



Travail Encadré de Recherche

Simulation par utilisation des Chaînes de Markov

DOURDOIGNE Matthieu et LACOUR Julien

Année Académique 2018-2019

Table des matières

1	Introduction	2
2	Rappels sur les Chaînes de Markov	3
3	Les méthodes	4
3.1	Monte Carlo Markov Chains	4
3.2	L'échantillonneur de Gibbs	5
3.3	L'échantillonneur de Gibbs à balayage constant	6
3.4	Les chaînes de Métropolis	6
3.5	L'algorithme de Propp-Wilson	7
3.6	Sandwiching, ou comment simplifier l'algorithme de Propp-Wilson	9
3.7	Recuit simulé (Simulated annealing)	9
4	Les modèles ou problèmes	12
4.1	Le modèle des sphères dures à deux dimensions	12
4.1.1	Mise en œuvre sous Scilab	12
4.2	Modèle d'Ising	13
4.2.1	Mise en œuvre sous Scilab	14
4.3	Résultats obtenus	17
4.3.1	Résultats pour le modèle des sphères dures	17
4.3.2	Résultats pour le modèle d'Ising	17
5	Conclusion	18

1 Introduction

Lors de ce travail de recherche encadré, nous avons découvert plusieurs méthodes de simulation de loi s'appuyant sur les Chaînes de Markov, en particulier par l'intermédiaire de plusieurs problèmes connus tels que le Problème des Sphères dures, le Problème du voyageur et le modèle d'Ising.

Mais que signifie "simuler une loi" ?

Lorsqu'on parle de simulation de loi, il faut en fait entendre, "simulation d'une variable aléatoire suivant une certaine loi" dans un premier temps. Ensuite, les méthodes de simulation, dites classiques, sont des algorithmes dont les sorties (X_1, X_2, X_3, \dots) représentent des réalisations des variables aléatoires distribuées selon la loi de probabilité que l'on étudie.

Il y a d'autres méthodes de simulation de loi, dont une que l'on présentera plus en détails par la suite, à savoir, les méthodes de simulation Monte Carlo Markov Chains, plus communément appelées "méthodes de simulation *MCMC*". En quelques mots, on aura alors un algorithme simulant une Chaîne de Markov ayant pour loi stationnaire, la loi de probabilité recherchée.

Au cours de nos recherches, nous nous sommes plus particulièrement penchés sur deux sujets, à savoir le problème des Sphères dures, ainsi que le modèle d'Ising, tous deux en dimension 2, pour lesquels nous proposerons des programmes de simulation permettant l'étude des dits problèmes.

Dans un premier temps, nous ferons quelques rappels sur les Chaînes de Markov, qui sont nécessaires à la justification des méthodes étudiées. Notamment, nous rappellerons certaines propriétés telles que l'irréductibilité, la réversibilité ou encore l'apériodicité, mais également des théorèmes, tels que les théorèmes ergodique et de convergence en loi...

Ensuite, nous amènerons ainsi les différentes méthodes de simulation que nous avons découvert en justifiant au mieux la construction de celles-ci à l'aide des rappels réalisés au préalable.

Dans la troisième partie de ce mémoire, nous présenterons les modèles et problèmes que nous avons étudié lors de ce travail de recherche.

Enfin, nous présenterons une synthèse de ce qui a été fait précédemment en réalisant une étude de certains problèmes se basant sur les résultats issus de nos programmes avec des représentations.

2 Rappels sur les Chaînes de Markov

Chaîne récurrente positive : Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov à valeurs dans S , un ensemble fini ou dénombrable, définie sur $(\Omega, \mathcal{F}, \mathbb{P})$ et irréductible.

On dit que la chaîne est récurrente positive si tous ses états sont récurrents positifs. On rappelle qu'un élément s de S est dit récurrent positif si $\mathbb{E}_s[T_s]$ est fini, avec T_s le temps d'atteinte de l'état s .

Chaîne apériodique : Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov à valeurs dans S , un ensemble fini ou dénombrable, définie sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, irréductible, récurrente positive et d'unique loi stationnaire π . On dit que $(X_n)_{n \in \mathbb{N}}$, ou que P sa matrice de transition, est apériodique si et seulement s'il existe un état s de S tel que $\text{pgcd}\{n \in \mathbb{N}^*; p_{i,i}^{(n)} > 0\} = 1$.

Loi de probabilité réversible pour une chaîne de Markov : Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov à valeurs dans $S = \{s_1, s_2, \dots, s_k\}$, de matrice de transition P . Une loi de probabilité π sur S est dite réversible pour la chaîne $(X_n)_{n \in \mathbb{N}}$, si on a :

$$\forall i, j \in \{1, \dots, k\}, \quad \pi_i P_{i,j} = \pi_j P_{j,i}$$

Théorème loi réversible et stationnaire Soit $S = \{s_1, s_2, \dots, s_k\}$. Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov à valeurs dans S , de matrice de transition P . Si π est une loi de probabilité réversible pour la chaîne $(X_n)_{n \in \mathbb{N}}$, alors c'est aussi une loi stationnaire pour la chaîne.

Démonstration : Rappelons qu'une loi de probabilité π est stationnaire pour la chaîne de Markov à valeurs dans S , de matrice de transition P , si elle vérifie $\pi P = \pi$, plus précisément, si on a :

$$\forall j \in \{1, \dots, k\}, \quad \sum_{i=1}^k \pi_i P_{i,j} = \pi_j$$

Soit j dans $\{1, \dots, k\}$, on a :

$$\pi_j = \pi_j \sum_{i=1}^k P_{j,i} = \sum_{i=1}^k \pi_j P_{j,i} = \sum_{i=1}^k \pi_i P_{i,j}$$

Ainsi, si π est une loi réversible sur S , c'est une loi stationnaire.

Théorème ergodique : Soit $S = \{s_1, \dots, s_k\}$ un ensemble fini et soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov irréductible, récurrente positive et à valeurs dans S . On note π son unique loi stationnaire. Alors, pour toute fonction $f : S \rightarrow \mathbb{R}$ π -intégrable, pour toute loi initiale μ_0 sur S , on a :

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) = \mathbb{E}_\pi[f(X_0)] = \sum_{s \in S} \pi(s) f(s) \quad \mathbb{P}_{\mu_0} - p.s.$$

Théorème de convergence en loi vers la loi stationnaire : Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov à valeurs dans S , un ensemble fini ou dénombrable, définie sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, irréductible, récurrente positive et apériodique. Soit π son unique loi stationnaire.

Alors pour toute loi initiale μ_0 sur S , pour tout élément s de S ,

$$\mathbb{P}_{\mu_0}(X_n = s) \xrightarrow{n \rightarrow +\infty} \pi(s) \Leftrightarrow (X_n)_{n \in \mathbb{N}} \text{ converge en loi vers } \pi \text{ quand } n \text{ tend vers } +\infty.$$

3 Les méthodes

3.1 Monte Carlo Markov Chains

Le principe des Monte Carlo Markov Chains, comme dit précédemment, est de simuler non pas un échantillon d'une variable aléatoire suivant la loi étudiée, mais une chaîne de Markov, ayant certaines propriétés et dont la loi recherchée, que l'on va nommer π pour la suite, est la loi stationnaire.

On va se placer dans l'ensemble $S = \{s_1, s_2, \dots, s_k\}$.

Dans le cadre général, pour simuler une variable aléatoire X distribuée suivant la loi π sur l'ensemble S , on procède de la manière suivante : on commence par énumérer les éléments de S : s_1, s_2, \dots, s_k et on définit X ainsi

$$X = \Psi(U)$$

avec U une variable aléatoire suivant la loi uniforme sur $[0,1]$ et $\Psi : [0,1] \rightarrow S$ la fonction définie par :

$$\Psi(u) = \begin{cases} s_1 & \text{si } u \in [0, \pi(s_1)[\\ s_2 & \text{si } u \in [\pi(s_1), \pi(s_1) + \pi(s_2)[\\ \cdot & \\ \cdot & \\ \cdot & \\ s_i & \text{si } u \in [\sum_{l=1}^{i-1} \pi(s_l), \sum_{l=1}^i \pi(s_l)[\\ \cdot & \\ \cdot & \\ \cdot & \\ s_k & \text{si } u \in [\sum_{l=1}^{k-1} \pi(s_l), 1] \end{cases}$$

En effet, la variable X ainsi construite suit bien la loi π , en utilisant la fonction indicatrice, on peut facilement vérifier que $\mathbb{P}(X = s) = \pi(s)$:

$$\begin{aligned} \mathbb{P}(X = s) &= \mathbb{P}(\Psi(U) = s) = \int_0^1 \mathbf{1}_{\{\Psi(U)=s\}}(x) \, dx \\ &= \pi(s) \end{aligned}$$

C'est à partir de cette méthode que l'on va construire la chaîne de Markov.

Cependant, il y a un problème avec ce genre de simulation, à savoir, lorsque l'ensemble S comporte un grand nombre d'éléments, l'évaluation de Ψ prend beaucoup de temps. C'est dans ces conditions que les méthodes MCMC sont utiles.

L'idée est la suivante : on suppose que l'on peut construire une chaîne de Markov irréductible et apériodique dont la loi stationnaire est π . Puisque S est fini, on peut appliquer le théorème de convergence en loi qui nous assure donc qu'en parcourant la chaîne pendant un temps assez long, sa distribution va tendre vers π .

Ceci nous donne donc une approximation de la loi π , mais une approximation que l'on peut rendre aussi précise que l'on souhaite en faisant croître n .

Dans la plupart des méthodes MCMC, on va construire la chaîne de sorte que sa loi stationnaire ait une condition encore plus forte, la réversibilité. Ceci permet de construire la chaîne plus facilement.

Dans la section suivante, nous allons vous présenter une des grandes classes de méthode MCMC, à savoir l'Echantillonneur de Gibbs.

3.2 L'échantillonneur de Gibbs

Les algorithmes du type de l'échantillonneur de Gibbs sont très utiles dans la simulation de loi π lorsque l'espace d'état est de la forme S^V (à l'image du modèle des sphères dures ou du modèle d'Ising), c'est-à-dire un espace d'état dont les éléments sont des configurations possibles sur un graphe dont les sommets forment l'espace V et S est l'ensemble des valeurs pouvant être prises en ces sommets.

L'échantillonneur de Gibbs donne une chaîne de Markov dont la transition d'un instant n à un instant $n+1$ se fait comme suit :

1. Choisir un sommet $v \in V$ de façon aléatoire.
2. Attribuer une nouvelle valeur (ou non) à v (notée $X_{n+1}(v)$) suivant la valeur de la loi de distribution π en v et les contraintes liées aux autres sommets (exemple du modèle des sphères dures : il n'est pas possible d'assigner 1 à un sommet si un voisin a déjà cette valeur).
3. Laisser les valeurs des autres sommets comme dans la configuration X_n , autrement dit, $X_{n+1}(w) = X_n(w), \forall w \in V; w \neq v$.

Si cet algorithme nous donne une chaîne de Markov apériodique et irréductible, comme π est une loi réversible pour celle-ci donc, d'après le théorème rappelé plus tôt, c'est la loi stationnaire de la chaîne.

Montrons que π est bien une loi réversible pour la chaîne ainsi construite. Il faut donc montrer :

$$\forall \xi_0, \xi_1 \in S^V, \pi_{\xi_0} \mathbb{P}(X_1 = \xi_1 | X_0 = \xi_0) = \pi_{\xi_1} \mathbb{P}(X_1 = \xi_0 | X_0 = \xi_1)$$

Soit ξ, ξ' deux éléments de S^V . On pose $p_{\xi, \xi'} := \mathbb{P}(X_1 = \xi' | X_0 = \xi)$.

Si ξ et ξ' sont différents mais ne diffèrent pas de la valeur d'un seul sommet, d'après la construction, on a $p_{\xi, \xi'} = p_{\xi', \xi} = 0$, ce qui n'implique rien sur π_ξ ou $\pi_{\xi'}$.

Si $\xi = \xi'$, il est évident, que l'on a ce qui est demandé, à savoir : $\pi_i p_{\xi, \xi} = \pi_i p_{\xi, \xi}$.

Enfin, si ξ et ξ' sont différents et diffèrent seulement de la valeur d'un seul sommet, on a : Supposons que j soit le sommet pour lequel ξ et ξ' diffèrent.

On pose la relation d'équivalence : $\xi \simeq_j \xi'$ dans le cas où ξ et ξ' sont deux éléments de S^V qui vérifient : $\forall i \in V, i \neq j : \xi(i) = \xi'(i)$. Alors on a :

$$p_{\xi, \xi'} = \frac{1}{\text{Card}(S^V)} \frac{\pi(\xi')}{\sum_{\substack{\gamma \in S^V \\ \gamma \simeq_j \xi}} \pi(\gamma)}$$

d'où :

$$\begin{aligned}
\pi_{\xi} p_{\xi, \xi'} &= \pi_{\xi} \frac{1}{\text{Card}(S^V)} \frac{\pi(\xi')}{\sum_{\substack{\gamma \in S^V \\ \gamma \simeq_j \xi}} \pi(\gamma)} = \frac{1}{\text{Card}(S^V)} \frac{\pi_{\xi'} \pi(\xi)}{\sum_{\substack{\gamma \in S^V \\ \gamma \simeq_j \xi'}} \pi(\gamma)} \\
&= \pi_{\xi'} \frac{1}{\text{Card}(S^V)} \frac{\pi(\xi)}{\sum_{\substack{\gamma \in S^V \\ \gamma \simeq_j \xi'}} \pi(\gamma)} = \pi_{\xi'} p_{\xi', \xi}
\end{aligned}$$

et ainsi on a bien la réversibilité de la loi π pour la chaîne.

3.3 L'échantillonneur de Gibbs à balayage constant

Il existe également une autre version de l'échantillonneur de Gibbs, cependant nous n'obtenons pas les mêmes propriétés sur la chaîne construite à partir de celle-ci.

En effet, dans cette version le choix de l'élément qui peut être modifié, au cours d'une étape de l'algorithme, ne se fait pas de manière aléatoire, mais de façon cyclique. Autrement dit, on ordonne les sommets de V et à chaque étape d'assignation d'une nouvelle valeur pour un sommet, on passe au sommet suivant.

Le problème avec cette méthode est que l'on obtient une chaîne de Markov non-homogène.

3.4 Les chaînes de Métropolis

Dans le cadre de la construction d'une chaîne réversible pour les méthodes MCMC, une procédure courante est la construction d'une chaîne de Metropolis.

Le contexte est le même qu'au début, nous cherchons à fabriquer une chaîne de Markov servant à simuler la loi de probabilité π sur $S = \{s_1, \dots, s_k\}$.

Commençons par construire un graphe G dont l'ensemble des sommets est S . Pour l'ensemble des arêtes E , il est pris arbitrairement, mais doit tout de même vérifier deux conditions :

- i) le graphe doit être fait de sorte que l'on ait l'irréductibilité de la chaîne ,
- ii) les sommets ne doivent pas être le point d'arrivée d'un trop grand nombre d'arêtes.

Le seconde point sert à prévenir un trop grand nombre de calculs qui rendrait alors trop lourd la siulation.

On a la chaîne de Metropolis, pour un graphe G donné, de matrice de transition P , dont les éléments sont donnés par :

$$P_{i,j} = \begin{cases} \frac{1}{d_i} \min\left\{\frac{\pi_j d_i}{\pi_i d_j}, 1\right\} & \text{si } s_i \neq s_j \text{ et } \langle s_i, s_j \rangle \in E \\ 1 - \sum_{\substack{l \\ \langle s_i, s_l \rangle \in E}} \frac{1}{d_i} \min\left\{\frac{\pi_l d_i}{\pi_i d_l}, 1\right\} & \text{si } i = j \\ 0 & \text{si } s_i \neq s_j \text{ et } \langle s_i, s_j \rangle \notin E \end{cases}$$

On rappelle que deux sommets sont dits *voisins* si le segment, dont ils sont les sommets, est une arête de E et on note d_i le nombre de voisins du sommet s_i et ceux pour i allant de 1 à k .

Ce qui nous donne en fait la transition suivante pour la chaîne :
 Soit $n \in \mathbb{N}^*$, soit $i_0 \in \{1, \dots, k\}$. Supposons que $X_n = s_{i_0}$. On a :

$$X_{n+1} = \begin{cases} s_j & \text{avec probabilité } \min\{\frac{\pi_j d_i}{\pi_i d_j}, 1\} \\ s_i & \text{avec probabilité } 1 - \min\{\frac{\pi_j d_i}{\pi_i d_j}, 1\} \end{cases}$$

Montrons maintenant que la loi π est réversible, ce qui, d'après les rappels fais en amont, implique que π est l'unique loi stationnaire. Nous allons donc montrer que π vérifie bien le critère de réversibilité.

Soit $i, j \in \{1, \dots, k\}$, on a : pour $i = j$, $\pi_i P_{i,i} = \pi_i P_{i,i}$ est vraie.

Pour $i \neq j$, si s_i et s_j ne sont pas voisins, on a $P_{i,j} = P_{j,i} = 0$, donc cela n'implique aucune contrainte sur π_i et π_j .

Enfin, lorsque $i \neq j$ et s_i et s_j sont voisins, regardons la critère de réversibilité suivant la valeur de $\frac{\pi_i d_j}{\pi_j d_i}$, en particulier suivant si cette fraction est supérieure ou non à 1 :

si $\frac{\pi_i d_j}{\pi_j d_i} \geq 1$, on a :

$$\begin{cases} \pi_i P_{i,j} = \pi_i \frac{1}{d_i} \frac{\pi_j d_i}{\pi_i d_j} = \pi_j \frac{1}{d_j} \\ \pi_j P_{j,i} = \pi_j \frac{1}{d_j} \end{cases}$$

et si $\frac{\pi_i d_j}{\pi_j d_i} < 1$, on a :

$$\begin{cases} \pi_i P_{i,j} = \pi_i \frac{1}{d_i} \\ \pi_j P_{j,i} = \pi_j \frac{1}{d_j} \frac{\pi_i d_j}{\pi_j d_i} = \pi_i \frac{1}{d_i} \end{cases}$$

Le critère de réversibilité est donc bien vérifié, ainsi π est une loi réversible pour la chaîne et donc une loi stationnaire. Ainsi, la chaîne de Metropolis peut-être utilisée pour les méthodes MCMC.

3.5 L'algorithme de Propp-Wilson

Par rapport aux autres méthodes MCMC, l'algorithme de Propp-Wilson est *exact*, autrement dit, les simulations que l'on effectue à l'aide de cet algorithme ne suivent pas une distribution qui approxime la loi d'étude, mais suivent exactement la loi π , ce qui permet d'obtenir des échantillons exacts. Cet algorithme possède également l'avantage considérable de déterminer de lui même quand s'arrêter, ce qui évite de devoir faire des études de vitesse de convergence. Cependant, cet algorithme demande de simuler plusieurs copies d'une même chaîne, ce qui peut poser des problèmes de mémoire dans le cas d'une mise en application sur machine.

On se place encore dans le contexte où l'ensemble est $S = \{s_1, \dots, s_k\}$ et que l'on souhaite étudier la distribution π . Comme pour toute méthode MCMC, nous allons construire une chaîne de Markov irréductible et apériodique sur S dont la mesure stationnaire est la loi d'étude π .

L'algorithme (très peu intuitif) fonctionne de la manière suivante : L'idée ici sera de "coupler par le passé", ce qui signifie que l'on prendra des temps de départ négatifs et qu'on mettra à jour la chaîne jusque 0, plutôt que de partir de 0 pour continuer indéfiniment. Soit $f : S \times [0, 1] \rightarrow S$ une fonction de mise à jour (construite selon la loi π), $(N_i)_{i \in \mathbb{N}^*}$ une suite de nombres entiers positifs strictement croissante (dans notre cas et pour la preuve plus bas, nous avons $N_i = 2^{i-1}$) et $(U_{-i})_{i \in \mathbb{N}}$ une suite de nombres i.i.d, distribués uniformément sur l'intervalle $[0, 1]$.

1. Prendre $i = 1$;
2. $\forall s \in S$, simuler la chaîne depuis $-N_m$ jusqu'à 0 avec la fonction f et les valeurs $U_{-N_i+1}, U_{-N_i+2}, \dots, U_{-1}, U_0$. On simule donc k chaînes ;
3. Si les k chaînes de la phase 2 terminent toutes sur le même s_j , l'algorithme renvoie s_i ;
4. Sinon, on prendra $i \leftarrow i + 1$ et on retourne en phase 2.

Une remarque très importante est que les paramètres $(U_{-i})_{i \in \mathbb{N}}$ ne changent pas entre les étapes : par exemple pour $m=3$, pour les mises à jour vont reprendre $U_{-3}, U_{-2}, U_{-1}, U_0$ qui sont utilisés lors de l'étape précédente, mais prendront les nouvelles valeurs $U_{-7}, U_{-6}, U_{-5}, U_{-4}$ pour les premières mises à jour à cette étape.

On obtient le résultat suivant :

Supposons que l'algorithme ci-dessus se termine avec probabilité 1. Soit une chaîne de Markov sur $S = \{s_1, \dots, s_k\}$ ayant f comme fonction de mise à jour et $\pi = \pi_1, \dots, \pi_k$ comme mesure stationnaire. Prenons $N_i = 2^{i-1}$ et $(U_{-i})_{i \in \mathbb{N}}$ dans l'algorithme ci-dessus et notons X sa sortie. Alors :

$$\forall i \in \{1, \dots, k\}, \mathbb{P}(X = s_i) = \pi_i$$

En effet, fixons un $s \in S = \{s_1, \dots, s_k\}$. Il suffira de prouver que $\forall \epsilon > 0, |\mathbb{P}(X = s_i) - \pi_i| \leq \epsilon$. Fixons donc $\epsilon > 0$. Comme l'algorithme se termine avec probabilité 1, $\exists M \in \mathbb{N}$ suffisamment grand tel que :

$$\mathbb{P}(A = \text{"l'algorithme n'a pas besoin de prendre } m > M") \geq 1 - \epsilon$$

Construisons maintenant une nouvelle chaîne de Markov en utilisant la même fonction f et les mêmes variables $(U_{-i})_{i \in \{-N_m+1, \dots, 0\}}$ mais en choisissant maintenant l'état initial suivant la distribution π et notons \tilde{X} la valeur de la chaîne à l'état 0.

Comme π est stationnaire, \tilde{X} suit la loi π . Si en plus l'algorithme n'a pas besoin d'aller plus loin que $m = M$, nous pouvons affirmer que $X = \tilde{X}$. Ainsi,

$$\mathbb{P}(X \neq \tilde{X}) < \epsilon$$

De cette façon, on obtient :

$$\begin{aligned}
\epsilon &\geq \mathbb{P}(X \neq \tilde{X}) = \mathbb{P}\left(\bigcup_{s \in S} \{X = s \cap \tilde{X} \neq s\}\right) \\
&\geq \mathbb{P}(X = s_i \cap \tilde{X} \neq s_i) \\
&\geq \mathbb{P}(X = s_i) \mathbb{P}(\tilde{X} \neq s_i) \\
&= \mathbb{P}(X = s_i) (1 - \mathbb{P}(\tilde{X} = s_i)) \\
&= \mathbb{P}(X = s_i) - \mathbb{P}(X = s_i) \mathbb{P}(\tilde{X} = s_i) \\
&\geq \mathbb{P}(X = s_i) - \mathbb{P}(\tilde{X} = s_i) \\
&= \mathbb{P}(X = s_i) - \pi_i
\end{aligned}$$

Mais également :

$$\begin{aligned}
\epsilon &\geq \mathbb{P}(X \neq \tilde{X}) = \mathbb{P}\left(\bigcup_{s \in S} \{\tilde{X} = s \cap X \neq s\}\right) \\
&\geq \mathbb{P}(\tilde{X} = s_i \cap X \neq s_i) \\
&\geq \mathbb{P}(\tilde{X} = s_i) \mathbb{P}(X \neq s_i) \\
&= \mathbb{P}(\tilde{X} = s_i)(1 - \mathbb{P}(X = s_i)) \\
&= \mathbb{P}(\tilde{X} = s_i) - \mathbb{P}(X = s_i) \mathbb{P}(X = s_i) \\
&\geq \mathbb{P}(\tilde{X} = s_i) - \mathbb{P}(X = s_i) \\
&= \pi - \mathbb{P}(X = s_i)
\end{aligned}$$

Ce qui montre bien que $|\mathbb{P}(X = s_i) - \pi_i| < \epsilon$ ce qui conclut la preuve.

3.6 Sandwiching, ou comment simplifier l'algorithme de Propp-Wilson

L'algorithme de Propp-Wilson, bien qu'exact, possède un grand désavantage : il demande de simuler autant de chaînes de Markov que S possède d'éléments. Ainsi, dans le cas où l'espace d'état est très grand, cet algorithme devient virtuellement impossible à mettre en place à cause de la mémoire requise, c'est pourquoi il nous faut une méthode pour simplifier cet algorithme et c'est ici que le processus de "Sandwiching" apparaît pour nous permettre d'employer l'algorithme sur des espaces très grands, sous réserve que l'on vérifie certaines propriétés.

Avant toute chose, il faut savoir que ce procédé ne s'applique pas dans certains cas. En effet, il faut que la chaîne soit monotone par rapport à une relation d'ordre, qui n'a pas nécessairement besoin d'être totale, mais qui doit tout de même définir un unique maximum et un unique minimum. Le problème pour l'utilisation du sandwiching revient donc à la possibilité de trouver une telle relation d'ordre.

Si une telle propriété est vérifiée, il devient alors possible de ne prendre en compte que les états maximum et le minimum dans l'algorithme de Propp-Wilson : en effet, si nous prenons deux états distincts, par exemple si l'on prend $s_i, s_j \in S, s_i \preceq s_j$, alors $\forall x \in [0, 1], f(s_i, x) \leq f(s_j, x)$. Ainsi, toutes les chaînes simulées restent comprises entre celles simulées depuis les états maximal et minimal, et donc quand ces deux chaînes se "rejoignent", toutes les autres les rejoignent et l'algorithme se termine à la fin de la boucle en cours.

3.7 Recuit simulé (Simulated annealing)

Soit $S = \{s_1, s_2, \dots, s_k\}$, soit $f : S \rightarrow \mathbb{R}$ une fonction. L'objectif ici est de trouver le minimum de f .

Lorsque la taille de S est petite, il n'y a pas de problème, il suffit d'évaluer f en s_j pour tout j dans $\{1, \dots, k\}$.

Cependant, lorsque S est un ensemble de grande taille, le temps de calcul et la mémoire prise par les $f(s_j)$ deviennent un problème. C'est ici que la méthode du Recuit Simulé s'avère d'une grande utilité.

L'idée est la suivante, on suppose que l'on a une chaîne de Markov sur S dont l'unique loi stationnaire va "donner plus de probabilité" aux états s_j pour lesquels l'image par f est petite.

En parcourant suffisamment longtemps la chaîne, on finira avec une grande probabilité dans un de ces états.

Supposons maintenant que l'on change de chaîne pour les temps suivants, en prenant une chaîne dont l'unique loi stationnaire "donne encore plus de probabilité" aux états dont l'image par f est petite. Ainsi, après un certain temps de parcours de la chaîne, on se retrouve probablement dans un état dont l'image est plus petite qu'à l'étape précédente. On réitère alors cette étape en donnant toujours plus de probabilité aux états qui minimisent f .

Il semble naturel qu'au bout d'un long moment, avec ce schéma, on finisse dans l'état qui minimise f .

Posons $P^{(j)}$ la matrice de transition de la chaîne de Markov de la j -ème étape et N_j le temps de parcours de cette chaîne.

On peut alors représenter l'ensemble des étapes comme une unique chaîne de Markov non homogène dont la matrice de transition P est donnée par :

$$P = \begin{cases} P^{(1)} & \text{si } n=1, \dots, N_1 \\ P^{(2)} & \text{si } n=N_1 + 1, \dots, N_1 + N_2 \\ \cdot \\ \cdot \\ \cdot \\ P^{(j)} & \text{si } n= \sum_{i=1}^{j-1} N_i + 1, \dots, \sum_{i=1}^j N_i \\ \cdot \\ \cdot \end{cases}$$

Cependant, trouver des lois de distribution favorisant les états qui minimisent f peut poser problème.

Il existe une distribution connue pouvant être utilisée pour cela : la distribution de Boltzmann définie comme suit.

La distribution de Boltzmann $\pi_{f,T}$ sur un ensemble fini S , avec $f : S \rightarrow \mathbb{R}$ sa fonction d'énergie et T le paramètre de température, strictement positif, est la probabilité sur S qui à chaque élément s de S associe :

$$\pi_{f,T}(s) = \frac{1}{Z_{f,T}} \exp\left(\frac{-f(s)}{T}\right)$$

où $Z_{f,T}$ est la constante de normalisation donnée par $Z_{f,T} = \sum_{s \in S} \exp\left(\frac{-f(s)}{T}\right)$. Une chaîne de Markov ayant pour unique loi stationnaire la distribution de Boltzmann peut être construite par des méthodes MCMC comme vu précédemment.

Théorème : Soit S un ensemble fini et soit $f : S \rightarrow \mathbb{R}$. Pour tout réel strictement positif T , soit $\alpha : \mathbb{R} \rightarrow \mathbb{R}$ de sorte que $\alpha(T)$ représente la probabilité qu'un élément Y choisi selon la distribution de Boltzmann $\pi_{f,T}$ sur S satisfait $f(Y) = \min_{s \in S} f(s)$ alors $\lim_{T \rightarrow 0} \alpha(T) = 1$.

Démonstration On se place dans $S = \{s_1, \dots, s_k\}$. Soit s l'élément de S qui minimise f . On pose $a = f(s)$ et on pose b l'image de l'élément qui minimise f dans $S \setminus \{s\}$. On a $a < b$, d'où

$$\lim_{T \rightarrow 0} \exp\left(\frac{a-b}{T}\right) = 0$$

puisque $a-b < 0$. On a alors :

$$\begin{aligned}
\pi_{f,T}(s) &= \frac{1}{Z_{f,T}} \exp\left(\frac{-a}{T}\right) = \frac{\exp\left(\frac{-a}{T}\right)}{\sum_{s' \in S} \exp\left(\frac{-f(s')}{T}\right)} \\
&= \frac{\exp\left(\frac{-a}{T}\right)}{\sum_{s' \in S \setminus \{s\}} \exp\left(\frac{-f(s')}{T}\right) + \exp\left(\frac{-a}{T}\right)} \\
&\geq \frac{\exp\left(\frac{-a}{T}\right)}{(k-1)\exp\left(\frac{-b}{T}\right) + \exp\left(\frac{-a}{T}\right)} \\
&= \frac{1}{1 + \exp\left(\frac{a-b}{T}\right)}
\end{aligned}$$

Ainsi, lorsque T va tendre vers 0, on va obtenir :

$$\lim_{T \rightarrow 0} \pi_{f,T}(s) = 1$$

ce que l'on cherchait à montrer.

Le modèle du recuit simulé permettant de trouver l'élément s de S qui minimise f peut être exécuté ainsi :

1. Construire une chaîne de Markov suivant le modèle MCMC, avec la distribution de Boltzmann $\pi_{f,T}$ sur S
2. Fixer une suite décroissante de réels strictement positifs $T_1 > T_2 > T_3 > \dots$, qui seront les paramètres de température et tels que $\lim_{i \rightarrow +\infty} T_i = 0$. Fixer également une suite de nombres entiers positifs N_1, N_2, \dots qui seront les temps de parcours
3. Partant d'un état initial aléatoire dans S , parcourir pendant un N_1 la chaîne à la température T_1 , puis pendant un temps N_2 on parcourt la chaîne à température T_2 , et ainsi de suite..

Il est important de ne pas faire décroître la température trop rapidement, auquel cas on risque de finir dans un minimum local qui ne sera pas le minimum global.

4 Les modèles ou problèmes

4.1 Le modèle des sphères dures à deux dimensions

Le modèle des sphères dures a été introduit en physique statistique pour comprendre le comportement des particules de gaz dont le rayon est non-négligeable et qui ne peuvent se chevaucher.

Les valeurs 1, dans ce qui suit, représenteront la présence d'une particule en un emplacement tandis que la valeur 0 signifiera que l'emplacement est vide.

Soit $G = (V, E)$ un graphe, avec $V = \{v_1, \dots, v_k\}$ l'ensemble des sommets et $E = \{e_1, \dots, e_l\}$ l'ensemble des arêtes.

On va assigner de manière aléatoire la valeur 0 ou 1 à chaque sommet de sorte que deux sommets voisins n'aient pas tous deux la valeur 1. Une assignation de 0 et de 1 aux sommets de G sera appelée une *configuration* et on désignera par $\{0, 1\}^V$ l'ensemble des configurations.

Une configuration vérifiant "si un sommet prend la valeur 1, ses voisins doivent avoir la valeur 0" est dite *réalisable*. La manière de choisir une configuration aléatoire sera de prendre toutes les configurations réalisables avec équi-probabilité.

On note μ_G la mesure de probabilité sur $\{0, 1\}^V$ issue du tel modèle, on a donc, pour $\xi \in \{0, 1\}^V$,

$$\mu_G(\xi) = \begin{cases} \frac{1}{Z_G} & \text{si } \xi \text{ est une configuration réalisable} \\ 0 & \text{sinon} \end{cases}$$

où Z_G est le nombre de configurations réalisables sur G .

Une des questions les plus légitimes que l'on puisse se poser est : "Quel est le nombre de 1 attendus d'une configuration aléatoire ?"

4.1.1 Mise en œuvre sous Scilab

Pour répondre à la question posée ci-dessus, nous allons employer une des méthodes citées plus haut qui est facilement mise en œuvre pour ce problème : L'échantillonneur de Gibbs. L'intérêt sera d'estimer grâce au théorème ergodique le nombre moyen de "1" sur une grille de dimension donnée.

Montrons dans un premier temps que nous pouvons employer cette méthode dans ce cas : La première chose à faire est de construire la fonction de mise à jour et de vérifier que la chaîne qui en est issue est bien irréductible et apériodique. La fonction de mise à jour sera ici :

$$f(M, p, x, y) = \begin{cases} P & \text{si } p < \frac{1}{2} \\ Q & \text{si } p > \frac{1}{2} \end{cases}$$

Avec $\forall (i, j) \neq (x, y), Q(i, j) = M(i, j) = P(i, j)$, $P(x, y) = 0$ et $Q(x, y) = 1$ si les voisins de $Q(x, y)$ sont tous égaux à 0, sinon $Q(x, y) = M(x, y)$. On posera donc $X_{n+1} = f(X_n, U_{n+1})$ tel que $U_{n+1} = (p_{n+1}, x_{n+1}, y_{n+1})$ avec $(p_i)_{i \in \mathbb{N}} \sim U[0, 1]$ et $(x_i)_{i \in \mathbb{N}} \sim U\{1, \dots, n\}$ et $(y_i)_{i \in \mathbb{N}} \sim U\{1, \dots, m\}$ si la grille est de dimension $m \times n$.

La chaîne est donc bien apériodique, du fait qu'un état puisse donner accès à lui-même. De plus, la chaîne est clairement irréductible. L'échantillonneur de Gibbs peut donc bien être employé ici.

L'algorithme va se composer de deux fonctions : La mise à jour, qui va mettre à jour notre chaîne de Markov, puis l'algorithme de mise en œuvre qui va demander en entrée les dimensions de notre grille ainsi que le nombre d'itérations que l'on souhaite.

```

function[M]=miseajour(M,p,x,y)
    if p<1/2 then
        M(x,y)=0;
    else
        if and([M(x+1,y)==0,M(x-1,y)==0,M(x,y+1)==0,M(x,y-1)==0]) then
            M(x,y)=1;
        end
    end
end

function[M]=spheresdures(nblig,nbcol,nbiter)
    M=zeros(nblig+2,nbcol+2)
    for i=1 :nbiter
        x=floor(nblig*rand()+2)
        y=floor(nbcol*rand()+2)
        p=rand()
        M=miseajour(M,p,x,y)
    end
    M=M(2 :nblig+1,2 :nbcol+1)
end

```

Remarque : On utilisera une grille de taille $(n+2) \times (m+2)$ pour avoir une grille de taille $n \times m$ pour prendre une "ceinture de zéros" et éviter les problèmes de bords dans les calculs, puis on prendra uniquement la grille intérieure (i.e. sans cette ceinture).

4.2 Modèle d'Ising

Le modèle d'Ising est un modèle de physique statistique utilisé permettant de représenter des phénomènes dans lesquels un ensemble d'éléments, possédant une valeur parmi deux possibles, ont un effet collectif suivant la disposition des valeurs.

On se place encore une fois dans le cadre d'un graphe $G = (V, E)$. Dans ce graphe, chaque sommet se voit attribué d'une valeur -1 ou 1 que nous appellerons des *spins*. Ces spins représentent en fait l'orientation des particules d'un matériau ferromagnétique. Le modèle d'Ising est une manière de tirer de façon aléatoire $\xi \in \{-1, 1\}^V$, c'est-à-dire une disposition particulière des spins sur le graphe, dans l'ensemble des dispositions possibles.

Dans ce modèle, deux éléments proviennent justement de cette interprétation physique. Le premier élément est l'inverse de la température, que l'on notera β et qui sera considérée comme étant un nombre positif fixé. Le second élément est une fonction représentant l'*énergie* dégagée par la configuration prise par les spins, on notera H la fonction énergie définie par :

$$H(\xi) = - \sum_{\langle x,y \rangle \in E} \xi(x)\xi(y)$$

où $\langle x, y \rangle \in E$ signifie que le segment $[x,y]$ est une arête du graphe.

Chaque arête ajoute 1 à la valeur de l'énergie si les points d'extrémité sont de signes opposés et soustrait 1 dans le cas contraire. Ainsi, l'énergie sera plus faible lorsque dans la configuration comporte beaucoup de spins ayant des voisins de même signe.

Le modèle d'Ising signifie donc tirer de manière aléatoire une configuration X des spins dans $\{-1, 1\}^V$ suivant la mesure de probabilité $\pi_{G,\beta}$ définie par :

$$\forall \xi \in \{-1, 1\}^V, \pi_{G,\beta}(\xi) = \frac{1}{Z_{G,\beta}} \exp(\beta \sum_{\langle x,y \rangle \in E} \xi(x)\xi(y))$$

avec $Z_{G,\beta}$ la constante de normalisation donnée par $Z_{G,\beta} = \sum_{\nu \in \{-1,1\}^V} \exp(-\beta H(\nu))$.

Suivant le paramètre β , la distribution $\pi_{G,\beta}$ ne donnera pas la même probabilité aux configurations de $\{-1, 1\}^V$. En effet, dans le cas où $\beta = 0$, la distribution $\pi_{G,\beta}$ serait alors la loi uniforme sur $\{-1, 1\}^V$. A l'inverse, si β tend vers l'infini, la distribution va favoriser les deux configurations extrémales opposées où tous les sommets ont la même valeur.

4.2.1 Mise en œuvre sous Scilab

Les intérêts de la modélisation de ce modèle sont multiples et surtout intéressantes dans le domaine de la physique car il permet la représentation du comportement d'un liquide ferromagnétique.

En effet, ce modèle va mettre en évidence la transition de phase, ce qui signifie que selon le paramètre β , la grille ne se comportera pas de la même façon. Par exemple, sur une grille de dimension $n \times n$, on peut mettre en évidence l'existence d'un seuil critique β_c tel que si $\beta < \beta_c$, plus la grille deviendra grande et plus les "+1" et les "-1" vont s'équilibrer, et au contraire si $\beta > \beta_c$, alors plus la grille deviendra grande et plus la proportion d'un des deux spins sera grande.

Nous allons ici utiliser l'algorithme de Propp-Wilson avec le principe du sandwiching afin de réduire considérablement le temps d'exécution des calculs. Avant toute chose, définissons et justifions l'utilisation de cette méthode :

Premièrement, il nous faut une relation d'ordre. Prenons par exemple la relation \preceq ainsi définie : $\forall \alpha, \gamma \in S, \alpha \preceq \gamma \iff \forall v \in V, \alpha(v) \leq \gamma(v)$. On remarque ainsi l'existence d'une configuration minimale, où il n'y a que des "-1", et d'une maximale, celle où il n'y a que des "+1" : nous ne prendrons donc en compte que deux chaînes partant de ces états.

Ensuite, définissons une fonction de mise à jour qui va nous donner une chaîne apériodique, irréductible et qui respecte la relation d'ordre quand nous passons d'une étape à la suivante. Pour cela, nous pouvons implémenter un échantillonneur de Gibbs afin de mettre à jour la chaîne au sein de l'algorithme de Propp-Wilson. La fonction de mise à jour sera donc :

$$f_\beta(A, p, x, y) = \begin{cases} M & \text{si } p < \frac{\exp(2\beta(k_+(A,x,y) - k_-(A,x,y)))}{\exp(2\beta(k_+(A,x,y) - k_-(A,x,y))) + 1} \\ Q & \text{sinon} \end{cases}$$

Avec ici $X_{n+1} = f_\beta(X_n, U_{n+1})$ tel que :

- - $U_{n+1} = (p_{n+1}, x_{n+1}, y_{n+1})$ avec $(p_i)_{i \in \mathbb{N}} \sim U[0, 1]$ et $(x_i)_{i \in \mathbb{N}} \sim U\{1, \dots, n\}$ et $(y_i)_{i \in \mathbb{N}} \sim U\{1, \dots, m\}$ si la grille est de dimension $m \times n$.
- - $k_+(A, x, y)$ le nombre de "+1" voisin(s) de A(x,y)
- - $k_-(A, x, y)$ le nombre de "-1" voisin(s) de A(x,y)

- - $\forall(i, j) \neq (x, y), M(i, j) = Q(i, j) = A(i, j)$
- - $M(x, y) = +1$
- - $Q(x, y) = -1$

On remarque qu'un élément donne accès à lui même donc que l'on a l'apériodicité, de même que tous les états communiquent d'où l'irréductibilité. Il reste donc à montrer que la fonction de mise à jour respecte bien la relation d'ordre.

Prenons donc A et B deux états ($A, B \in S$) tels que $A \preceq B$. Notons $A_2 = f_\beta(A, p, x, y)$ et $B_2 = f_\beta(B, p, x, y)$ avec p, x et y fixés. Alors :

- - $\forall(i, j) \neq (x, y)$, les valeurs ne changent pas donc $A(i, j) \leq B(i, j)$;
- - Comme $A \preceq B \iff \forall v \in V, A(v) \leq B(v)$, on a $k_+(A, x, y) \leq k_+(B, x, y)$ et $k_-(A, x, y) \geq k_-(B, x, y)$. Ainsi, $\frac{\exp(2\beta(k_+(A, x, y) - k_-(A, x, y)))}{\exp(2\beta(k_+(A, x, y) - k_-(A, x, y))) + 1} \leq \frac{\exp(2\beta(k_+(B, x, y) - k_-(B, x, y)))}{\exp(2\beta(k_+(B, x, y) - k_-(B, x, y))) + 1}$ du fait que la fonction $z \mapsto \frac{z}{z+1}$ soit strictement croissante pour $z \geq 0$.

Nous pouvons en conclure que si $A \preceq B$, alors $A_2 \preceq B_2$ et la fonction conserve bien la relation d'ordre : nous pouvons appliquer le sandwiching.

```
function [M]=Plus(n)
    M=ones(n,n);
    V=zeros(1,n+2);
    W=zeros(n,1);
    M=[W,M,W];
    M=[V;M;V];
endfunction
```

```
function [N]=Minus(n)
    N=-Plus(n);
endfunction
```

```
function[a]=kplus(M,x,y)
    a=0
    if M(x+1,y)==1 then
        a=a+1
    end
    if M(x-1,y)==1 then
        a=a+1;
    end
    if M(x,y+1)==1 then
        a=a+1;
    end
    if M(x,y-1)==1 then
        a=a+1
    end
endfunction
```

```
function[a]=kmoins(M,x,y)
    a=0
```



```

    if M(x+1,y)==-1 then
        a=a+1 ;
    end
    if M(x-1,y)==-1 then
        a=a+1 ;
    end
    if M(x,y+1)==-1 then
        a=a+1 ;
    end
    if M(x,y-1)==-1 then
        a=a+1 ;
    end
endfunction

function[M]=Update(A,Beta,p,x,y)
    a=kplus(A,x,y) ;
    b=kmoins(A,x,y) ;
    cond=exp(2*Beta*(a-b))/(exp(2*Beta*(a-b))+1) ;
    if p<cond then
        A(x,y)=1 ;
    else
        A(x,y)=-1 ;
    end
    M=A ;
endfunction

function[M]=SandwichIsing(n,Beta)
    i=0 ;
    u=[rand() ;floor(n*rand()+2 ;floor(n*rand()+2)] ;
    P=Update(Plus(n),Beta,u(1,1),u(2,1),u(3,1)) ;
    Q=Update(Minus(n),Beta,u(1,1),u(2,1),u(3,1)) ;
    while isequal(P,Q)==%F
        i=i+1 ;
        for j=2**(i-1)+1 :2**i
            u=[[rand() ;floor(n*rand()+2 ;floor(n*rand()+2],u] ;
        end
        P=Plus(n) ;
        Q=Minus(n) ;
        for k=1 :2**i
            P=Update(P,Beta,u(1,k),u(2,k),u(3,k)) ;
            Q=Update(Q,Beta,u(1,k),u(2,k),u(3,k)) ;
        end
    end
    M=Q(2 :n+1,2 :n+1)
endfunction

```

Remarque : Encore une fois, nous utilisons l'astuce de la "ceinture de zéros" afin d'éviter les problèmes de bords.

4.3 Résultats obtenus

4.3.1 Résultats pour le modèle des sphères dures

Il suffira ici de regarder en moyenne le nombre de 1 qui vont apparaître sur la grille, c'est à dire employer le théorème ergodique avec la fonction suivante :

$$f : \mathcal{M}_{n \times m}(\mathbb{R}) \rightarrow \mathbb{N}$$
$$M \mapsto \sum_{1 \leq i \leq n} \sum_{1 \leq j \leq m} \mathbb{1}_{M_{i,j}=1}$$

Cette fonction est bornée sur un espace d'état fini donc f est pi - intégrable, de plus la chaîne est irréductible sur un espace d'états fini donc récurrente positive : le théorème ergodique peut donc s'appliquer ici :

Pour ce faire on peut utiliser un programme simple du type :

```
function[k]=NbSpheres(nblig,nbcol)
    k=0
    for i=1 :5000
        k = k + sum(spheresdures(nblig,nbcol,500));
    end
    k=1/5000*k
end
```

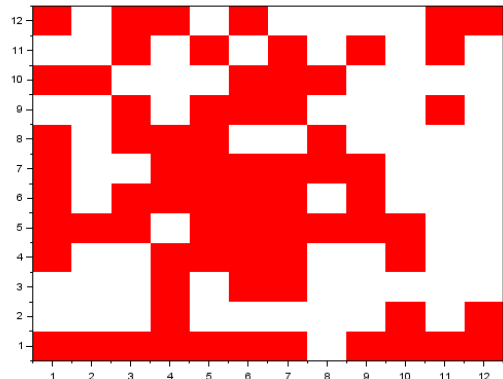
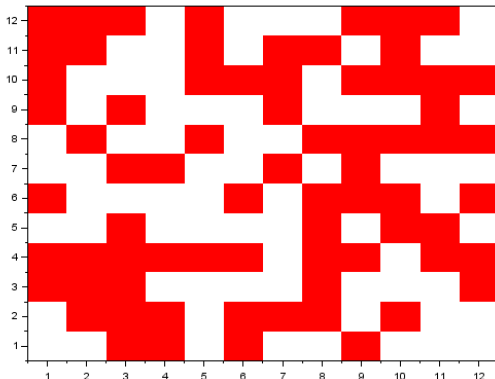
En s'intéressant uniquement dans un premier temps aux grilles carrées de dimension $n \times n$, nous avons ici obtenu qu'en moyenne nous avons $\frac{n^2}{4}$ sphères dures.

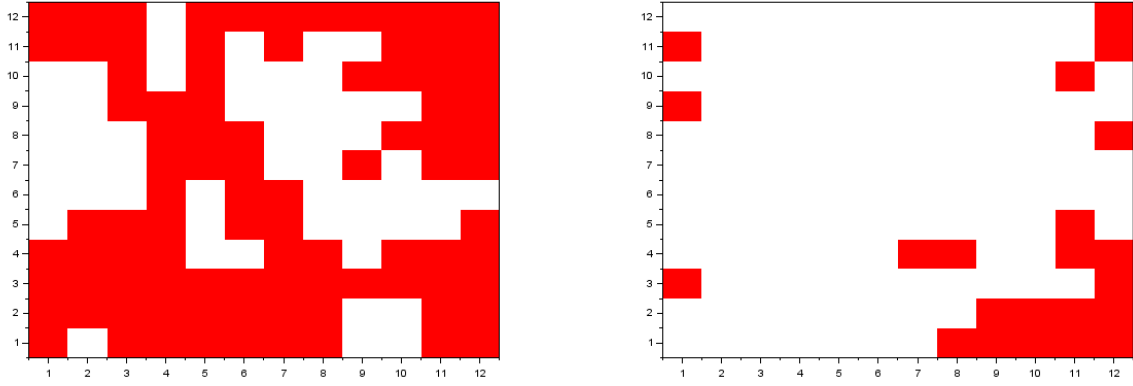
Nous avons également remarqué que pour des grilles non carrées de dimension $n \times m$, nous avons un résultat analogue car nous obtenons en moyenne $\frac{n \times m}{4}$.

Ce résultat semble intuitif, puisque la grille comprendra toujours au maximum $\frac{n \times m}{2}$ sphères et au minimum 0, et donc en moyenne $\frac{n \times m}{4}$ sphères.

4.3.2 Résultats pour le modèle d'Ising

Nous avons pu faire des simulations pour des grilles de dimension 12x12 au maximum afin de pouvoir avoir des valeurs de β entre 0 et 0.5.





Ces quatre grilles représentent une réalisation de l'algorithme de Propp-Wilson respectivement pour $\beta = 0$, $\beta = 0.1$, $\beta = 0.3$, $\beta = 0.5$ (Les cases rouges représentant un spin "+1" et les blanches un spin "-1"). On peut remarquer que pour les deux premières, la proportion entre les deux spins est sensiblement la même, cependant pour $\beta = 0.3$ et $\beta = 0.5$, on remarque que l'un des deux spins prend une proportion supérieure à l'autre : dans le cas $\beta = 0.3$, ce sont les spins "+1" et dans le cas $\beta = 0.5$, ce sont les spins "-1". Cela permet aussi de constater que les deux spins peuvent prendre l'ascendant sur l'autre par symétrie.

Nous avons ainsi constaté que pour β aux alentours de 0.4, nous avons l'un des deux spins qui prend l'ascendant par rapport à l'autre quant à la proportion de présence sur la grille.

5 Conclusion

Les chaînes de Markov se sont avérées être un outil très puissant et polyvalent pour la simulation de lois. En effet, les nombreux théorèmes tels que la convergence en loi ou le théorème ergodique sont tous deux très polyvalents dans leurs utilisations, et très faciles à utiliser en pratique pour obtenir des résultats. Le réel travail réside en fait dans la construction de chaînes de Markov suivant des hypothèses précises afin que nous puissions utiliser ces théorèmes.

L'intérêt ici est que nous pouvons obtenir des méthodes assez simples et intuitives telles que les chaînes de Monte-Carlo, qui cependant sont assez difficiles à mettre en œuvre dès que l'espace d'états est trop grand, mais également certaines méthodes peu intuitives dont la mise en œuvre est beaucoup plus facile. Ces simulations peuvent ensuite être manipulées afin de répondre aux questions que nous pouvons nous poser, par exemple "À partir de quelle valeur de β l'un des deux spins prend plus de proportion qu'un autre de manière significative dans le modèle d'Ising?".

En conclusion, ces méthodes sont très utiles dans le cadre d'études de modèles apportés en physique pour étudier le comportement complexe de certains phénomènes tels que ceux présentés ici, et pour pouvoir en extraire des données par l'utilisation de théorèmes relatifs aux chaînes de Markov.

Références

- [1] OLLE HÄGGSTRÖM, Finite Markov Chains and Algorithmic Applications
- [2] Wikipedia, page "Gibbs sampling",
https://en.wikipedia.org/wiki/Gibbs_sampling
- [3] <http://www.cs.tau.ac.il/~amnon/Classes/2010-Seminar-Random-Walk/Presentations/Propp-Wilson.pdf>