17	25	61	59	76	58	2	80	25	17	112	21	76	25	16	20
22	88	26	62	87	33	60	67	84	1	53	43	18	6	35	21	42	37	53	38
73	40	28	10	13	80	69	72	24	20	30	116	114	31	37	42	45	61	29	64
37	45	79	2	39	33	8	68	1	7	47	4	44	6	96	16	51	80	18	45
18	48	20	7	2	62	77	9	53	85	60	17	22	112	48	16	76	20	1	8
49	40	13	42	43	72	11	3	50	27	71	53	20	24	57	60	25	44	17	3

Tous les neurones sont actifs.

Les neurones observés sont marqués



Intervalles moyens inter-spikes

Coordonnées	Intervalle	Variance
(10,10)	58.0	1333.40
(11,11)	55.3	1131.37
(10,11)	52.4	1292.92
(12,12)	53.1	1180.19
(10,12)	47.4	1123.16
(5, 5)	88.1	2395.73
(5, 6)	91.2	2997.99
(6, 6)	94.3	2211.87
(6, 7)	78.8	2625.15
(7, 6)	85.4	3208.38

10 premières autocorrélations des neurones

Coordonnées	Autocorrélations									
(10,10)	.25	.09	.01	.08	.23	.38	.07	-.10	-.01	.19
(11,11)	-.06	.06	.08	.09	-.35	.15	-.17	-.07	.02	-.03
(10,11)	.14	-.07	-.11	.37	.21	-.06	-.02	.11	.17	-.05
(12,12)	.03	.05	.01	.02	.13	-.09	-.02	-.21	-.14	-.30
(10,12)	-.30	.19	-.12	-.27	.08	-.31	.01	-.11	.03	.08
(5, 5)	.00	-.08	-.15	.01	.64	.11	.07	-.17	-.14	.22
(5, 6)	.33	.04	.02	.08	.17	-.14	.06	.08	-.47	-.34
(6, 6)	-.08	-.20	-.15	-.16	.01	-.02	-.33	-.52	-.26	-.30
(6, 7)	-.23	-.24	-.05	.06	-.22	-.16	-.06	-.02	-.26	-.07
(7, 6)	-.04	-.03	.07	-.12	.02	-.41	-.19	-.44	-.26	.11

Nombre moyen de spikes par fenêtre de 100

Coordonnées	Nombre	Variance
(10,10)	1.7	1.05
(11,11)	1.8	.90
(10,11)	1.9	.92
(12,12)	1.9	1.07
(10,12)	2.1	1.37
(5, 5)	1.1	.52
(5, 6)	1.1	.90
(6, 6)	1.1	.45
(6, 7)	1.2	.69
(7, 6)	1.1	.58

Matrice d'intercorrélations des neurones observés

(10,10)	1.00									
(11,11)	-.11	1.00								
(10,11)	-.39	.03	1.00							
(12,12)	-.08	-.20	-.10	1.00						
(10,12)	-.30	-.04	-.20	.05	1.00					
(5, 5)	.20	-.21	-.05	-.05	-.11	1.00				
(5, 6)	.00	.09	.11	-.07	-.31	.02	1.00			
(6, 6)	.02	-.21	-.21	.05	-.11	-.13	-.05	1.00		
(6, 7)	-.17	.13	.15	.10	.05	.42	.15	-.28	1.00	
(7, 6)	-.17	-.08	-.05	.02	.18	.07	-.34	.04	-.10	1.00

(10,10) (11,11) (10,11) (12,12) (10,12) (5, 5) (5, 6) (6, 6) (6, 7) (7, 6)

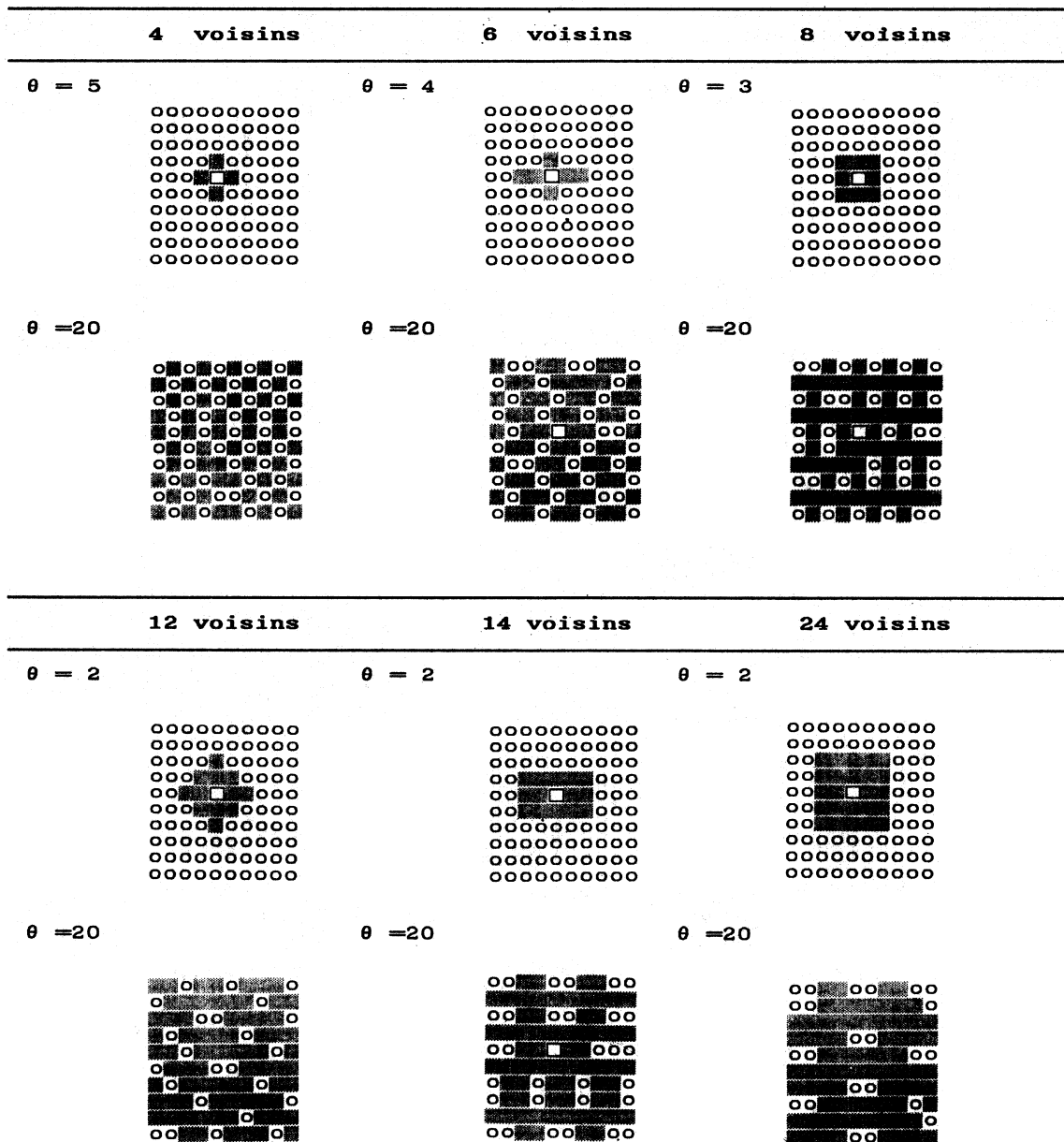
APPENDICE 3 - CAS DE n^2 NEURONES

1 SEUL NEURONE EXCITE

Seul le neurone central est excité, de loi \mathcal{F}_1 avec $E\mathcal{F}_1 < E\mathcal{F}$.

Ci-dessous on a représenté l'état du système (10×10) au bout de 4000 itérations, pour $E\mathcal{F}_1 = 10$, $E\mathcal{F} = 30$, et différentes formes de voisinages.

Dans chaque cas on prend $\theta_0 < \theta < \theta_1$, puis $\theta > \theta_1$.



Les zones grises représentent les neurones totalement inhibés.
 Les zones blanches représentent les neurones actifs et excités.
 Les petits cercles marquent les neurones actifs ayant une activité propre normale.

APPENDICE 4 - CAS DE n^2 NEURONES

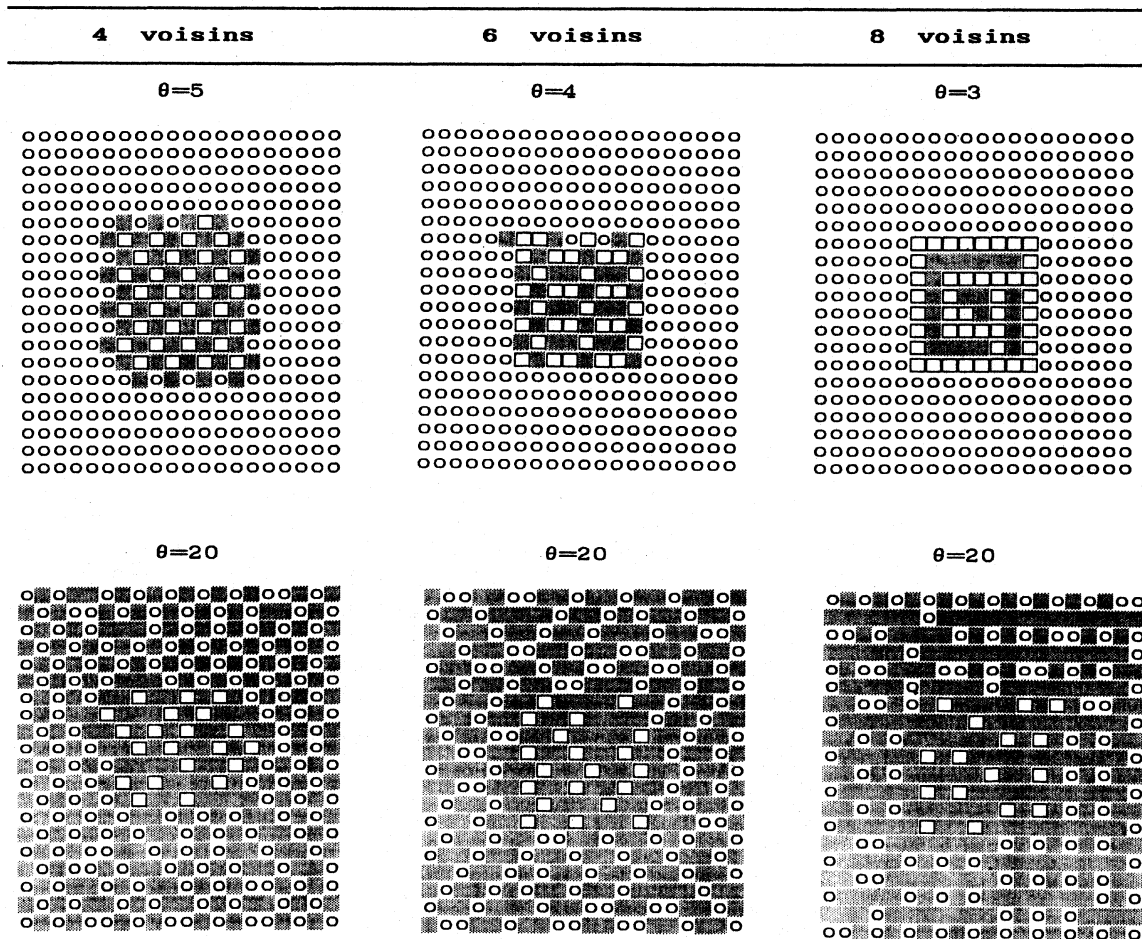
UNE PLAQUE DE NEURONES EXCITES

On excite une "plaque" de neurones en les réinitialisant avec la loi \mathcal{F}_1 , telle que $E\mathcal{F}_1 < E\mathcal{F}$.

Ci-dessous on a représenté l'état du système (20×20) au bout de 4000 itérations, pour $E\mathcal{F}_1 = 10$, $E\mathcal{F} = 30$ et différentes formes de voisinages.

On prend $\theta_0 < \theta < \theta_1$ puis $\theta > \theta_1$.

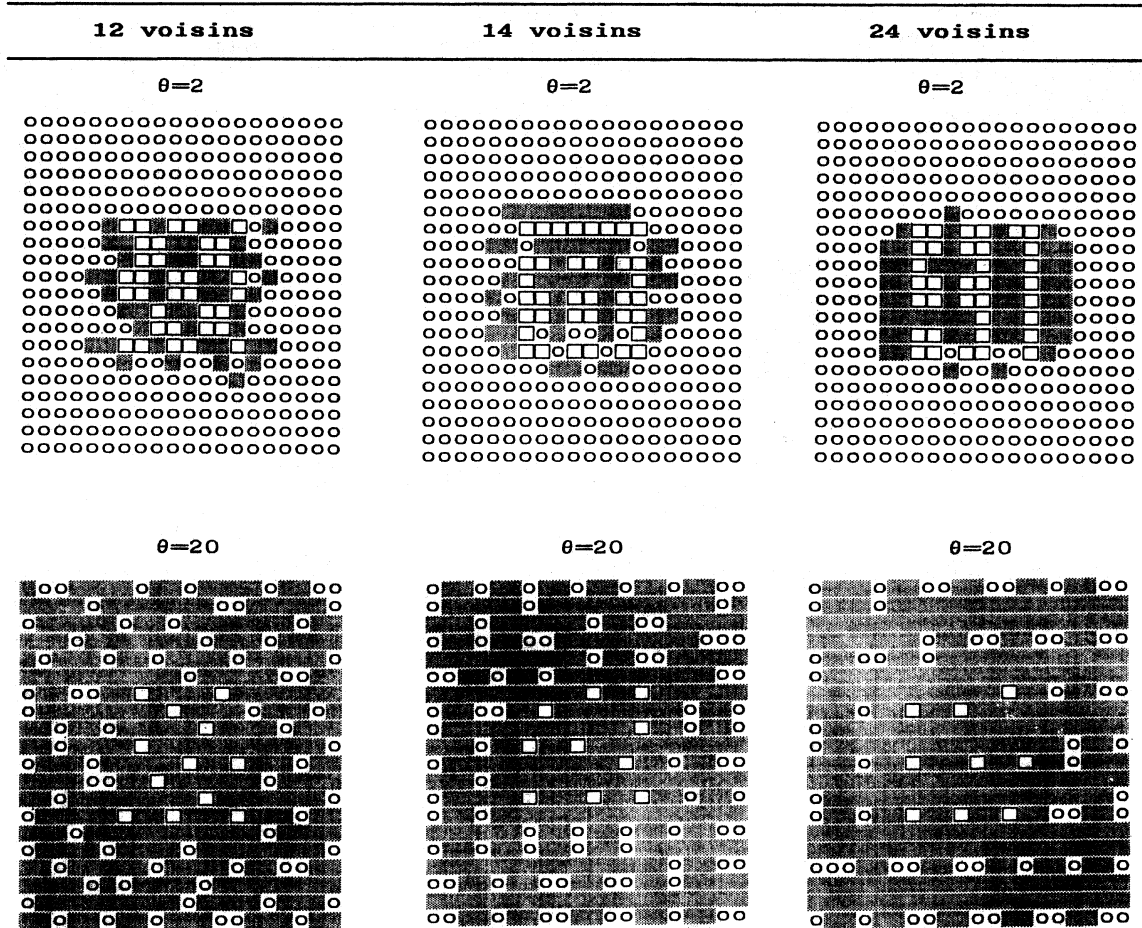
Plaque excitée : $7 < i < 14$, $7 < j < 14$.



Les zones grises représentent les neurones totalement inhibés.

Les zones blanches représentent les neurones actifs et excités.

Les petits cercles marquent les neurones actifs ayant une activité propre normale.



Les zones grises représentent les neurones totalement inhibés.
 Les zones blanches représentent les neurones actifs et excités.
 Les petits cercles marquent les neurones actifs ayant une activité propre normale.

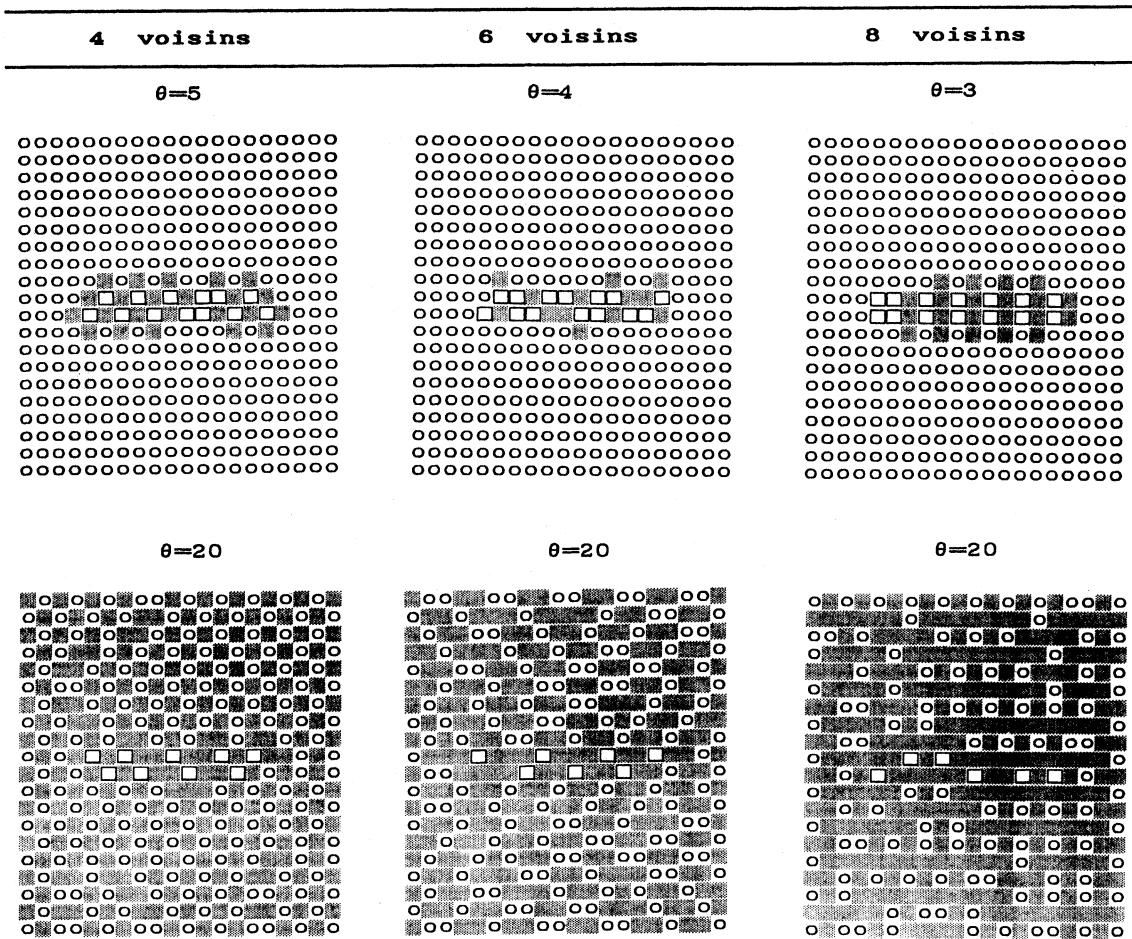
DEUX LIGNES DE NEURONES EXCITES

On excite une "plaque" de neurones en les réinitialisant avec la loi \mathcal{F}_1 , telle que $E\mathcal{F}_1 < E\mathcal{F}$.

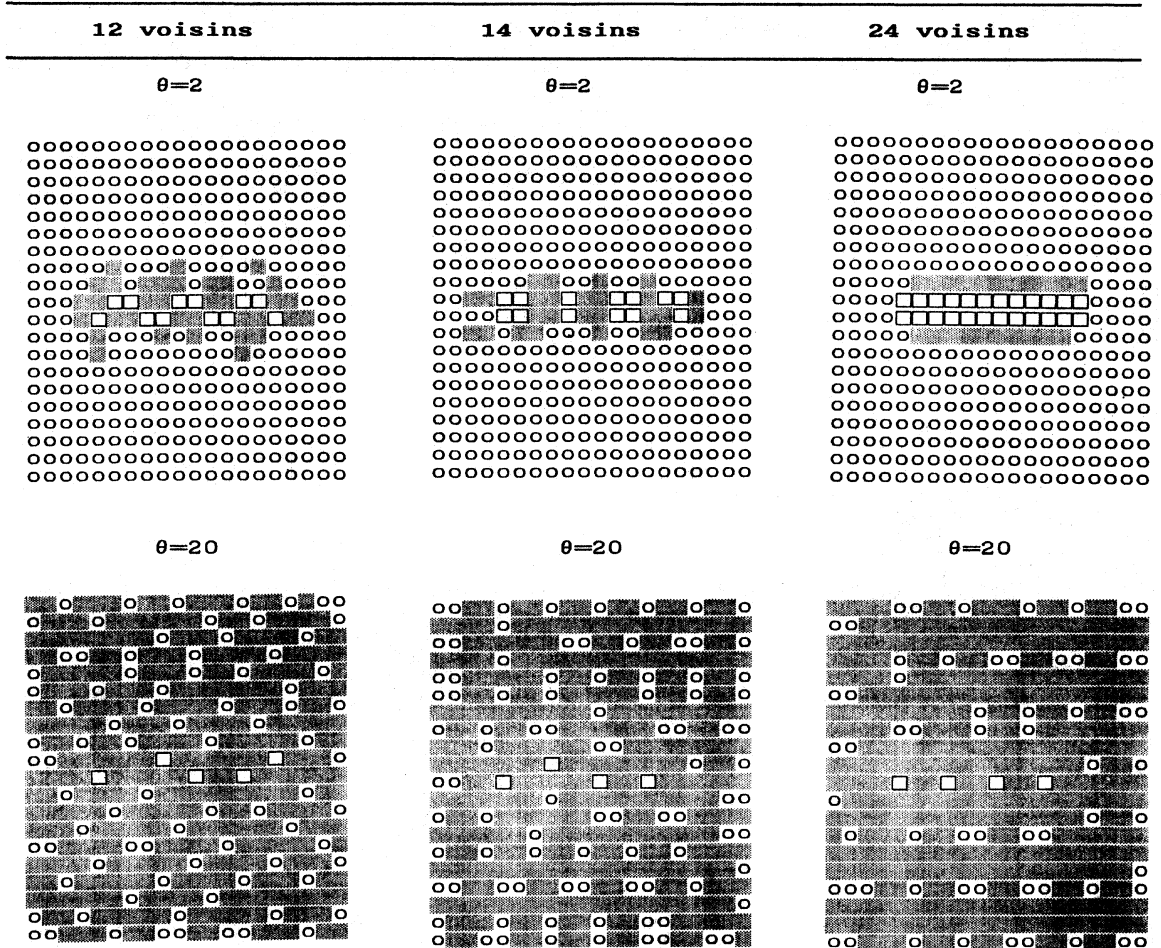
Ci-dessous on a représenté l'état du système (20×20) au bout de 4000 itérations, pour $E\mathcal{F}_1 = 10$, $E\mathcal{F} = 30$ et différentes formes de voisinages.

On prend $\theta_0 < \theta < \theta_1$ puis $\theta > \theta_1$.

Lignes excitées : $i=10,11, 5 \leq j \leq 16$



Les zones grises représentent les neurones totalement inhibés.
 Les zones blanches représentent les neurones actifs et excités.
 Les petits cercles marquent les neurones actifs ayant une activité propre normale.



Les zones grises représentent les neurones totalement inhibés.
 Les zones blanches représentent les neurones actifs et excités.
 Les petits cercles marquent les neurones actifs ayant une activité propre normale.

N° d'impression : 987

3EME Trimestre 1988