

Correspondance de Markov–Gibbs et simulation de champs aléatoires

On appelle *champ* une famille de variables aléatoires indicées par les sommets d'un graphe. L'objectif du devoir est d'étudier un certain type de champ aléatoire : ces champs sont caractérisés par une propriété qui généralise la définition des chaînes de Markov. On montrera que les lois de ces champs sont toujours données par des expressions semblables à celle définissant le champ du modèle d'Ising (mesure de Gibbs). Par ailleurs, il est toujours possible de simuler de tels champs en utilisant des chaînes de Markov, et la dernière partie du devoir est consacrée à ces techniques de simulation.

1 Indépendance conditionnelle et propriété de Markov

Soit X, Y, Z trois variables aléatoires à valeurs dans des ensembles dénombrables $\mathfrak{X}, \mathfrak{Y}, \mathfrak{Z}$. On dira que X et Y sont *indépendants conditionnellement à Z* si, pour toute valeur $z \in \mathfrak{Z}$ telle que $\mathbb{P}[Z = z] \neq 0$ et tous $x \in \mathfrak{X}$ et $y \in \mathfrak{Y}$, on a

$$\mathbb{P}[X = x \text{ et } Y = y \mid Z = z] = \mathbb{P}[X = x \mid Z = z] \times \mathbb{P}[Y = y \mid Z = z].$$

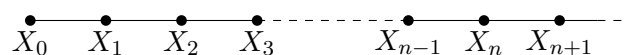
1. Montrer que X et Y sont indépendants conditionnellement à Z si et seulement si, pour tous $y \in \mathfrak{Y}$ et $z \in \mathfrak{Z}$ tels que $\mathbb{P}[Y = y \text{ et } Z = z] \neq 0$, et pour tout $x \in \mathfrak{X}$,

$$\mathbb{P}[X = x \mid Y = y \text{ et } Z = z] = \mathbb{P}[X = x \mid Z = z].$$

2. Soit \mathfrak{X} un ensemble dénombrable et $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov de matrice P sur \mathfrak{X} . Montrer que pour tous entiers $N > n \geq 1$, les variables $X_n \in \mathfrak{X}$ et $(X_0, X_1, \dots, X_{n-2}, X_{n+2}, \dots, X_N) \in \mathfrak{X}^{N-2}$ sont indépendantes conditionnellement à $(X_{n-1}, X_{n+1}) \in \mathfrak{X}^2$. Donner la valeur de la loi conditionnelle

$$\mathbb{P}[X_n = x \mid X_{n-1} = a, X_{n+1} = b].$$

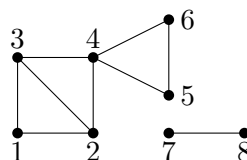
Ainsi, si l'on représente le temps discret par une droite, alors chaque variable X_n est indépendante des autres variables conditionnellement aux valeurs prises par les deux variables *voisines* (X_{n-1}, X_{n+1}) .



La section suivante étend cette définition à des familles de variables indicées par les sommets de graphes plus généraux.

2 Champs de Markov et champs de Gibbs

Dans tout ce qui suit, $G = (V, E)$ est un graphe simple fini : V est un ensemble fini de sommets, et E est une partie de l'ensemble des paires $\{a, b\}$ de sommets distincts ($a \neq b$). On a dessiné ci-dessous un exemple avec 8 sommets et 9 arêtes. Les *voisins* d'un sommet v sont les sommets w tels que $\{v, w\}$ appartient à E . On notera N_v l'ensemble des voisins d'un sommet v ; dans l'exemple ci-dessous, $N_1 = \{2, 3\}$ et $N_4 = \{2, 3, 5, 6\}$. On notera aussi $\bar{N}_v = \{v\} \sqcup N_v$ et $M_v = V \setminus \bar{N}_v$.



Soit \mathfrak{X} un ensemble dénombrable. Un *champ* sur le graphe G et à valeurs dans \mathfrak{X} est un ensemble de variables aléatoires $(X_v)_{v \in V}$ à valeurs dans \mathfrak{X} . Sa loi est la fonction

$$\mathbb{P}[(x_v)_{v \in V}] = \mathbb{P}[\forall v \in V, X_v = x_v]$$

sur \mathfrak{X}^V . On dit que le champ $(X_v)_{v \in V}$ a la *propriété de Markov* si, pour tout sommet $v \in V$, les variables X_v et $(X_w)_{w \in M_v}$ sont indépendantes conditionnellement à $(X_w)_{w \in N_v}$. Ainsi, conditionnellement aux variables voisines, toute variable X_v est indépendante des autres variables X_w . La *loi locale* d'un champ de Markov est la fonction

$$\begin{aligned} \Pi : \prod_{v \in V} \mathfrak{X}^{\overline{N}_v} &\rightarrow [0, 1] \\ (x_v, (x_w)_{w \in N_v}) &\mapsto \mathbb{P}[X_v = x_v \mid \forall w \in N_v, X_w = x_w]. \end{aligned}$$

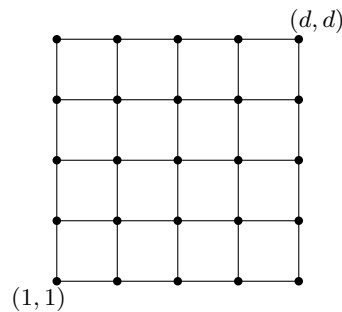
Pour que la loi locale soit bien définie en $(x_v, (x_w)_{w \in N_v})$, il faut que $\mathbb{P}[\forall w \in N_v, X_w = x_w]$ soit non nulle. Dans ce qui suit, pour s'assurer de cela, on supposera toujours que le champ $(X_v)_{v \in V}$ est *positif* : pour toute famille $(x_v)_{v \in V} \in \mathfrak{X}^V$, $\mathbb{P}[(x_v)_{v \in V}] > 0$. Autrement dit, toutes les valeurs de champs possibles ont une probabilité strictement positive. Alors, la loi locale Π est une fonction bien définie et strictement positive sur tout $\prod_{v \in V} \mathfrak{X}^{\overline{N}_v}$.

3. Soit $(X_v)_{v \in V}$ un champ de Markov positif de graphe $G = (V, E)$ et d'espace de valeurs \mathfrak{X} . Quitte à énumérer les sommets, on peut supposer que $V = [1, N] = \{1, 2, \dots, N\}$. Montrer que pour toutes familles $(x_v)_{v \in V}$ et $(y_v)_{v \in V}$ dans \mathfrak{X}^V , on a

$$\mathbb{P}[(x_v)_{v \in V}] = \left(\prod_{i=1}^N \frac{\mathbb{P}[X_i = x_i \mid X_1 = x_1, \dots, X_{i-1} = x_{i-1}, X_{i+1} = y_{i+1}, \dots, X_N = y_N]}{\mathbb{P}[X_i = y_i \mid X_1 = x_1, \dots, X_{i-1} = x_{i-1}, X_{i+1} = y_{i+1}, \dots, X_N = y_N]} \right) \mathbb{P}[(y_v)_{v \in V}].$$

En déduire que si la loi locale $\Pi : \prod_{v \in V} \mathfrak{X}^{\overline{N}_v} \rightarrow (0, 1]$ est fixée, alors la loi globale $\mathbb{P} : \mathfrak{X}^V \rightarrow (0, 1]$ du champ $(X_v)_{v \in V}$ est uniquement déterminée.

Parmi les champs de Markov, une sous-famille intéressante est formée par les *champs de Gibbs* associés à des potentiels. Dans un graphe $G = (V, E)$, une *clique* $C \subset V$ est un ensemble de sommets tous voisins les uns des autres : si $v, w \in C$ et $v \neq w$, alors $\{v, w\} \in E$. On note $\mathfrak{C}(G)$ l'ensemble des cliques de G ; c'est une partie de l'ensemble $\mathfrak{P}(V)$ des parties de V , qui contient en particulier tous les singletons-sommets $\{v\}$ et toutes les paires-arêtes $\{v, w\}$. Par exemple, si $G = [1, d]^2$ est la grille carrée



alors $\mathfrak{C}(G)$ est uniquement composé des sommets et des arêtes. Dans le graphe dessiné au début de la section, $\{2, 3, 4\}$ et $\{4, 5, 6\}$ sont aussi des cliques. Un *potentiel* sur le graphe G est une famille de fonctions $\mathcal{V} = (\mathcal{V}_C)_{C \in \mathfrak{C}(G)}$ indicées par les cliques de G , avec chaque fonction

$$\mathcal{V}_C : \mathfrak{X}^C \rightarrow \mathbb{R}.$$

La *mesure de Gibbs* associée à un potentiel \mathcal{V} est la mesure de probabilité sur l'ensemble des champs \mathfrak{X}^V définie par :

$$\mathbb{P}_{\mathcal{V}}[(x_v)_{v \in V}] = \frac{1}{Z_{\mathcal{V}}} \exp \left(- \sum_{C \in \mathfrak{C}(G)} \mathcal{V}_C((x_v)_{v \in C}) \right),$$

où $Z_{\mathcal{V}}$ est la constante de normalisation $\sum_{(x_v)_{v \in V} \in \mathfrak{X}^V} \exp \left(- \sum_{C \in \mathfrak{C}(G)} \mathcal{V}_C((x_v)_{v \in C}) \right)$.

4. Montrer qu'un champ $(X_v)_{v \in V}$ donné par une mesure de Gibbs de potentiel \mathcal{V} a la propriété de Markov. Donner dans ce cadre la valeur de la loi locale Π en fonction du potentiel \mathcal{V} . On pourra simplifier l'expression de $\mathbb{P}[X_v = x_v \mid (X_w)_{w \in M_v} = (x_w)_{w \in M_v}$ et $(X_w)_{w \in N_v} = (x_w)_{w \in N_v}]$ pour ne faire apparaître que des potentiels $\mathcal{V}_C((x_v)_{v \in C})$ avec C clique contenant v , et donc incluse dans le voisinage \bar{N}_v .

5. On suppose que G est la grille carrée $[1, d]^2$, que $\mathfrak{X} = \{-1, 1\}$ et que le potentiel \mathcal{V} est défini comme suit :

$$\begin{aligned} \mathcal{V}_{\{v\}}(x_v) &= -Hx_v; \\ \mathcal{V}_{\{v,w\}}(x_v, x_w) &= -Jx_vx_w \end{aligned}$$

avec H, J constantes réelles. Ainsi,

$$\mathbb{P}[(x_v)_{v \in V}] = \frac{1}{Z} \exp \left(H \sum_{v \in V} x_v + J \sum_{\{v,w\} \in E} x_vx_w \right).$$

C'est la *loi d'Ising* avec champ magnétique H et interaction J entre les voisins. Montrer que la loi locale s'écrit :

$$\Pi(x_v, (x_w)_{w \in N_v}) = \frac{\exp(x_v (H + J \sum_{w \in N_v} x_w))}{2 \cosh(H + J \sum_{w \in N_v} x_w)}.$$

3 Le théorème d'Hammersley-Clifford

L'objectif de cette section est de montrer que réciproquement, si $(X_v)_{v \in V}$ est un champ de Markov positif sur un graphe fini $G = (V, E)$, alors il existe un potentiel \mathcal{V} tel que $(X_v)_{v \in V}$ soit le champ de Gibbs associé à ce potentiel.

6. Soit Φ et Ψ deux fonctions de l'ensemble $\mathfrak{P}(V)$ des parties de V vers \mathbb{R} . Montrer que les deux assertions suivantes sont équivalentes :

— pour toute partie A ,

$$\Psi(A) = \sum_{B \mid B \subset A} \Phi(B).$$

— pour toute partie A ,

$$\Phi(A) = \sum_{B \mid B \subset A} (-1)^{|A \setminus B|} \Psi(B),$$

où $|A \setminus B| = |A| - |B|$ désigne le cardinal de $A \setminus B$. C'est la formule d'*inclusion-exclusion*.

7. On fixe un champ de Markov positif de loi globale \mathbb{P} , ainsi qu'un élément arbitraire 0 de l'espace des valeurs \mathfrak{X} . Si $(x_v)_{v \in V}$ est un champ fixé et B une partie de V , notons $(x_v^B)_{v \in V}$ le champ qui vaut $x_v^B = x_v$ pour $v \in B$, et qui vaut $x_v^B = 0$ pour $v \notin B$. Pour $A \subset V$, posons alors

$$\mathcal{V}_A((x_v)_{v \in A}) = \sum_{B \mid B \subset A} (-1)^{|A \setminus B|} \log \left(\frac{\mathbb{P}[(0)_{v \in V}]}{\mathbb{P}[(x_v^B)_{v \in V}]} \right).$$

Montrer que pour toute valeur de champ $(x_v)_{v \in V}$, on a

$$\mathbb{P}[(x_v)_{v \in V}] = \mathbb{P}[(0)_{v \in V}] \exp \left(- \sum_{A \subset V} \mathcal{V}_A((x_v)_{v \in A}) \right).$$

On se donne une partie $A \subset V$ contenant deux éléments $s \neq t$ qui ne sont pas voisins; autrement dit, A n'est pas une clique. En séparant les parties $B \subset A$ selon qu'elles contiennent t ou non, montrer que

$$\begin{aligned} \mathcal{V}_A((x_v)_{v \in A}) &= \sum_{B \mid B \subset A \setminus \{t\}} (-1)^{|A \setminus B|} \log \left(\frac{\mathbb{P}[(x_v^{B \sqcup \{t\}})_{v \in V}]}{\mathbb{P}[(x_v^B)_{v \in V}]} \right) \\ &= \sum_{B \mid B \subset A \setminus \{t\}} (-1)^{|A \setminus B|} \log \left(\frac{\Pi(x_v, (x_w^B)_{w \in N_t})}{\Pi(0, (x_w^B)_{w \in N_t})} \right), \end{aligned}$$

où Π est la loi locale du champ de Markov. En remarquant que les termes de la dernière somme ne dépendent que de variables associées à la partie N_t , et en séparant les parties $B \subset A \setminus \{t\}$ selon qu'elles contiennent s ou non, en déduire que $\mathcal{V}_A((x_v)_{v \in A}) = 0$ pour tout vecteur $(x_v)_{v \in A}$. Conclure que $\mathbb{P} = \mathbb{P}_{\mathcal{V}}$ est la mesure de Gibbs associée au potentiel \mathcal{V} .

4 Simulation des champs de Markov-Gibbs

Pour conclure notre étude des champs de Markov, fixons un champ positif sur un graphe fini $G = (V, E)$, avec espace de valeurs \mathfrak{X} . On note comme précédemment \mathbb{P} la loi globale et Π la loi locale. Soit $(\xi_n)_{n \in \mathbb{N}}$ une suite de sommets indépendants et de même loi, telle que $\mathbb{P}[\xi_0 = v] > 0$ pour tout sommet v . On définit une chaîne de Markov $X_n = (X_{n,v})_{v \in V}$ sur l'espace des champs \mathfrak{X}^V comme suit : si X_n est déjà construit, alors X_{n+1} est le champ avec les mêmes valeurs que X_n en tout sommet $v \neq \xi_n$, et avec une valeur aléatoire de loi

$$\Pi(\cdot, (X_{n,v})_{v \in N_{\xi_n}})$$

en le sommet ξ_n . Autrement dit, on modifie une seule valeur du champ à chaque étape, en utilisant la loi locale pour choisir cette valeur.

8. Montrer que la chaîne de Markov $(X_n)_{n \in \mathbb{N}}$ est irréductible sur l'espace \mathfrak{X}^V , et que la loi globale \mathbb{P} du champ positif est réversible pour cette chaîne. En déduire que la chaîne est récurrente positive. Montrer qu'elle est aussi apériodique, et expliquer comment utiliser la chaîne $(X_n)_{n \in \mathbb{N}}$ pour simuler un champ de loi \mathbb{P} . C'est la technique d'*échantillonnage de Gibbs des champs de Markov*.
9. On reprend l'exemple de la question 5., par exemple avec une taille de grille $d = 50$. Écrire un programme (dans n'importe quel langage, par exemple Python ou Python/Sage) qui prend pour paramètres H, J et n et qui renvoie le champ aléatoire X_n après n transitions.
 - On pourra éventuellement travailler sur une grille torique $((\mathbb{Z}/d\mathbb{Z})^2)$ au lieu de $[1, d]^2$ pour que tous les sommets ait le même nombre de voisins.
 - Si l'on le souhaite, on pourra également modifier la façon dont les sites à mettre à jour sont choisis : si au lieu d'en prendre un au hasard à chaque fois, on les prend tous les uns à la suite des autres périodiquement, cela ne modifie pas les propriétés attendues du champ X_n pour n grand, et a tendance à accélérer un peu la convergence.

Commenter les champs ainsi obtenus (on pourra les dessiner avec une couleur pour les sites de valeur $+1$ et une autre pour les sites de valeur -1). Quelle est l'influence de H et de J sur ces champs aléatoires? On appelle magnétisation d'un champ $(x_v)_{v \in G}$ la valeur moyenne $M = \frac{1}{|V|} \sum_{v \in V} x_v$. Vérifier par l'expérience le phénomène de *magnétisation spontanée* suivant : pour J assez grand,

$$\lim_{H \rightarrow 0_+, H > 0} \left(\lim_{d \rightarrow +\infty} \mathbb{E}_{H,J}[M] \right) > 0.$$

Corrigé

1. Si l'on développe deux des probabilités conditionnelles dans la définition d'indépendance conditionnelle, on obtient :

$$\frac{\mathbb{P}[X = x \text{ et } Y = y \text{ et } Z = z]}{\mathbb{P}[Z = z]} = \frac{\mathbb{P}[Y = y \text{ et } Z = z]}{\mathbb{P}[Z = z]} \mathbb{P}[X = x | Z = z]$$

pour toute valeur $z \in \mathfrak{Z}$ telle que $\mathbb{P}[Z = z] \neq 0$. Si $\mathbb{P}[Y = y \text{ et } Z = z] = 0$, alors les deux termes de l'équation valent 0, donc l'égalité est vraie. Sinon, on peut diviser les deux termes par $\mathbb{P}[Y = y | Z = z]$ pour obtenir l'équation équivalente

$$\frac{\mathbb{P}[X = x \text{ et } Y = y \text{ et } Z = z]}{\mathbb{P}[Y = y \text{ et } Z = z]} = \mathbb{P}[X = x | Z = z],$$

et le terme de gauche est $\mathbb{P}[X = x | Y = y \text{ et } Z = z]$.

2. Fixons des valeurs $y_0, y_1, \dots, y_{n-2}, a, x, b, y_{n+2}, \dots, y_N$ dans \mathfrak{X} . On note π_0 la mesure initiale de la chaîne $(X_n)_{n \in \mathbb{N}}$; $X = X_n$, $Y = (X_0, \dots, X_{n-2}, X_{n+2}, \dots, X_N)$ et $Z = (X_{n-1}, X_{n+1})$. La formule habituelle pour les probabilités trajectorielles d'une chaîne de Markov donne :

$$\begin{aligned} \mathbb{P}[Z = (a, b)] &= \sum_{x_0, \dots, x_{n-2}, x} \pi_0(x_0) P(x_0, x_1) \cdots P(x_{n-3}, x_{n-2}) P(x_{n-2}, a) P(a, x) P(x, b) \\ &= (\pi P^{n-1})(a) P^2(a, b). \end{aligned}$$

De même,

$$\begin{aligned} \mathbb{P}[X = x \text{ et } Z = (a, b)] &= \sum_{x_0, \dots, x_{n-2}} \pi_0(x_0) P(x_0, x_1) \cdots P(x_{n-3}, x_{n-2}) P(x_{n-2}, a) P(a, x) P(x, b) \\ &= (\pi P^{n-1})(a) P(a, x) P(x, b), \end{aligned}$$

donc $\mathbb{P}[X = x | Z = z] = \mathbb{P}[X_n = x | X_{n-1} = a \text{ et } X_{n+1} = b] = \frac{P(a, x) P(x, b)}{P^2(a, b)}$. Finalement, pour l'indépendance conditionnelle, on a

$$\begin{aligned} \mathbb{P}[X = x \text{ et } Y = (y_0, \dots, y_{n-2}, y_{n+2}, \dots, y_N) \text{ et } Z = (a, b)] \\ = \pi_0(y_0) P(y_0, y_1) \cdots P(y_{n-3}, y_{n-2}) P(y_{n-2}, a) P(a, x) P(x, b) P(b, y_{n+2}) \cdots P(y_{N-1}, y_N) \end{aligned}$$

et

$$\begin{aligned} \mathbb{P}[Y = (y_0, \dots, y_{n-2}, y_{n+2}, \dots, y_N) \text{ et } Z = (a, b)] \\ = \sum_x \pi_0(y_0) P(y_0, y_1) \cdots P(y_{n-3}, y_{n-2}) P(y_{n-2}, a) P(a, x) P(x, b) P(b, y_{n+2}) \cdots P(y_{N-1}, y_N) \\ = \pi_0(y_0) P(y_0, y_1) \cdots P(y_{n-3}, y_{n-2}) P(y_{n-2}, a) P^2(a, b) P(b, y_{n+2}) \cdots P(y_{N-1}, y_N) \end{aligned}$$

donc le ratio vaut de nouveau $\frac{P(a, x) P(x, b)}{P^2(a, b)}$. Ainsi, $\mathbb{P}[X = x | Y = y \text{ et } Z = z] = \mathbb{P}[X = x | Z = z]$ si l'événement $\{Y = y \text{ et } Z = z\}$ a une probabilité non nulle.

3. Dans le produit, on peut réécrire chaque ratio sous la forme

$$\frac{\mathbb{P}[x_1, \dots, x_i, y_{i+1}, \dots, y_N]}{\mathbb{P}[x_1, \dots, x_{i-1}, y_i, \dots, y_N]},$$

la positivité du champ garantit qu'aucun dénominateur ne s'annule. Alors, le terme de droite de l'identité est un produit télescopique

$$\frac{p_1}{p_0} \times \frac{p_2}{p_1} \times \cdots \times \frac{p_N}{p_{N-1}} \times p_0 = p_N$$

avec $p_i = \mathbb{P}[x_1, \dots, x_i, y_{i+1}, \dots, y_N]$. On a donc bien établi la formule annoncée, et on remarque que par définition d'un champ de Markov,

$$\mathbb{P}[X_i = x_i | X_1 = x_1, \dots, X_{i-1} = x_{i-1}, X_{i+1} = y_{i+1}, \dots, X_N = y_N] = \Pi(x_i, (a_j)_{j \in N_i})$$

avec $a_j = x_j$ si $j < i$ et $a_j = y_j$ si $j > i$. De même,

$$\mathbb{P}[X_i = y_i | X_1 = x_1, \dots, X_{i-1} = x_{i-1}, X_{i+1} = y_{i+1}, \dots, X_N = y_N] = \Pi(y_i, (a_j)_{j \in N_i}).$$

Autrement dit, si la loi locale Π est fixée, alors la loi globale \mathbb{P} est entièrement déterminée (les rapports de probabilité sont tous déterminés, et la normalisation $\sum_{x \in \mathfrak{X}^V} \mathbb{P}[x] = 1$ fixe toutes les valeurs).

4. Fixons un sommet v et calculons les probabilités conditionnelles de la question 1. pour une mesure de Gibbs $\mathbb{P} = \mathbb{P}_V$. On fixe des valeurs $(x_w)_{w \in V}$ et on note $X = X_v$, $Y = (X_w)_{w \in M_v}$ et $Z = (X_w)_{w \in N_v}$. On a :

$$\begin{aligned} \mathbb{P}[X = x_v \text{ et } Y = (x_w)_{w \in M_v} \text{ et } Z = (x_w)_{w \in N_v}] &= \frac{1}{Z_V} \exp\left(-\sum_C \mathcal{V}_C((x_w)_{w \in C})\right) \\ \mathbb{P}[Y = (x_w)_{w \in M_v} \text{ et } Z = (x_w)_{w \in N_v}] &= \sum_{y \in \mathfrak{X}} \frac{1}{Z_V} \exp\left(-\sum_C \mathcal{V}_C((\tilde{x}_w(y))_{w \in C})\right) \end{aligned}$$

avec $\tilde{x}_w(y) = x_w$ si $w \neq v$, et $\tilde{x}_w(y) = y$ si $w = v$. Par conséquent,

$$\mathbb{P}[X = x_v | Y = (x_w)_{w \in M_v} \text{ et } Z = (x_w)_{w \in N_v}] = \frac{\exp(-\sum_C \mathcal{V}_C((x_w)_{w \in C}))}{\sum_{y \in \mathfrak{X}} \exp(-\sum_C \mathcal{V}_C((\tilde{x}_w(y))_{w \in C}))}.$$

Dans cette expression, les sommes \sum_C portent sur l'ensemble des cliques $\mathfrak{C}(G)$. Or, si v n'est pas dans la clique C , alors les vecteurs $(x_w)_{w \in C}$ et $(\tilde{x}_w(y))_{w \in C}$ sont identiques, donc les contributions correspondantes des potentiels se compensent. On peut donc ne considérer que les cliques contenant v :

$$\mathbb{P}[X = x_v | Y = (x_w)_{w \in M_v} \text{ et } Z = (x_w)_{w \in N_v}] = \frac{\exp\left(-\sum_{C|v \in C} \mathcal{V}_C((x_w)_{w \in C})\right)}{\sum_{y \in \mathfrak{X}} \exp\left(-\sum_{C|v \in C} \mathcal{V}_C((\tilde{x}_w(y))_{w \in C})\right)}.$$

Si $v \in C$, alors $C \subset \bar{N}_v$, donc la loi conditionnelle ne dépend que des valeurs x_v et $(x_w)_{w \in N_v}$. Ceci démontre que

$$\mathbb{P}[X = x_v | Y = (x_w)_{w \in M_v} \text{ et } Z = (x_w)_{w \in N_v}] = \mathbb{P}[X = x_v | Z = (x_w)_{w \in N_v}],$$

et un champ de Gibbs a donc la propriété de Markov. La valeur de la loi locale est

$$\Pi(x_v, (x_w)_{w \in N_v}) = \frac{\exp\left(-\sum_{C|v \in C} \mathcal{V}_C((x_w)_{w \in C})\right)}{\sum_{y \in \mathfrak{X}} \exp\left(-\sum_{C|v \in C} \mathcal{V}_C((\tilde{x}_w(y))_{w \in C})\right)}.$$

5. Les cliques contenant un sommet v de la grille sont $\{v\}$ et les arêtes $\{v, w\}$ avec w voisin de v . Le numérateur de l'expression de la loi locale est donc

$$\exp\left(Hx_v + J \sum_{w \in N_v} x_v x_w\right) = \exp\left(x_v \left(H + J \sum_{w \in N_v} x_w\right)\right).$$

Le dénominateur est la somme sur les deux valeurs possibles de $x_v \in \{-1, 1\}$ de cette expression, c'est-à-dire

$$\exp\left(+\left(H + J \sum_{w \in N_v} x_w\right)\right) + \exp\left(-\left(H + J \sum_{w \in N_v} x_w\right)\right) = 2 \cosh\left(H + J \sum_{w \in N_v} x_w\right).$$

Le résultat est donc une conséquence immédiate de la formule générale montrée à la question précédente.

6. Si l'on considère Ψ et Φ comme des vecteurs de $\mathbb{R}^{\mathfrak{P}(V)}$, alors les deux relations s'écrivent $\Psi = M\Phi$ et $\Phi = N\Psi$, où M et N sont les deux matrices

$$M_{A,B} = 1_{B \subset A} \quad ; \quad N_{A,B} = 1_{B \subset A} (-1)^{|A \setminus B|}.$$

Il s'agit donc de montrer que les deux matrices M et N sont inverses l'une de l'autre. Or,

$$(MN)_{A,C} = \sum_B M_{A,B} N_{B,C} = \sum_B 1_{C \subset B \subset A} (-1)^{|B \setminus C|}.$$

Si C n'est pas incluse dans A , alors la somme ci-dessus n'a que des termes nuls, donc vaut 0. Supposons maintenant $C \subset A$ et notons $c = |C|$ et $a = |A|$. Les parties B qui contribuent à la somme ont un cardinal b compris entre c et a , et pour tout $b \in [c, a]$, il y a $\binom{a-c}{b-c}$ parties B de cardinal b contribuant (il faut choisir $b-c$ éléments dans A et pas dans C pour compléter C en une partie B). Ainsi,

$$(MN)_{A,C} = \sum_{b=c}^a (-1)^{b-c} \binom{a-c}{b-c} = \sum_{k=0}^{a-c} (-1)^k \binom{a-c}{k} = (1-1)^{a-c} = \begin{cases} 1 & \text{si } a=c, \\ 0 & \text{sinon.} \end{cases}$$

Or, si $C \subset A$ et $a=c$, alors $A=C$, donc on conclut que $(MN)_{A,C}$ vaut 0 sauf si $A=C$. Ainsi, MN est bien la matrice identité.

7. Par la formule d'inclusion exclusion, on a pour toute partie A :

$$\log \left(\frac{\mathbb{P}[(0)_{v \in V}]}{\mathbb{P}[(x_v^A)_{v \in V}]} \right) = \sum_{B \subset A} \mathcal{V}_B((x_v)_{v \in B}).$$

En particulier, avec $A=V$ et en prenant l'exponentielle, on obtient la formule :

$$\mathbb{P}[(x_v)_{v \in V}] = \mathbb{P}[(0)_{v \in V}] \exp \left(- \sum_{A \subset V} \mathcal{V}_A((x_v)_{v \in A}) \right).$$

Il reste à montrer que les potentiels \mathcal{V}_A s'annulent si A n'est pas une clique. Fixons deux éléments $s \neq t$ dans A tels que $\{s, t\} \notin E$. On a pour commencer :

$$\begin{aligned} \mathcal{V}_A((x_v)_{v \in A}) &= \sum_{B|B \subset A} (-1)^{|A \setminus B|} \log \left(\frac{\mathbb{P}[(0)_{v \in V}]}{\mathbb{P}[(x_v^B)_{v \in V}]} \right) \\ &= \sum_{B|B \subset A, t \in B} (-1)^{|B \setminus A|} \log \left(\frac{\mathbb{P}[(0)_{v \in V}]}{\mathbb{P}[(x_v^B)_{v \in V}]} \right) + \sum_{B|B \subset A, t \notin B} (-1)^{|A \setminus B|} \log \left(\frac{\mathbb{P}[(0)_{v \in V}]}{\mathbb{P}[(x_v^B)_{v \in V}]} \right) \\ &= \sum_{B|B \subset A \setminus \{t\}} -(-1)^{|A \setminus B|} \log \left(\frac{\mathbb{P}[(0)_{v \in V}]}{\mathbb{P}[(x_v^{B \sqcup \{t\}})_{v \in V}]} \right) + \sum_{B|B \subset A \setminus \{t\}} (-1)^{|A \setminus B|} \log \left(\frac{\mathbb{P}[(0)_{v \in V}]}{\mathbb{P}[(x_v^B)_{v \in V}]} \right) \\ &= \sum_{B|B \subset A \setminus \{t\}} (-1)^{|A \setminus B|} \log \left(\frac{\mathbb{P}[(x_v^{B \sqcup \{t\}})_{v \in V}]}{\mathbb{P}[(x_v^B)_{v \in V}]} \right). \end{aligned}$$

Par conditionnement,

$$\begin{aligned} \mathbb{P}[(x_v^{B \sqcup \{t\}})_{v \in V}] &= \mathbb{P}[X_t = x_t \mid (X_w)_{w \notin B \sqcup t} = 0, (X_v)_{v \in B} = (x_v)_{v \in B}] \mathbb{P}[(X_w)_{w \notin B \sqcup t} = 0, (X_v)_{v \in B} = (x_v)_{v \in B}] \\ \mathbb{P}[(x_v^B)_{v \in V}] &= \mathbb{P}[X_t = 0 \mid (X_w)_{w \notin B \sqcup t} = 0, (X_v)_{v \in B} = (x_v)_{v \in B}] \mathbb{P}[(X_w)_{w \notin B \sqcup t} = 0, (X_v)_{v \in B} = (x_v)_{v \in B}], \end{aligned}$$

donc le rapport des deux probabilités vaut

$$\frac{\mathbb{P}[X_t = x_t \mid (X_w)_{w \notin B \sqcup t} = 0, (X_v)_{v \in B} = (x_v)_{v \in B}]}{\mathbb{P}[X_t = 0 \mid (X_w)_{w \notin B \sqcup t} = 0, (X_v)_{v \in B} = (x_v)_{v \in B}]}.$$

Or, puisque le champ est de Markov, ces probabilités ne dépendent que des valeurs au voisinage de t . Ainsi,

$$\frac{\mathbb{P}[(x_v^{B \sqcup \{t\}})_{v \in V}]}{\mathbb{P}[(x_v^B)_{v \in V}]} = \frac{\Pi(x_v, (x_w^B)_{w \in N_t})}{\Pi(0, (x_w^B)_{w \in N_t})}.$$

Séparons maintenant les parties $B \subset A \setminus \{t\}$ selon qu'elles contiennent s ou non. Remarquons que si B ne contient pas s , alors les rapports de lois locales

$$\frac{\Pi(x_v, (x_w^{B \sqcup \{s\}})_{w \in N_t})}{\Pi(0, (x_w^{B \sqcup \{s\}})_{w \in N_t})} \quad \text{et} \quad \frac{\Pi(x_v, (x_w^B)_{w \in N_t})}{\Pi(0, (x_w^B)_{w \in N_t})}$$

sont les mêmes, car s n'est pas dans N_t , et les variables apparaissant comme paramètres des lois locales ne contiennent pas x_s . On a donc :

$$\begin{aligned} \mathcal{V}_A((x_v)_{v \in A}) &= \sum_{B \mid B \subset A \setminus \{t\}} (-1)^{|A \setminus B|} \log \left(\frac{\Pi(x_v, (x_w^B)_{w \in N_t})}{\Pi(0, (x_w^B)_{w \in N_t})} \right) \\ &= \sum_{B \mid B \subset A \setminus \{s \sqcup t\}} (-1)^{|A \setminus B|} \log \left(\frac{\Pi(x_v, (x_w^B)_{w \in N_t})}{\Pi(0, (x_w^B)_{w \in N_t})} \right) \\ &\quad + \sum_{B \mid B \subset A \setminus \{s \sqcup t\}} -(-1)^{|A \setminus B|} \log \left(\frac{\Pi(x_v, (x_w^{B \sqcup \{s\}})_{w \in N_t})}{\Pi(0, (x_w^{B \sqcup \{s\}})_{w \in N_t})} \right) \\ &= 0 \end{aligned}$$

car les termes se compensent deux-à-deux. Ainsi, tout champ de Markov positif est le champ de Gibbs d'un certain potentiel.

8. Comme la loi π des variables ξ_n charge tous les sommets, on peut modifier à chaque étape n'importe quelle valeur X_v du champ. De plus, comme le champ est positif, la loi locale $\Pi(\cdot, \text{voisinage})$ charge toutes les valeurs possibles dans \mathfrak{X} , donc on peut modifier avec probabilité positive un sommet arbitraire en lui imposant une valeur arbitraire. En répétant cet argument pour chaque sommet, on peut passer en un nombre fini d'étapes de n'importe quel champ $(x_v)_{v \in V}$ à n'importe quel autre champ $(y_v)_{v \in V}$; ainsi, la chaîne est irréductible.

Montrons maintenant que la loi \mathbb{P} est réversible pour cette dynamique. Notons P la matrice de transition; par définition, $P((x_v)_{v \in V}, (y_v)_{v \in V}) = 0$ sauf si les deux champs $(x_v)_{v \in V}$ et $(y_v)_{v \in V}$ diffèrent exactement un site (ou sont égaux). Il suffit donc de montrer que, si $x_w = y_w$ pour tout $w \neq v$ avec v un site fixé, alors

$$\mathbb{P}[(x_v)_{v \in V}] P((x_v)_{v \in V}, (y_v)_{v \in V}) = \mathbb{P}[(y_v)_{v \in V}] P((y_v)_{v \in V}, (x_v)_{v \in V}).$$

Les deux probabilités de transition sont par définition données respectivement par

$$\pi(v) \Pi(y_v, (x_w)_{w \in N_v}) \quad \text{et} \quad \pi(v) \Pi(x_v, (x_w)_{w \in N_v}).$$

Or, par la propriété de Markov pour le champ, on a

$$\begin{aligned} \Pi(y_v, (x_w)_{w \in N_v}) &= \mathbb{P}[X_v = y_v \mid (X_w)_{w \in N_v} = (x_w)_{w \in N_v}] \\ &= \mathbb{P}[X_v = y_v \mid (X_w)_{w \in N_v} = (x_w)_{w \in N_v} \quad \text{et} \quad (X_w)_{w \in M_v} = (x_w)_{w \in M_v}] \\ &= \frac{\mathbb{P}[(X_v)_{v \in V} = (y_v)_{v \in V}]}{\mathbb{P}[(X_w)_{w \neq v} = (x_w)_{w \neq v}]} \end{aligned}$$

Donc, le terme de gauche de l'équation de réversibilité est

$$\pi(v) \frac{\mathbb{P}[(x_v)_{v \in V}] \mathbb{P}[(y_v)_{v \in V}]}{\mathbb{P}[(X_w)_{w \neq v} = (x_w)_{w \neq v}]}$$

et on obtient la même expression pour le terme de droite, car $(x_w)_{w \neq v} = (y_w)_{w \neq v}$. Ainsi, \mathbb{P} est une mesure de probabilité réversible et donc invariante pour la chaîne de Markov, ce qui implique la récurrence positive. Finalement, à chaque étape, on peut choisir un sommet $\xi_n = v$ et conserver la même valeur x_v avec probabilité positive. Donc, l'ensemble des temps de retour possibles contient 1 et la chaîne est apériodique. Par le théorème de convergence des chaînes de Markov récurrentes positives apériodiques, la loi de X_n tend vers la loi \mathbb{P} lorsque n tend vers l'infini : pour tout champ $(x_v)_{v \in V}$,

$$\lim_{n \rightarrow \infty} \mathbb{P}[(X_{n,v})_{v \in V} = (x_v)_{v \in V}] = \mathbb{P}[(x_v)_{v \in V}].$$

Pour simuler la loi \mathbb{P} du champ de Markov, on peut donc lancer la chaîne $(X_n)_{n \in \mathbb{N}}$ (à partir d'un état arbitraire, par exemple $(0)_{v \in V}$) et l'observer en temps grand.

9. Le programme suivant, écrit en Python, définit une classe `Field` qui permet la manipulation de configurations dans $\mathfrak{X}^V = \{\pm 1\}^{(\mathbb{Z}/d\mathbb{Z})^2}$.

```
import numpy as np
import numpy.random as rand
import matplotlib.pyplot as plt

class Field:

    def __init__(self, d):
        self.dim = d
        self.grid = 2*rand.randint(0, 2, size=(d, d)) - 1

    def __repr__(self):
        return f"Configuration_on_a_{self.dim}*{self.dim}_grid"

    def draw(self):
        fig, ax = plt.subplots(figsize = (10, 10))
        ax.set_axis_off()
        ax.set_aspect(1)
        for i in range(self.dim):
            for j in range(self.dim):
                if self.grid[i, j] == 1:
                    ax.add_patch(plt.Rectangle((i, j), 1, 1))
        ax.set_xlim(0, self.dim)
        ax.set_ylim(0, self.dim)
        plt.show()

    def magnetisation(self):
        return np.sum(self.grid) / (self.dim * self.dim)

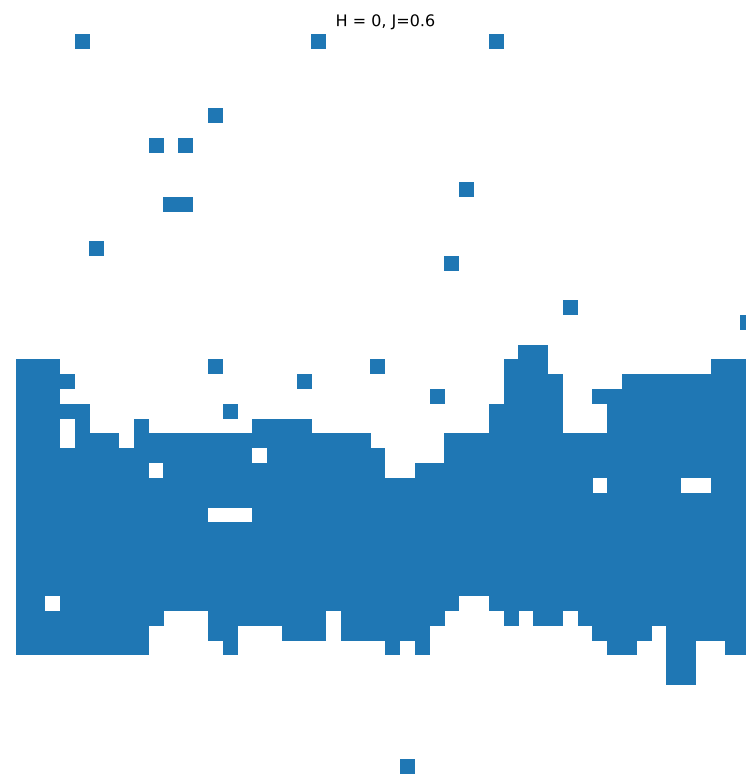
    def update(self, site, H, J):
        x = self.grid
        d = self.dim
        i, j = site
        theta = H + J * (x[(i+1)%d, j] + x[(i-1)%d, j]
            + x[i, (j+1)%d] + x[i, (j-1)%d])
        threshold = np.exp(theta) / (np.exp(theta) + np.exp(-theta))
        val = 2*(rand.random() <= threshold) - 1
        self.grid[i, j] = val

    def periodic_update(self, H, J, n=1):
        for _ in range(n):
            for i in range(self.dim):
                for j in range(self.dim):
                    self.update((i, j), H, J)

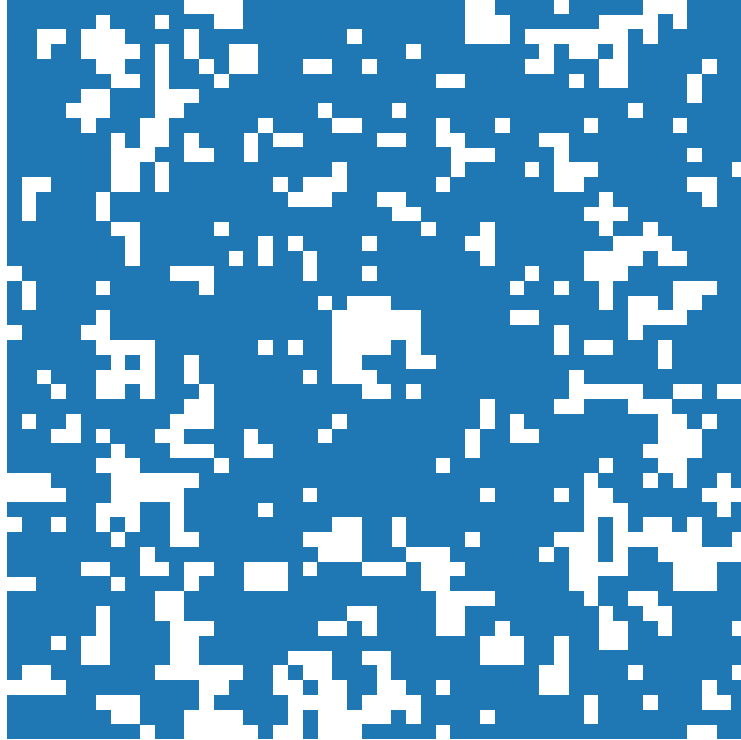
def GibbsSampler(d, n, H, J):
```

```
res = Field(d)
res.periodic_update(H, J, n)
return res
```

On a dessiné ci-dessous quatre simulations avec $d = 50$ (sur un tore) et $n = 1000$ mises à jour périodiques.



$H = 0.1, J=0.3$



$H = 0.1, J=0.6$



Plus le paramètre J est grand, et plus il y a tendance à avoir de grandes zones toutes de même valeur $+1$ ou -1 . En l'absence de champ magnétique ($H = 0$), les deux valeurs ont la même probabilité, et des régions positives et négatives apparaissent en même proportion. Si $H > 0$, alors les zones positives sont privilégiées, et pour J assez grand et $H > 0$, ceci force quasiment toute la grille à être positive. On aurait un comportement opposé avec $H < 0$.

Mettons finalement en évidence le phénomène de magnétisation spontanée. Avec une grille de taille

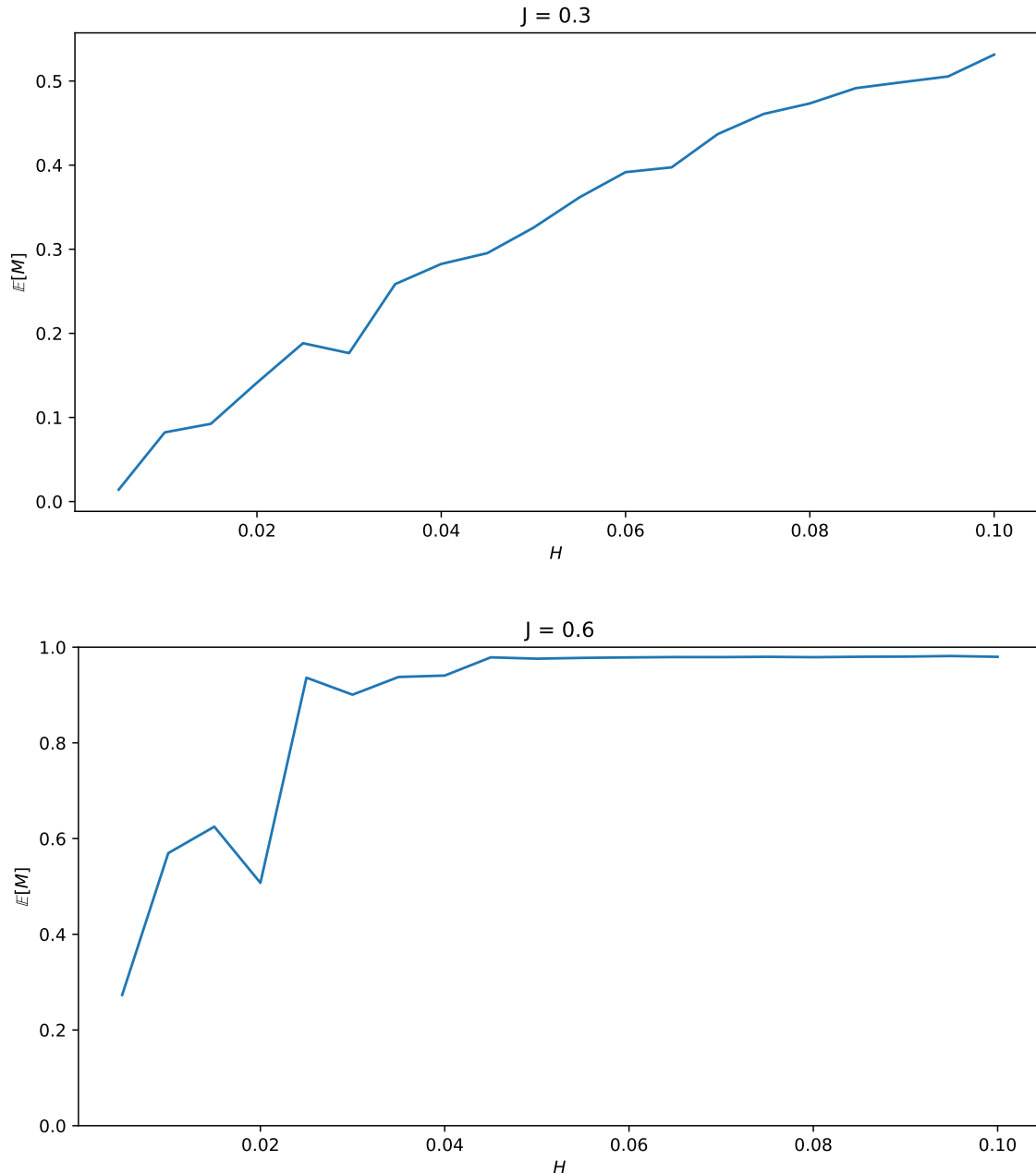
$d = 30$, et pour 20 valeurs de H entre 0 et 0.1, on calcule les moyennes empiriques sur 50 essais de la magnétisation :

```

J = 0.3
Hlist = np.arange(0.005, 0.105, 0.005)
Mlist = np.zeros(20)
for i in range(20):
    Mlist2 = np.zeros(50)
    for j in range(50):
        Mlist2[j] = GibbsSampler(30, 1000, Hlist[i], J).magnetisation()
    Mlist[i] = np.mean(Mlist2)

```

Avec les valeurs $J = 0.3$ et $J = 0.6$, on obtient les graphes suivants de $\mathbb{E}[M]$ en fonction de H :



Les irrégularités des graphes sont liées à l'approximation des espérances par des moyennes empiriques. Par ailleurs, on considère à la fois que $d = 30$ est une dimension assez grande pour avoir une approximation de $\lim_{d \rightarrow \infty} \mathbb{E}_{H,J}[M]$, et aussi que 1000 pas de l'échantillonneur de Gibbs périodique produise une approximation assez bonne de la loi d'Ising. Quoiqu'il en soit :

- pour $J = 0.3$, il semble que $\lim_{H \rightarrow 0^+} (\lim_{d \rightarrow \infty} \mathbb{E}_{H,J}[M]) = 0$: lorsqu'on diminue la valeur du champ magnétique H , la magnétisation moyenne $\mathbb{E}[M]$ diminue également, et elle tend vers 0 si H tend vers 0.
- en revanche, pour $J = 0.6$, il semble que $\lim_{H \rightarrow 0^+} (\lim_{d \rightarrow \infty} \mathbb{E}_{H,J}[M]) > 0$: lorsque la valeur du champ magnétique H tend vers 0, il reste une magnétisation moyenne résiduelle $\mathbb{E}[M]$ qui ne tend pas vers 0. Notons également que $\mathbb{E}[M]$ est très proche de la valeur maximale 1 même pour des petites valeurs de H .

On a ainsi illustré le phénomène de magnétisation spontanée pour des valeurs de J assez grandes.